

# **Interrater-Reliabilität bei der Beurteilung der Körpersprache nach dem Fasziendistorsionsmodell (FDM)**

Master Thesis zur Erlangung des Grades  
Master of Science in Osteopathie

an der **Donau Universität Krems –  
Zentrum für chin. Medizin &  
Komplementärmedizin**

niedergelegt  
an der **Wiener Schule für Osteopathie**

von **Stefan Anker**  
Wien, Juni 2011

Betreuung: Mag<sup>a</sup>. Claudia Gamsjäger

## **Eidesstattliche Erklärung**

Hiermit versichere ich, die vorgelegte Masterthese selbständig verfasst zu haben.

Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten oder nicht veröffentlichten Arbeiten anderer übernommen wurden, wurden als solche gekennzeichnet. Sämtliche Quellen und Hilfsmittel, die ich für die Arbeit genützt habe, sind angegeben. Die Arbeit hat mit gleichem Inhalt weder im In- noch Ausland einer anderen Prüfungsbehörde vorgelegen.

Diese Arbeit stimmt mit der von dem/der Gutachter/in beurteilten Arbeit überein.

Wien, 1. Juni 2011 \_\_\_\_\_

Stefan Anker

## **Interrater-Reliabilität bei der Beurteilung der Körpersprache nach dem Faszindistorsionsmodell (FDM)**

Schlüsselbegriffe: Faszindistorsionsmodell, Osteopathie, Interrater-Reliabilität, Körpersprache,

Das Faszindistorsionsmodell (FDM) ist ein in der Osteopathie angewendetes, anatomisches Modell, welches die Ursache von körperlichen Beschwerden in pathologischen Verformungen von Faszie sucht. Diese Faszindistorsionen können durch ein spezielles Befundungssystem diagnostiziert werden. Ein prägnantes, vom osteopathischen Ansatz der manuellen Befundung abweichendes Kriterium stellt dabei die visuelle Diagnostik der Körpersprache des Patienten dar. Reliabilitätsuntersuchungen manueller Testmethoden werden in Review-Artikeln kritisch hinsichtlich ihrer Ergebnisse und Methodologie beurteilt. In dieser Studie wird im Gegensatz dazu ein visueller Diagnoseansatz bezüglich seiner Interrater-Reliabilität untersucht. Zusätzlich werden Einflussfaktoren wie der Quellberuf der Osteopathen, ihre Erfahrungsdauer mit dem FDM und zusätzliche verbale Aussagen der gefilmten Patienten auf die Reliabilität erforscht.

Dafür werden 25 Osteopathen aus Deutschland, Österreich und der Schweiz mittels eines Online-Fragebogens, Patientenvideos vorgeführt. Auf diesen Videos beschreiben Patienten vorwiegend nonverbal ihre körperlichen Beschwerden. Die Tester diagnostizieren anhand der FDM-Richtlinien Faszindistorsionen auf diesem Bildmaterial. Die Ergebnisse werden statistisch anhand des Kappa-Index nach Siegel und Castellan ausgewertet und zeigen akzeptable Werte in der Verlässlichkeit ( $k=0,51$ ). Dabei sind die Ergebnisse von Osteopathen mit dem Quellberuf Arzt oder Heilpraktiker signifikant besser als die Ergebnisse von Osteopathen mit dem Quellberuf Physiotherapeut. Die Dauer der Erfahrung der Tester mit dem FDM beeinflusst die Reliabilität nicht. Verbale Aussagen, die von Patienten parallel zur nonverbalen Kommunikation getätigt werden, haben einen signifikanten Einfluss auf die Reliabilität, wobei diese Ergebnisse durch die verwendete Methodologie nur als Ansatzpunkt für weitere Forschung gesehen werden können. Vor dem Hintergrund fehlender Aussagen über die Validität kann in dieser Studie von einer akzeptablen Interrater-Reliabilität der visuellen FDM-Diagnose gesprochen werden.

Stefan Anker

## **Interrater reliability in evaluating the body language based on the Fascial Distortion Model (FDM)**

Key words: Fascial distortion model, osteopathy, interrater reliability, body language

The Fascial Distortion Model (FDM) is an anatomical model applied in osteopathy which seeks the cause of physical symptoms in pathological deformations of fascia. These fascial distortions can be evaluated with a special system of diagnosis. The visual utilisation of the patient's body language is concisely different to the common osteopathic approach through manual diagnosis. Review articles about the reliability of manual test methods are very critical regarding their findings and methodology. In contrast this study examines a visual diagnostic approach in terms of its interrater reliability. In addition, such factors as the professional background of osteopaths, the experience level with the FDM and additional verbal statements of patients are tested concerning their influence on reliability.

Patient videos were presented via online questionnaire to 25 osteopaths from Germany, Austria and Switzerland. On these most patients describe their physical complaints nonverbally. The testers diagnose fascial distortions based on the FDM guidelines. The results are statistically analyzed using the kappa index according to Siegel and Castellan, and show acceptable levels of reliability ( $k = 0.51$ ). Here, the results of osteopaths with a medical profession or non medical practicion as a background are significantly better than the results of osteopaths who are physical therapist by trade. The duration of experience of the tester with the FDM does not affect the reliability. Verbal statements that are made by patients parallel to the non-verbal communication, have a significant impact on the reliability. These results can only be seen as an encouragement for further research. Seen against the background of a lack of information about validity, this study presents an acceptable interrater reliability of visual FDM diagnosis.

## **DANKE ...**

Mag<sup>a</sup>. Claudia Gamsjäger für die konstruktive Kritik  
und ihren Einsatz für das Gelingen dieser Masterthese

Mag<sup>a</sup>. Cornelia Pollany, Michael Vogt und Matthias  
von Ehrenstein für ihre Hilfe im Finale dieser These

den zahlreichen Kollegen aus der European FDM Association,  
die mich bei der Durchführung dieser Studie unterstützt haben

Melinda für die Unterstützung während der  
Arbeit an dieser Masterthese und darüber hinaus.

# Inhaltsverzeichnis

- 1. Einleitung 9**
  - 1.1. Aufbau der Masterthese
  - 1.2. Hinweise für Leserinnen und Leser
  
- 2. Grundlagen 12**
  - 2.1. Gütekriterien diagnostischer Testverfahren
    - 2.1.1. Validität
      - 2.1.1.1. Interne Validität
      - 2.1.1.2. Externe Validität
    - 2.1.2. Sensitivität und Spezifität
    - 2.1.3. Positiver und negativer Vorhersagewert
    - 2.1.4. Reliabilität
      - 2.1.4.1. Interrater- und Intrarater-Reliabilität
  - 2.2. Bedeutung von Reliabilitätsuntersuchungen für die Osteopathie
  - 2.3. Untersuchungen im Bereich manueller Tests an der Wirbelsäule
    - 2.3.1. Kritik und Kriterien für verbesserte Reliabilität
  
- 3. Zur Bedeutung von Gestik und Körpersprache 21**
  - 3.1. Definitionen
    - 3.1.1. Körpersprache
    - 3.1.2. Gestik
  - 3.2. Charakteristik von Gesten
    - 3.2.1. Gestenkategorien
    - 3.2.2. Temporaler Ablauf und Inhalt von Gestik
  - 3.3. Forschung im Bereich Gestik und medizinischer Kommunikation
  - 3.4. Gestik und Körpersprache in der Osteopathie
  - 3.5. Gestik und Körpersprache als reliables Diagnosekriterium
    - 3.5.1. Untersuchung aus dem Bereich der Kardiologie
  - 3.6. Zusammenfassung und Konsequenzen für diese Masterthese
  
- 4. Das Faziendistorsionsmodell (FDM) nach Stephen Typaldos 35**
  - 4.1. Entwicklungsgeschichte
    - 4.1.1. Publikationen
  - 4.2. Modellvorstellung im FDM
    - 4.2.1. Faszie aus Sicht des FDM
  - 4.3. Typologie der Faziendistorsionen
    - 4.3.1. Triggerbänder (TB)

- 4.3.2. Hernierte Triggerpunkte (HTP)
- 4.3.3. Continuumdistorsionen (CD)
- 4.3.4. Faltdistorsionen (FD)
- 4.3.5. Zylinderdistorsionen (CyD)
- 4.3.6. Tektonische Fixierungen (TF)
- 4.4. Die FDM-Diagnose

## **5. Fragestellung 51**

## **6. Methodologie 53**

- 6.1. Erhebungsmaterial
  - 6.1.1. Videoaufnahmen
    - 6.1.1.1. Aufnahmesituation
    - 6.1.1.2. Auswahl der gefilmten Patienten
    - 6.1.1.3. Nachbearbeitung der Videos
    - 6.1.1.4. Auswahl der verwendeten Videos für den Fragebogen
  - 6.1.2. Onlinefragebogen
    - 6.1.2.1. Abwicklung der Befragung
    - 6.1.2.2. Aufbau des Online-Fragebogens
- 6.2. Tester
  - 6.2.1. Auswahl der Tester
  - 6.2.2. Demografische Angaben
    - 6.2.2.1. Geschlecht
    - 6.2.2.2. Quellberuf
    - 6.2.2.3. Erfahrung mit dem FDM
  - 6.2.3. Rücklaufquote und Untersuchungsablauf
- 6.3. Statistik
  - 6.3.1. Statische Auswertung der Daten
    - 6.3.1.1. Der Kappa-Index (k-Index) nach Siegel und Castellan (1988)
    - 6.3.1.2. Prüfung der teststatistischen Voraussetzungen
    - 6.3.1.3. Zur Betrachtung der Untergruppen bei den Testern
    - 6.3.1.4. Zur qualitativen Betrachtung der Ergebnisse anhand der einzelnen Videos

## **7. Ergebnisse 70**

- 7.1. Interrater-Reliabilität der Osteopathen
  - 7.1.1. Interrater-Reliabilität und Quellberuf
    - 7.1.1.1. Interrater-Reliabilität der Ärzte
    - 7.1.1.2. Interrater-Reliabilität der Physiotherapeuten
    - 7.1.1.3. Interrater-Reliabilität der Heilpraktiker

- 7.1.1.4. Vergleich der Interrater-Reliabilität bei unterschiedlichen Quellberufen
- 7.1.2. Interrater-Reliabilität und Erfahrung mit dem FDM
  - 7.1.2.1. Interrater-Reliabilität bei mindestens drei Jahren an Erfahrung
  - 7.1.2.2. Interrater-Reliabilität bei unter drei Jahren an Erfahrung
  - 7.1.2.3. Interrater-Reliabilität im Vergleich bei Osteopathen mit weniger als drei und mit mehr als zwei Jahren an Erfahrung mit dem FDM
  - 7.1.2.4. Interrater-Reliabilität mit mehr als drei Jahren an Erfahrung
  - 7.1.2.5. Interrater-Reliabilität mit drei und weniger Jahren an Erfahrung
  - 7.1.2.6. Interrater-Reliabilität im Vergleich bei Osteopathen mit weniger als vier und mit mehr als drei Jahren an Erfahrung mit dem FDM
- 7.1.3. Einfluss der verbalen Aussagen der Patienten auf die Interrater-Reliabilität
  - 7.1.3.1. Interrater-Reliabilität bei Videomaterial ohne Ton
  - 7.1.3.2. Interrater-Reliabilität bei Videomaterial mit Ton
  - 7.1.3.3. Interrater-Reliabilität im Vergleich bei Videomaterial mit und ohne Ton
- 7.1.4. Qualitative Analyse der Ergebnisse anhand der einzelnen Videos

## **8. Diskussion 82**

- 8.1. Zusammenfassung der Ergebnisse
- 8.2. Diskussion der Ergebnisse
- 8.3. Methodenreflexion

## **9. Zusammenfassung und Ausblick 90**

## **10. Literaturverzeichnis 91**

## **11. Abbildungsverzeichnis 97**

## **12. Tabellenverzeichnis 98**

## **13. Anhang 101**

## **14. Englische Kurzzusammenfassung 111**



# 1. Einleitung

Wenn Patienten zum Osteopathen kommen und ihre Beschwerden beschreiben, findet das auf verschiedenen Kommunikationsebenen statt. Patienten sprechen über ihre körperlichen Beschwerden, sie zeigen Bewegungen vor, welche sie unzureichend oder nur unter Schmerzen durchführen können und sie gestikulieren mit ihren Händen und zeigen an ihrem Körper, wo die Beschwerden lokalisiert sind. Dabei deuten sie Punkte und Linien am Ort der Schmerzen an, drücken Gelenke mit den Händen, kneten Weichteile oder bohren ihre Finger zwischen Muskeln.

Welchen Einfluss haben diese Informationen auf den osteopathischen Befund? Können diese Informationen bei der Diagnosestellung entscheidend weiterhelfen? Oder sind sie im Gegensatz zum zentralen Element der osteopathischen Diagnose, der Palpation, kaum von Bedeutung?

Stephen Typaldos stellte sich als Osteopath und Notfallmediziner ähnliche Fragen. Das Resultat seiner Forschung war das Faszienmodell (FDM), welches eine systematische Sammlung subjektiver Patienteninformationen als Leitlinie für die Diagnose und Therapie von körperlichen Beschwerden enthält.

Das FDM ist ein anatomisches Modell, welches primär in Verformungen von Faszie die Ursache dieser Beschwerden sieht. Dem FDM eigen ist ein spezieller Befundungsprozess, der sich stark an der subjektiven verbalen und nonverbalen Kommunikation des Patienten orientiert. Die meisten Informationen, die Patienten während der osteopathischen Behandlung äußern, haben im FDM Bedeutung und führen unter anderem zu einer Diagnose von einer oder mehrerer ursächlichen Faszienstörungen. Diese werden als pathologische Veränderungen von Faszie beschrieben (Typaldos 1999, 2002) und können prinzipiell (derzeit meist manuell) korrigiert werden. Diese Korrektur führt nach Typaldos zu sofortigen und messbaren Resultaten in der Therapie (Typaldos 1999, 2002).

Ein Teil der Diagnose im FDM beruht auf einer Art visueller Diagnostik, da nach Typaldos (1999, 2002) Patienten Faszienstörungen unter anderem durch Zeigen am eigenen Körper andeuten. Das FDM stellt somit die Körpersprache des Patienten in den Mittelpunkt seiner Diagnostik. Diese Vorgangsweise deckt sich nicht mit der klassisch angewandten osteopathischen Basisuntersuchung, welche im Gegensatz dazu auf manuellen Testverfahren beruht.

In der mehrjährigen Osteopathieausbildung an der Wiener Schule für Osteopathie (WSO) lernt man eine große Zahl dieser manuellen Diagnoseverfahren kennen, welche teilweise in wissenschaftlichen Untersuchungen auf ihre Inter- und Intra-rater-Reliabilität hin untersucht werden. Reliable Tests sind als ein Grundpfeiler der *evidence based medicine*, auch in der Osteopathie von großem Interesse. Sie sind Voraussetzung für adäquat ausgewählte und nachvollziehbare Therapieverfahren und deren Prognose (Johnston 1982, Sackett et al. 1996, Lucas und Bogduk 2011).

Diese Reliabilitätsstudien zeigen jedoch mehr oder weniger akzeptable Resultate, wie Review-Artikel<sup>1</sup> im Bereich manueller Testverfahren im Bereich der Wirbelsäule aufzeigen (Seffinger et al. 2004, Hollerwöger 2006, Stochkendahl et al. 2006, Haneline und Young 2009). Der visuelle Ansatz der FDM-Diagnostik wurde derartigen Untersuchungen bisweilen noch nicht unterzogen und es stellt sich die Frage, ob die Diagnose der FDM-Körpersprache ein reliabler Parameter in der Diagnose von Faziendistorsionen ist.

Was bringt dieser alternative Diagnoseansatz aber für die Osteopathie in der Praxis? Ist das FDM prinzipiell in die Osteopathie integrierbar?

Faktisch scheint das FDM durchaus in der Osteopathie angekommen zu sein. Es wird an mehreren Osteopathieschulen in Europa<sup>2</sup> gelehrt und von vielen Osteopathen erfolgreich in der Praxis eingesetzt.

Modellhafte Vorstellungen waren und sind Teil der Osteopathie und der Medizin. Ähnlich wie die zu seiner Zeit innovativen Überlegungen von Sutherland die Osteopathie im kranialen Bereich beeinflusste, wollte Typaldos mit seinem Modell den Grundlagen der Medizin (und damit in weiten Teilen auch der Osteopathie) einen neuen Ansatz gegenüberstellen, der im Gegensatz zu Gewebeschädigung mit den damit verbundenen physiologischen Adaptions- und Reparaturvorgängen als Ursache für körperliche Beschwerden und Schmerzen, die anatomische Form der Faszie und ihre Auswirkung auf Struktur und Funktion in den Vordergrund rückt.

---

<sup>1</sup> *Reviews fassen den aktuellen Forschungsstand zu einem bestimmten Thema zusammen. Enthaltene Studien werden nach definierten Kriterien ausgewählt und bewertet: Ziel ist es, die bestmögliche Beweislage für eine wissenschaftliche Fragestellung zu produzieren (Kleibel und Mayer 2005)*

<sup>2</sup> *Das Faziendistorsionsmodell (FDM) ist Teil der Osteopathie-Ausbildung am Deutschen Fortbildungsinstitut für Osteopathie (DFO), an der Osteopathieschule Deutschland (OSD) und an der Wiener Schule für Osteopathie (WSO).*

Diese Masterthese beschäftigt sich mit dem FDM als Modell, welches eine anatomische Vorstellung von körperlichen Beschwerden und Funktionseinschränkungen anbietet. Dieser Ansatz ist für Osteopathen interessant, weil es eine weitere Möglichkeit eröffnet, Patienten gezielt und effektiv zu behandeln. Zudem beschreibt Typaldos (1999) den Vorteil, durch diese alternative Vorstellung, konkret angewendete osteopathische Techniken hinsichtlich ihrer anatomischen Wirkung auf ein Beschwerdebild zu überprüfen und damit osteopathische Maßnahmen gezielter einsetzbar und deren Effekt vorhersehbar zu machen.

Entscheidend für jegliche Anwendung des FDM in der Praxis ist eine adäquate Diagnostik im Sinne des Modells. Deshalb untersucht diese Masterthese erstmalig die Interrater-Reliabilität der FDM-Körpersprache, um zu erforschen, inwieweit dieser visuelle Diagnoseansatz übereinstimmende Ergebnisse liefern kann. Dazu werden Osteopathen in Deutschland, Österreich und der Schweiz mittels eines videounterstützten Online-Fragebogens befragt. Als Basis dafür dienen Videos, die Patienten bei der Beschreibung ihrer Beschwerden zeigen. Anhand dieser Videos sollen die Tester beurteilen, welche Faszienstörungen sie erkennen. Diese Ergebnisse werden mit Hilfe einer Kappa-Statistik<sup>1</sup> ausgewertet.

Als Grundlage für die Ergebnisdiskussion werden die Probleme von Reliabilitätsuntersuchungen im Bereich manueller Testverfahren, Ergebnisse aus der Gestik- und Kommunikationsforschung und die inhaltlichen Vorgaben des Faszienstörungsmodells vorgestellt.

### **1.1. Aufbau der Masterthese**

Im zweiten Kapitel über die Grundlagen von Reliabilitätsuntersuchungen werden die Gütekriterien derartiger Studien geklärt. Im Weiteren wird auf den Stellenwert von Reliabilitätsuntersuchungen in der Osteopathie eingegangen. Es werden Reviews zum Thema Reliabilität manueller Tests an der Wirbelsäule, deren wissenschaftliche Erkenntnisse und die Problemfelder derartiger Reliabilitätsuntersuchungen präsentiert.

Das dritte Kapitel widmet sich dem Thema Gestik und Körpersprache. Es soll geklärt werden, ob sich FDM-Körpersprache unter dem Begriff Gestik einordnen lässt und somit Erkenntnisse aus der Gestikforschung in diese Masterthese miteinbezogen werden können. Weiters werden Reliabilitätsstudien zum Thema Gestik als

---

<sup>1</sup> Vgl. Kap. 6.3.1.

Diagnosekriterium in der Medizin und Ergebnisse der Gestikforschung im Bereich der medizinischen Kommunikation vorgestellt.

Das Faszien-distorsionsmodell (FDM) ist Schwerpunkt des vierten Kapitels. Geschichtliche Entwicklung, Modellvorstellung des FDM, der Befundungsprozess und die spezifische Körpersprache der einzelnen Faszien-distorsionen sind Inhalt und Basis für die wissenschaftliche Untersuchung dieser Masterthese.

Das fünfte Kapitel beschreibt die Forschungsfrage und Null-Hypothesen.

Kapitel sechs widmet sich der Methodologie, dem Forschungsdesign, der Datenerhebung, dem Studienverlauf und der Datenverarbeitung.

In Kapitel sieben findet sich die deskriptive Statistik mit allen Ergebnissen.

Kapitel acht diskutiert die Ergebnisse dieser Masterthese in Zusammenhang mit den Erkenntnissen aus den vorangegangenen Kapiteln, befasst sich mit Problemfeldern dieser Untersuchung und gibt Ausblick auf resultierende Forschungsfragen für die Zukunft.

## **1.2. Hinweise für Leserinnen und Leser**

Im Rahmen dieser Arbeit wird zur besseren Lesbarkeit auf eine geschlechtsspezifische Differenzierung, zB OsteopathInnen, verzichtet. Selbstverständlich gelten entsprechende Begriffe im Sinne der Gleichbehandlung für beide Geschlechter.

## **2. Grundlagen**

Ziel dieses Kapitels ist es, die Gütekriterien diagnostischer Testverfahren zu erklären und damit eine Basis für das weitere Verständnis dieser Arbeit zu schaffen. Weiters wird der Stellenwert von Reliabilitätsstudien in der Osteopathie beleuchtet. Obwohl auch andere osteopathische Diagnosemethoden in Reliabilitätsstudien untersucht wurden (zB Hartman und Norton 2002, Sommerfeld et al. 2004, Liem 2005), konzentrieren sich die vorgelegten Untersuchungen auf das Thema Reliabilität manueller Tests an der Wirbelsäule. Hierzu liegen mehrere Review-Artikel vor. Diese Studien hinterfragen die Qualität der verarbeiteten wissenschaftlichen Untersuchungen und sind deshalb eine gute Basis, die Schwierigkeiten bei der Untersuchung manueller Diagnoseverfahren darzustellen. Die daraus resultierenden

Anforderungen an qualitativ hochwertige Reliabilitätsstudien sollen weitest gehend in die Methodologie dieser Masterthese einfließen. Die Basis liefert eine Literaturrecherche in den Onlinedatenbanken OSTMED-DR, Osteopathic Research Web und PubMed (vgl. Tab. 1) mit den Begriffen *reliability* OR *review* AND *osteopath\** OR *(manual) test\** OR *diagnos\**. Wenn die kompletten Artikel nicht zur Verfügung stehen, wurden sie per Email von den Autoren angefordert.

OSTMED-DR	<a href="http://www.ostmed-dr.com:8080/vital/access/manager/Index">http://www.ostmed-dr.com:8080/vital/access/manager/Index</a>
Osteopathic Research Web	<a href="http://www.osteopathic-research.com/">http://www.osteopathic-research.com/</a>
PubMed	<a href="http://www.ncbi.nlm.nih.gov/PubMed">http://www.ncbi.nlm.nih.gov/PubMed</a>

Tab. 1: Homepage-Adressen der in dieser Masterthese verwendeten Online-Datenbanken

## 2.1. Gütekriterien diagnostischer Testverfahren

Diese beziehen sich prinzipiell auf die Validität und Reliabilität eines Tests (Weiß 2008).

### 2.1.1. Validität

Validität sagt aus, inwieweit ein Test das misst, was er messen soll. Voraussetzung dafür ist, dass das zu messende Phänomen klar, auch durch alternative Messmethoden nachzuweisen ist (Sommerfeld 2006). Validität ist davon bestimmt, inwieweit ein Test richtig-positive Ergebnisse (entspricht der Sensitivität) bzw. richtig-negative Ergebnisse (entspricht der Spezifität) liefern kann. In der Praxis entscheidend ist der Vorhersagewert eines Tests, der die Wahrscheinlichkeit ausdrückt, dass ein Test den richtigen Status anzeigt und in Zusammenhang mit der Prävalenz<sup>1</sup> steht (Weiß 2008).

#### 2.1.1.1. Interne Validität

Interne Validität gibt Auskunft darüber, inwieweit die Ergebnisse und die daraus gezogenen Schlussfolgerungen für die Studienpopulation korrekt sind und ist deshalb absolut notwendig für die innere Stringenz einer Untersuchung. Ob die Studie sinnvoll ist, wird damit nicht beantwortet (Weiß 2008).

<sup>1</sup> Prävalenz wird beschrieben als: „Rate der zu einem bestimmten Zeitpunkt oder in einem bestimmten Zeitabschnitt an einer bestimmten Krankheit Erkrankten (im Vergleich zur Zahl der Untersuchten)“ (Duden-Online 2011)

### **2.1.1.2. Externe Validität**

Externe Validität beschreibt die Generalisierbarkeit der Studie und die Möglichkeit, Studienergebnisse in die Praxis zu übertragen (Weiß 2008).

### **2.1.2. Sensitivität und Spezifität**

Unter Sensitivität versteht sich die Fähigkeit eines Tests, ein vorhandenes Merkmal zu erkennen. Im Gegensatz dazu gibt die Spezifität Auskunft darüber, ob ein Test ein nicht vorhandenes Merkmal als nicht vorhanden anzeigt (Sommerfeld 2006).

Ein diagnostischer Test kann trotz hoher Sensitivität und Spezifität in der Praxis nur in Zusammenhang mit der Prävalenz beurteilt werden. Diese wiederum beschreibt statistische Zusammenhänge beispielsweise zwischen dem Auftreten einer Erkrankung und einem Risikofaktor, wobei sie kein Maß für das Risiko selbst ist. So kann ein richtig-positives Testergebnis bei einem Patienten mit einer niederen Prävalenz anders zu beurteilen sein als bei hoher Prävalenz (Weiß 2008).

Ausserdem gilt es zu berücksichtigen, ob das Patientenspektrum der Testpopulation mit dem in der Praxis vorgefundenen Spektrum übereinstimmt. Auch das könnte einen Test trotz hoher Sensitivität und Spezifität ungeeignet erscheinen lassen. Ein weiteres Problem könnte die Erwartungshaltung des Untersuchers (des Osteopathen) in der Praxis sein, der eine Diagnose bestätigen bzw. verwerfen möchte (sog. Informationsbias). Trotz hoher Sensitivität und Spezifität, reicht ein diagnostischer Test nicht aus, eine Diagnose festzulegen (Weiß 2008).

### **2.1.3. Positiver und negativer Vorhersagewert**

Dieser Wert beschreibt die Wahrscheinlichkeit, mit der ein Testergebnis den richtigen Status anzeigt. Während die Prävalenz die Wahrscheinlichkeit einer Erkrankung vor der Testung anzeigt, gibt der Vorhersagewert Auskunft über die Wahrscheinlichkeit, nachdem der Test durchgeführt wurde. Der positive Vorhersagewert gibt Information über die Wahrscheinlichkeit, dass ein positives Testergebnis tatsächlich die Erkrankung anzeigt und der negative Vorhersagewert die Wahrscheinlichkeit, dass ein negatives Testergebnis tatsächlich keine Erkrankung anzeigt (Weiß 2008).

### **2.1.4. Reliabilität**

Reliabilität (Verlässlichkeit) beschreibt die Reproduzierbarkeit und Konsistenz eines Tests unter gleichen bzw. ähnlichen Bedingungen (Sommerfeld 2006). Unter

Reproduzierbarkeit wird ein gleichbleibendes Ergebnis bei wiederholter Durchführung des Tests verstanden. Eine perfekte Reliabilität ist bei der Testung an lebenden Personen nahezu unmöglich. Reliabilität sollte sich zudem an einem unabhängigen Referenzstandard messen (Sommerfeld 2006). Beispielsweise könnte für einen manuellen Test, der eine Affektion des Discus intervertebralis am Patienten feststellen soll, als Referenzstandard eine MRT-Untersuchung herangezogen werden. Für einige osteopathische Tests gibt es jedoch keinen unabhängigen Referenzstandard und damit keine Auskunft über ihre Validität. Dieses Fehlen wirkt sich negativ auf die Reproduzierbarkeit und folglich auf die Reliabilität aus (Fritz und Wainner 2001).

#### **2.1.4.1. Interrater- und Intrarater-Reliabilität**

Interrater-Reliabilität gibt Auskunft über die Stabilität von Daten, die von mehreren Testern mit demselben Test an derselben Testperson erhoben werden. Dahingegen untersucht die Intrarater-Reliabilität die Stabilität der Daten, die derselbe Tester an derselben Testperson bei mehrmaliger Testung produziert (Sommerfeld 2006).

## **2.2. Bedeutung von Reliabilitätsuntersuchungen für die Osteopathie**

Die Reliabilität diagnostischer Tests hat einen hohen Stellenwert in der Osteopathie. Erst mit einer verlässlichen Diagnose kann die passende Behandlung für einen Patienten gefunden bzw. festgestellt werden, ob der Patient prinzipiell auf eine Behandlung anspricht. Reliable Tests sind ebenfalls Voraussetzung für die Prognosestellung und die weitere Verlaufsbetreuung des Patienten (Lucas und Bogduk 2011).

Individuelle klinische Expertise mitsamt der Erfahrung und klinischen Praxis des Arztes (hier Osteopathen), die Berücksichtigung der Ansprüche des Patienten und die beste verfügbare externe Evidenz, sowohl diagnostischer Verfahren als auch rehabilitativer und präventiver Maßnahmen sind die drei Grundpfeiler der *Evidence based Medicine (EbM)*. Grundlagenforschung und Ergebnisse klinischer Forschung werden herangezogen, um die beste externe Evidenz zu gewährleisten. Das kann dazu führen, dass bisher als reliabel geltende Diagnoseverfahren überdacht, verändert oder ersetzt werden (Sackett et al. 1996).

Da sich die Diagnosestellung in der Osteopathie vorwiegend auf manuelle Tests und die Palpation am Patienten stützt, sind Reliabilitätsstudien in diesem Bereich



von großem Interesse. Zugespielt formuliert rühren derartige Reliabilitätsstudien an den Grundfesten der Osteopathie (Johnston 1982).

Johnston (1982) beschreibt, dass durch die Arbeiten von Denslow und Korr über spinale Reflexe das Modell der osteopathischen Läsion in den 40er Jahren des vorigen Jahrhunderts auf ein neues, neurologisches Fundament gestellt wurden. In der Folge intensivierte sich das Interesse an exakter segmentaler Palpation dysfunktionaler Wirbelsäulensegmente und deren Verlässlichkeit bei der Testdurchführung durch mehrere Untersucher. Osteopathische Tests hatten und haben das Ziel, klare Informationen über das Patientenproblem zu gewinnen und einen logischen Handlungshintergrund für die Anwendung manipulativer Techniken zu bieten. Interrater-Reliabilitätsuntersuchungen sind Basis für ein Testprotokoll (im Fall der Osteopathie sind das manuelle Tests), mit welchem Dysfunktionen möglichst eindeutig, exakt und übereinstimmend manuell diagnostiziert werden können.

Die Ergebnisse von Reliabilitätsstudien im Bereich manueller Tests sollten jedoch nicht zu einer Einteilung in „gute“ und „schlechte“ Tests führen. Ein „schlechter“ Test kann in einer bestimmten Therapiesituation, trotz aller Vorbehalte seine Reliabilität betreffend, wichtige Informationen im Sinne der *evidence based practice* liefern (Fritz und Wainner 2001).

Reliabilitätstests unterstützen bzw. entkräften grundsätzlich keine osteopathischen Modelle oder Hypothesen, sondern untersuchen die Zuverlässigkeit des klinischen Tests unter wissenschaftlichen Bedingungen mit dem Ziel, Aussagen über seine Anwendbarkeit in der Praxis zu erlangen. Inwieweit die Methodologie derartiger Studien dem Vorgehen in die Praxis entspricht, bleibt jedoch zu hinterfragen. Dieses Problem wird von Kritikern wiederholt thematisiert (Johnston 1982, Korr 1991, Hollerwöger 2006, Stockkendahl et al. 2006) und spiegelt sich in den oft gegensätzlichen methodologischen Anforderungen in Bezug auf die interne und externe Validität (vgl. Kap. 2.1.1.).

### **2.3. Untersuchungen im Bereich manueller Tests an der Wirbelsäule**

Seffinger et al. (2004) gehen in ihrem Review der Frage nach, wie hoch die Inter- bzw. Intrarater-Reliabilität manueller Palpationstests an der Wirbelsäule ist. Nach einer Suche in 13 medizinischen Datenbanken (OSTMED-DR, PubMed, EMBASE, CINAHL, u.a.) und der Anwendung der Ein- und Ausschlusskriterien der Autoren bleiben 49 Artikel, die in diesem Review bearbeitet werden. Diese Studien behandeln die Interrater- bzw. Intrarater-Reliabilität von segmentalen Beweglich-



keitstests, Schmerzpalpationstests und Gewebespannungstests. Die Untersuchungen werden zuerst mit Hilfe eines von den Autoren entwickelten Bewertungsschemas in Studien hoher und niederer Qualität eingeteilt. Dabei zeigen sich Qualitätsmängel vor allem im Studiendesign, in der Beschreibung der Tester und der Patienten und in der Präsentation der Ergebnisse. In 14 Studien werden ausschließlich symptomatische Patienten beurteilt und in weiteren 9 Studien sowohl symptomatische als auch asymptomatische Patienten in die Untersuchung miteinbezogen.

Im weiteren Verlauf werden die Ergebnisse der hoch-qualitativen Studien statistisch miteinander verglichen, zwei Drittel dieser Untersuchungen verwenden die Kappa-Statistik.

Zusammenfassend zeigt sich akzeptable Reliabilität<sup>1</sup> ( $k=0,40$  oder höher) bei Schmerzprovokationstests, bei globalen Bewegungstests und Palpationstests anatomischer Orientierungspunkte. Prinzipiell haben Schmerzprovokationstests die höchste Reliabilität, Palpationstests der paraspinalen Weichteile die niedrigste. Regionale Bewegungseinschränkungen werden übereinstimmender diagnostiziert als segmentale Bewegungsdefizite. Die Intrarater-Reliabilität ist höher als die Interrater-Reliabilität. Die Erfahrung der Tester, deren Grundprofession, ein der Untersuchung vorgeschaltetes Training oder die Untersuchung ausschließlich symptomatischer Patienten haben keinen eindeutigen Einfluss auf die Reliabilität.

Viele Studien mit akzeptabler Reliabilität können nicht durch andere Untersucher mit ähnlichen Ergebnissen reproduziert werden. Die Validität der verwendeten manuellen Tests ist in vielen Fällen fraglich bzw. nicht ausreichend dokumentiert. Einheitliche Richtlinien für die Methodologie solcher Untersuchungen könnten die Qualität steigern und damit den Einsatz reliabler Palpationstests am Patienten ermöglichen.

Zu vergleichbaren Ergebnissen kommen Stochkendahl et al. (2006) in ihrem Review zum Thema Reproduzierbarkeit von Reliabilitätsstudien im Bereich der manuellen Untersuchung der Wirbelsäule. 48 Studien aus einer Grundgesamtheit von 900 in medizinischen Datenbanken auffindbaren Untersuchungen werden dafür analysiert. Auf Basis eines selbst entwickelten 6-Punkte-Schemas werden die Studien hinsichtlich Studiendesign, Studienpopulation und der statistischen Auswertung in solche hoher und solche niederer Qualität eingeteilt. Die Ergebnisse der

---

<sup>1</sup> Laut Fjellner et al. (1999) werden Werte von zumindest  $k=0,4$  als Indikator für eine akzeptable interindividuelle Verlässlichkeit angesehen.

qualitativ höherwertigen Untersuchungen werden in einer Meta-Analyse<sup>1</sup> miteinander verglichen. Die Ergebnisse der Analyse zeigen unter anderem, dass es eine hohe Evidenz für akzeptable Interrater-Reliabilität bei Schmerzprovokationstests an Knochen und Weichteilen und ebenfalls hohe Evidenz für inakzeptable Inter- und Intrarater-Reliabilität für segmentale Bewegungstests, statische Palpation und Gewebespannungstests gibt. Prinzipiell haben der Beruf der Tester, ihre Erfahrung, vorgeschaltetes Training und die Untersuchung von symptomatischen bzw. asymptomatischen Patienten keinen Einfluss auf die Reliabilität.

Die Autoren ziehen den Schluss, dass ein einzelner Test eine unzureichende Basis für eine Diagnose darstellt. Grund dafür ist unter anderem der diesen Studien immanente Konflikt zwischen interner und externer Validität. Während Blindierung, Randomisierung, Standardisierung des Testablaufs und paralleles Testen den Anforderungen für hohe interne Validität nachkommen, sind derartige Bedingungen in der Praxis kaum anzutreffen. Die Autoren kritisieren weiters die in vielen Studien mangelhafte Standardisierung der Untersuchungsabläufe und das weitgehende Fehlen eines unabhängigen Referenzstandards.

Nach den gleichen Qualitätskriterien wie Stockendahl et al. (2006) analysieren Haneline und Young (2009) in einem Review die Inter- und Intrarater-Reliabilität statischer Palpationstests im Bereich der Wirbelsäule und des Iliosakral-Gelenks. Von 29 die Ein- und Ausschlusskriterien erfüllenden Studien aus den Jahren 1965 bis 2007 fokussieren sich 14 Untersuchungen auf die reliable Palpation von Schmerzpunkten, 10 Untersuchungen beschäftigen sich mit der Reliabilität anatomischer Referenzpunkte und 5 Untersuchungen mit der Position knöcherner Strukturen im Bereich der Wirbelsäule.

Es zeigt sich, dass die Reliabilität der meisten Studien den akzeptablen Kappa-Wert von 0,40 nicht erreichen. Die Palpation von Schmerzen zeigt (statistisch nicht signifikant) bessere Inter- bzw. Intrarater-Reliabilität, als die Testung der Position anatomischer Orientierungspunkte. Die Autoren bringen 2 Punkte in die Diskussion ein. Einerseits wird bei einigen Untersuchungen eine exakte Definition des palperten dysfunktionalen oder schmerzhaften Wirbelsegments verlangt. Hier scheitern laut den Autoren viele Tester, was sich negativ auf deren Untersuchungsergebnisse auswirkt, obwohl die Tester womöglich die „richtige“ Stelle palpieren haben. Ein zweiter Diskussionspunkt ist die Frage, ob es möglich ist, an Pa-

---

<sup>1</sup> Eine Meta-Analyse fasst quantitative Untersuchungsergebnisse mehrerer Studien zusammen. Dadurch soll geklärt werden, ob ein bestimmter Effekt aufgrund dieser Untersuchungen existiert (Kleibel und Mayer 2005).

tienten objektiv wiederholt ein Schmerzsegment zu palpieren, ohne dass der Patient (auch durch die im Verlauf der Palpation zunehmende Sensitivität des schmerzenden Bereichs) durch seine Reaktion den Untersucher beeinflusst. Die Autoren formulieren, dass diese Art der Palpation eher eine Aussage darüber gibt, wie reliabel Patienten eine schmerzhaft Stelle wiedererkennen können.

Hollerwöger (2006) untersucht in seinem Review die Ergebnisse von Reliabilitätsstudien im Bereich manueller Tests an der Halswirbelsäule.

Nach Recherchen in medizinischen Datenbanken (PubMed, AMED, CINAHL, u.a.), Online-Journalen und einer Vorauswahl anhand der Ein- bzw. Ausschlusskriterien des Autors werden 15 Untersuchungen in diesem Review verarbeitet. In einem ersten Schritt werden sie einer qualitativen Prüfung nach den Vorgaben des QUADAS-Beurteilungprotokolls<sup>1</sup> nach Whiting et al. (2003) unterzogen. In einem zweiten Schritt erfolgt ein Outcome-Vergleich der ausgewählten Studien. Dafür unterteilt der Autor die Untersuchungen in solche, die das Ausmaß bzw. die Qualität segmentaler Wirbelsäulenbeweglichkeit untersuchen und Studien mit dem Ziel, Schmerz in diesem Bereich zu palpieren. Hollerwöger (2006) kritisiert in seiner Untersuchung das Fehlen unabhängiger Referenzstandards in vielen untersuchten Artikeln, fehlende Informationen über nicht in die Statistik einbezogene Untersuchungsergebnisse oder das Fehlen einer repräsentativen Studienpopulation. Vor diesem Hintergrund weichen die Ergebnisse der einzelnen Studien deutlich voneinander ab. Es zeigt sich keine eindeutige Tendenz die Reliabilität dieser Tests betreffend, sodass der Autor die Möglichkeit einer rein palpatorischen Diagnose zervikaler Dysfunktionen in Frage stellt. Dafür müssten seiner Meinung nach einerseits praxisorientiertere Studiendesigns, andererseits subjektive Parameter, beispielsweise die Patientenanamnese miteinbezogen werden.

### **2.3.1. Kritik und Kriterien für verbesserte Reliabilität**

Zusammenfassend können folgende qualitative Hauptprobleme von Reliabilitätsuntersuchungen manueller Tests an der Wirbelsäule festgestellt werden:

- Fehlen eines unabhängigen Referenzstandards (Seffinger et al. 2004, Hollerwöger 2006, Stockendahl et al. 2006)

---

<sup>1</sup> QUADAS ist ein Schema zur Qualitätsbeurteilung im Bereich von Diagnosemethoden. Neun wissenschaftliche Experten auf dem Gebiet der Diagnostik erarbeiteten nach Durchsicht einschlägiger Literatur einen aus 14 Punkten bestehenden Katalog, nach welchem diagnostische Studien in Reviews miteinander verglichen und qualitativ beurteilt werden können (Whiting et al. 2003).

- Mangelhafte Standardisierung des Untersuchungsablaufs (Hollerwöger 2006, Stochkendahl et al. 2006)
- Keine repräsentative Studienpopulation (Hollerwöger 2006, Stochkendahl et al. 2006)
- Mangelhafte Beschreibung des Studiendesigns (Seffinger et al. 2004)
- Mangelhafte Analyse und Präsentation der Studiendaten und –ergebnisse (Seffinger et al. 2004), fehlende Information über nicht in die Statistik einbezogener Daten (Hollerwöger 2006)

Diese Kritikpunkte decken sich größtenteils mit den drei Anforderungen, die Fritz und Wainner (2001) an qualitativ hochwertige Reliabilitätsstudien stellen:

- Ein mangelhafter, unpassender oder fehlender Referenzstandard beeinflusst die Validität der Untersuchung und damit ihre Reproduzierbarkeit (Fritz und Wainner 2001, Seffinger et al. 2004). Durch die Komplexität des menschlichen Bewegungsapparats könnte auch ein einzelner Referenzstandard zu wenig sein (Hollerwöger 2006).
- Eine unzureichende Standardisierung des Untersuchungsablaufs, fehlende Informationen über das Ziel des Tests, der Beurteilungskriterien für ein positives bzw. negatives Testergebnis, sowie fehlende Blindierung des Testers gegenüber dem Referenzstandard sind weitere Angelpunkte für mangelhafte Qualität von Reliabilitätsstudien (Fritz und Wainner 2001).
- Die Studienpopulation sollte laut Fritz und Wainner (2001) symptomatische Patienten umfassen, wobei sich der Einfluss dieses Kriteriums auf die Reliabilität nach Seffinger et al. (2004) nicht eindeutig bestätigen lässt. Stochkendahl et al. (2006) plädieren für einen Mix aus symptomatischen und asymptomatischen Patienten.

Sommerfeld (2006) führt weitere Punkte an, welche die Reliabilität in einer Untersuchung steigern:

- Der Messablauf soll standardisiert sein.
- Die Tester sollten den Umgang mit dem Messgerät trainiert haben.

- Das Messgerät sollte optimale Voraussetzungen bieten zB eindeutige Ablesbarkeit.
- Der Messvorgang sollte automatisiert ablaufen.
- Die Messungen sollten wiederholt werden.

Die hier beschriebenen Kriterien sollen soweit als möglich in die Methodologie dieser Studie miteinfließen. Die Durchführung der Messung mit Hilfe eines videounterstützten Onlinefragebogens sowie die Messung von visuell beurteilbaren Merkmalen in der FDM-Diagnostik könnten Vorteile im Bereich der Standardisierung mit sich bringen, die im weiteren Verlauf dieser Masterthese diskutiert werden.

### 3. Zur Bedeutung von Gestik und Körpersprache

Im FDM spielt Körpersprache eine zentrale Rolle. Typaldos (1999, 2002) verwendet in seinen Büchern diesen Begriff im Zusammenhang mit dem, was Patienten bei der Beschreibung ihrer körperlichen Beschwerden, vorwiegend mit den Händen, andeuten und aus diagnostischer Sicht Hinweis auf eine Faziendistorsion sein kann. Typaldos beschreibt Körpersprache als konsistente, unbewusste Bewegungen oder Haltungen (Typaldos 2002).

Bei der Literaturrecherche zum Thema Körpersprache ist auffallend, dass dieser Begriff im wissenschaftlichen Kontext selten aufscheint und stattdessen meist der Begriff Gestik (vor allem in der Gestikforschung) verwendet wird. Bei der Schlagwortsuche in PubMed werden unter anderem die MeSH-Terms *gesture* und *nonverbal communication* aufgelistet.

Um Erkenntnisse aus dem Bereich der Gestikforschung für diese Masterthese anwendbar zu machen, soll zuerst die Frage geklärt werden, ob FDM-Körpersprache als



Abb. 1: Körpersprache bei schmerzhafter Schulter (Typaldos 2002)

Gestik bezeichnet werden kann. Weiters soll der Stellenwert von Gestik als reliables Diagnoseinstrument in der Medizin beleuchtet werden. Basis dafür ist eine Literaturrecherche zum Thema Reliabilität von Gestik und Körpersprache in Online-Datenbanken (OSTMED-DR, Osteopathic Research Web, PubMed) mit den Begriffen reliability OR diagnos\* OR assessment in Zusammenhang mit body language OR gesture OR nonverbal communication OR kinesics. Erkenntnisse aus dem Bereich der medizinischen Kommunikation liefern weitere Informationen über wissenschaftliche Hintergründe der Körpersprache-Diagnose des FDM.

### **3.1. Definitionen**

In der wissenschaftlichen Literatur zum Thema Gestik, Körpersprache und non-verbale Kommunikation zeigt sich, dass diese Begrifflichkeiten von verschiedenen Autoren nicht immer einheitlich verwendet werden. Auch die Definitionen an sich überschneiden sich teilweise, sodass die hier präsentierten Beschreibungen als Annäherung zu verstehen sind.

#### **3.1.1. Körpersprache**

Körpersprache wird beispielsweise im Österreichischen Wörterbuch (2003) als „Ausdruck z.B. durch Mienenspiel, Körperhaltung“ beschrieben. Der Duden definiert Körpersprache als „in Körperhaltung, Bewegung, Gestik, Mimik sich ausdrückende psychische Konstitution, Gestimmtheit“ (Duden online 2011).

Körpersprache wurde vor allem in den 1970er Jahren populärwissenschaftlich derart verstanden, dass bestimmte Körperbewegungen und Körperhaltungen eine klare, allgemein (in einer Gruppe, Kulturkreis oder darüber hinausgehende) gültige Bedeutung haben. Diesem Ansatz kann gegenübergestellt werden, dass beispielsweise ein Nicken des Kopfes in verschiedenen Gesprächssituationen unterschiedliche Bedeutung haben kann, von Zustimmung bis zur Andeutung, dass man den verbalen Ausführungen eines Gesprächspartners folgt (Harrigan 2008). Körpersprache wird oft mit nonverbaler Kommunikation gleichgesetzt. Darunter werden Gestik, Mimik und optische Zeichen verstanden, die abseits der gesprochenen Sprache Teil der menschlichen Kommunikation sind (Navarro 2008, Bannert 2011). Knapp und Hall (2002) beschreiben nonverbales Verhalten als Verhalten ohne linguistischen Inhalt. Nonverbale Kommunikation kann den gesprochenen Inhalt unterstützen oder aber gegenteilige Informationen aussenden. Sie gilt ausserdem als schwerer kontrollierbar und wird deshalb als „echter“ angesehen (Bannert 2011).



In diesem Bereich bewegen sich auch Forschungen des Psychologen Paul Ekman (Ekman 2011), der sich im Bereich der Verhaltensforschung schwerpunktmässig mit dem Informationsgehalt von Gesichtsausdruck und Körpersprache in Bezug auf Emotionen oder dem Erkennen von dissonanter Kommunikation beschäftigt. Der Begriff der nonverbalen Kommunikation wird in Zusammenhang mit Rhetorik zB Neurolinguistisches Programmieren (NLP) und den nonverbalen, emotionalen Komponenten von interpersoneller Kommunikation verwendet (Navarro 2008).

### **3.1.2. Gestik**

Gestik (lat. gestus: Mienenspiel, Gebärdenspiel) beschreibt die nonverbale Kommunikation mit Armen, Händen und dem Kopf. Mimik, Körperbewegung und Haltung zählen im engeren Sinne nicht mehr zur Gestik (Nöth 1999).

Der Duden (Duden online 2011) beschreibt eine Geste als „spontane oder bewusst eingesetzte Bewegung des Körpers, besonders der Hände und des Kopfes, die jemandes Worte begleitet oder ersetzt [und eine bestimmte innere Haltung ausdrückt]“ oder als „Handlung oder Mitteilung, die etwas indirekt ausdrücken soll“. Als Gestik wird im Duden „die Gesamtheit der Gesten“ (Duden online 2011) definiert.

McNeill (1992) definiert Gestik als Bewegung der Arme und Hände während der verbalen Kommunikation. Kendon (1996) beschreibt Gestik als Bewegung vor allem der Arme und Hände in einem bestimmten, schematischen Ablauf mit klarem Beginn, Höhepunkt und Ende der Geste.

## **3.2. Charakteristik von Gesten**

Die inhaltliche Konzentration des Begriffs Gestik auf die oberen Extremitäten deckt sich nach diesen Definitionen mit der Art von Körpersprache, die in der FDM-Diagnose relevant ist. Deshalb wird im folgenden Gestik genauer untersucht.

### **3.2.1. Gestenkategorien**

Gesten können auf verschiedene Art und Weise eingeteilt werden. Ekman und Friesen (1969) entwickelten ein in weiten Teilen der Gestikforschung verwendetes Einteilungssystem:

- **Embleme** sind symbolische Gesten mit einer spezifischen Bedeutung, die von einem Großteil einer (kulturellen) Gruppe klar verstanden werden zB

Nicken oder das Zeichen für OK. Die Reliabilität im Codieren solcher Gesten ist sehr hoch, wenn innerhalb dieser (kulturellen) Gruppe getestet wird.

- **Illustratoren** sind Gesten, die verbal Artikuliertes gestisch darstellen zB ein Piktogramm für ein beschriebenes Objekt oder eine Zeigegeste, um einen Ort genauer festzulegen.
- **Adaptoren**, auch Selbstadaptoren genannt, sind Gesten, die mit Berührungen des eigenen Körpers einhergehen. Das reicht vom kurzen Kratzen bis zu längerem Streicheln bestimmter Körperstellen. Sie passieren mit weniger Aufmerksamkeit als bei Emblemen und können nach verschiedenen Kriterien eingeteilt und analysiert werden (zB zeitliche Dauer oder Körperregion, die berührt wird). Interessanterweise konnte in einer Untersuchung auch festgestellt werden, dass Patienten, während sie die Fragen des Arztes beantworten, mehr derartiger Gesten zeigen als in Ruhe (Harrigan 2008).
- **Regulatoren** sind Gesten, die in Zusammenhang mit dem Gesprächsfluss innerhalb einer Konversation stehen zB Augenkontakt oder Bewegungen der Augenbrauen.

McNeill<sup>1</sup> (1992) definiert alternativ mehrere Kategorien von Gesten:

- **Ikonische Gesten** (iconics) sind symbolhafte Gesten mit einem starken Zusammenhang zum semantischen Inhalt der Sprache zB eine Geste, die eine Aufwärtsbewegung versinnbildlicht, wenn parallel davon gesprochen wird.
- **Metaphorische Gesten** (metaphorics) ähneln den ikonischen Gesten, haben aber abstrakteren Inhalt.
- **Zeigende Gesten** (deictics) dienen der genauen räumlichen Bestimmung von Punkten bzw. Orten. Sie können sowohl abstrakt zB das Zeigen in

---

<sup>1</sup> David McNeill ist Psychologe und lehrt an der University of Chicago. Sein Forschungsschwerpunkt liegt auf dem Zusammenhang zwischen Sprechen und Denken. Gestik ist für ihn eine der Schnittstellen zwischen bildhafter Vorstellung und verbaler Sprache. Deshalb wird Gestik von ihm in Zusammenhang mit Sprache untersucht (McNeill 2011, 2005).



einem imaginären Raum, als auch konkret wie das Zeigen am eigenen Körper sein.

- **Rhythmische Gesten** (beats) betonen bestimmte Teile einer verbalen Aussage.
- **Verknüpfende Gesten** (cohesives) sind exakte Wiederholungen bereits benutzter Gesten im selben Kontext und verdeutlichen Zusammenhänge. Cohesives betonen Sachverhalte. Sie können an sich ikonische, metaphorische, rhythmische oder zeigende Gesten zum Inhalt haben. McNeill führt sie in der Folge (2005) nicht mehr als eigenständige Kategorie auf.

Kendon (1996) beschreibt unter Berufung auf die Arbeiten des Anthropologen Efron, dass innerhalb kultureller Gruppen Gestenmuster erkennbar sind, die über Embleme hinausgehen und unterteilt den Informationsgehalt von Gesten in 3 Gruppen:

- Alleinstehende gestische Äußerungen zB Embleme
- Gestische Äusserungen im Wechsel mit verbalen Äußerungen zB das Beenden eines Satzes mit einer Geste
- Gestische Äusserungen in Verbindung zu verbalen Äußerungen zB Gestikulieren während einer Rede, dabei ist der Inhalt der Gestik gemeinsam mit dem Inhalt der verbalen Äußerung zu beurteilen

### 3.2.2. Temporaler Ablauf und Inhalt von Gestik

Kendon (1996) beschreibt vier Merkmale, die Gesten von anderen körperlichen Bewegungen und Haltungen unterscheiden:

- Gesten sind Exkursionen dh sie haben eine Startposition und kehren wieder dahin zurück
- Gesten haben ein zentrales Element, einen *peak oder stroke*<sup>1</sup>
- Gesten haben einen klaren Beginn und ein klares Ende

---

<sup>1</sup> Der *stroke* bzw. *peak* wird synonym für den Gestenhöhepunkt, dem zentralen Inhalt der Geste verwendet (Kendon 1996, McNeill 1992).

- Gesten sind symmetrisch, dh wenn Gesten gefilmt werden, kann im Vorlauf bzw. Rücklauf des Filmes kaum ein Unterschied ausgemacht werden

McNeill (1992, 2005) beschreibt einen Ablauf der meisten Gesten in drei Phasen. Die erste, vorbereitende Phase führt die Hand zum Ausgangspunkt der Geste. Die zweite Phase, der *stroke*, ist die Geste an sich und kann wiederholt werden. Die dritte Phase ist die Rückbewegung der Hand zum Ausgangspunkt. Er betont, dass Gestik motorisch vor der gesprochenen Sprache beginnt. Der *stroke* fällt mit der signifikanten verbalen Aussage zusammen, wobei zuvor jedoch bereits die erste Phase der Gestenproduktion stattgefunden hat. Für McNeill (1992, 2005) ist das ein Hinweis für die primär bildhafte Entstehung einer verbalen Aussage im Gehirn. Er definiert einen *growth point*, die kleinste kommunikative Einheit mit bildhaften und linguistischen Elementen. Diese Einheit ist kontextabhängig und unterstreicht nach McNeill (2005) die gegenseitige Abhängigkeit von Gestik und Sprache.

McNeill (1992) beschreibt Gesten als

- kontextsensitiv, dh der Inhalt der Gestik ist im Zusammenhang mit der (verbalen) Aussage bzw. Situation zu sehen
- nicht kombinatorisch, dh im Gegensatz zur verbalen Sprache beinhaltet jede Geste einen eigenen Bedeutungsinhalt, ähnlich einem Buchstaben im Alphabet
- global und synthetisch, dh eine Geste besteht aus mehreren Bestandteilen (Einzelbewegungen), diese sind im Zusammenhang mit der gesamten Geste zu sehen (Anmerkung: ähnlich chinesischer Schriftzeichen)
- frei von formalen Standards dh das verschiedene Menschen drücken gleiches gestisch unterschiedlich aus.

Gesten sind kein Ausdruck von Gedanken, sondern ein eigener Informationsträger. Gesten sind bildhaft und werden durch das Gedächtnis und die Dynamik der Kommunikation beeinflusst. Diese beiden Faktoren können den gestischen Ausdruck bis hin zu deren vollkommener Absenz beeinflussen zB bei Demenzkranken (McNeill 2005).

Die Gestikforschung orientiert sich in weiten Teilen daran, dass Gestik in Zusammenhang mit verbaler Kommunikation steht. Dieser in der Linguistik entstandene Ansatz (Marcantonio 2008) spiegelt jedoch nur teilweise die Annahme des FDM, dass Gestik sowohl beobachtet innerhalb der Therapiesituation als auch unbeobachtet, bedeutungsvoll und codierbar ist, wider. Ob Gestik in diesen beiden, prinzipiell unterschiedlichen Situationen vergleichbar ist, bleibt fraglich und steht mit der in der Gestikforschung ungeklärten Frage in Verbindung, ob Gestik primär für den Sprecher oder den Zuhörer wichtig ist (McNeill 2005). Denn aus der Beobachterposition erscheinen Gesten oftmals abstrakt, obwohl sie für den Ausführenden persönlich Erlebtes widerzugeben scheinen.

*„By performing the gesture, a core idea is brought into concrete existence and becomes part of the speaker's own existence at that moment.“  
(McNeill 2005: 99)*

Basis vieler Untersuchungen von McNeill (1992) sind Erzählsituationen, in denen Probanden einen zuvor gesehenen Zeichentrickfilm nacherzählen. Diese Erinnerungssituation stimuliert offensichtlich die Gestenbildung. Daraus könnte geschlossen werden, dass das Kreieren einer derartigen Erzählsituation auch in der medizinischen Kommunikation ein probates Mittel für die Förderung der Produktion von Gesten seitens des Sprechers ist.

### **3.3. Forschung im Bereich Gestik und medizinischer Kommunikation**

Denkt man an eine Gesprächssituation und die Art und Weise, wie viele Körperbewegungen und -haltungen ein Gegenüber währenddessen präsentieren kann, scheint es nahezu unmöglich, diesen Bund an nonverbalen Informationen verstehen zu können. Die Gestikforschung bedient sich dreier Kriterien, um die Datensammlung zu erleichtern (Harrigan 2008):

- Die meisten nonverbalen Informationen werden von den Händen und dem Kopf übermittelt.
- Es gibt einen natürlichen Rahmen an Körperbewegungen, die bewertet werden können. Sehr ausgefallene Gesten wie abstrakte Verenkungen von Rumpf oder Armen werden in der Gestikforschung im Allgemeinen nicht berücksichtigt.

- Viele Teilbewegungen von Gesten zeigen sich in einem bestimmten zeitlichen Ablauf, was sie im Gegensatz zu Einzelbewegungen leichter codierbar machen (vgl. Kap. 3.2.2.)

Die Gestikforschung des vorigen Jahrhunderts durch Efron, Ekman und Friesen sowie Morris untersuchten Gesten als eigenständiges Kommunikationsmittel. Es zeigte sich, dass es kulturelle Unterschiede im nonverbalen Verhalten (Mimik, Gestik und Distanzverhalten) gibt, gleichzeitig scheint es aber gemeinsame anthropologische Grundlagen für nonverbale Kommunikation zu geben, die vom jeweiligen Kulturkreis modifiziert werden (Marcantonio 2008). Im Gegensatz zur Gebärdensprache, die ein eigenständiges, nonverbales, an den jeweiligen Kulturkreis gebundenes Zeichensystem ist, wird Gestik seit den 1980er Jahren besonders von der Linguistik in ihrem Zusammenspiel mit verbaler Sprache analysiert. Diese Entwicklung generiert sich auch aus der Beobachtung, dass die Unterdrückung der Körpersprache die kommunikativen Möglichkeiten reduziert. Interessant in diesem Zusammenhang ist auch die anthropologische Entwicklung der Gestik von einer ursprünglich durch einzelne Laute begleiteten Körpersprache zu einer dominant verbalen Sprache, die durch Gesten unterstützt wird (Marcantonio 2008). Die Studie dieser Masterthese beruht auf einer videodokumentierten Gesprächssituation eines Osteopathen mit einzelnen Patienten und entspricht dem linguistischen Ansatz zur Gestenuntersuchung und deren Codierung.

In der Analyse von Arzt-Patienten-Gesprächen wird die nonverbale Komponente der Kommunikation seltener studiert als ihr verbaler Anteil, obwohl nonverbale Zeichen den Verlauf dieses Gesprächs bis hin zur Diagnosestellung entscheidend beeinflussen können (Schmid Mast 2007). Anhand der Analyse eines derartigen Arzt-Patienten-Gesprächs gibt die Linguistin Stukenbrock (2008) Einblick, wie gestisches Zeigen (Deiktik) innerhalb der Arzt-Patient-Kommunikation eingebettet ist. Dabei wird ein Patient mit chronischen Schmerzen in einem videodokumentierten Gespräch von einem Chefarzt zu seinen Beschwerden befragt, zusätzlich befinden sich andere Ärzte als Beobachter im Raum.

Auf die Frage „Wo ist Ihr Hauptschmerz?“ antwortet der Patient sowohl gestisch als auch verbal. Diese Antwort kann in fünf Phasen gegliedert werden. In der ersten Phase wendet der Patient seinen Blick auf sein eigenes Knie. Er stellt damit einen Raum für die folgende Geste her und fokussiert die Aufmerksamkeit des Adressaten (des Chefarztes) auf diesen Körperbereich. Die zweite Phase besteht aus der Geste an sich (der Patient klopft mehrmals auf sein Knie). Stukenbrock (2008) betont den Vorteil deiktischer Gesten, die sich im Blickfeld des Patienten

befinden und der Patient spüren kann, weil sie einerseits dem Patienten selbst *feedback* erlauben und andererseits Klarheit für den Adressaten der Geste schaffen. Beeinflusst kann dies durch die Erreichbarkeit des Körperbereichs und eventuell kulturelle Tabus werden. In der dritten Phase folgt die verbale Deiktik („da“), die in der vierten Phase zu einer akzentuierten Geste des Patienten führt. Phase fünf besteht darin, dass der Patient wieder Blickkontakt mit dem Chefarzt aufnimmt, der seinerseits die Aufnahme der Information bestätigt. Als Teilsimultanität beschreibt Stukenbrock (2008) den phasenverschobenen und zugleich phasenüberlappenden Ablauf verbaler und gestischer Deiktik, die sich zu einer Aussage verdichten.

In einem weiteren Abschnitt der Untersuchung wird das Zeigen des Patienten am eigenen Körper ohne Sichtkontrolle demonstriert. Dabei präsentiert der Patient eine Geste im Bereich seines Rückens, welche im ersten Versuch weder für ihn selbst noch für sein Gegenüber durch die unbedingte Erhaltung des Blickkontakts zum Chefarzt sichtbar ist. Nach Goffmann (Goffmann 1963, zit. in: Stukenbrock 2008) ist der Blickkontakt die wichtigste Möglichkeit für den Sprecher, die Wahrnehmung des Zuhörenden zu beobachten und wird deshalb nur ungern aufgegeben. Im vorliegenden Beispiel dreht der Patient erst nach mehrmaliger Nachfrage den Rücken zum Chefarzt, um die Geste für diesen sichtbar zu machen.

In dieser Phase zeigt sich eine weitere Herausforderung in der Evaluierung von Gestik. Naivität und Unverstelltheit sind Voraussetzungen für spontane, klare Gesten. Außerdem scheint das Zeigen selbst der Modus größtmöglicher Authentizität zu sein. Vor allem Patienten mit einer längeren medizinischen Biographie vermengen ihre subjektiven Wahrnehmungen mit erworbenen Wissen, unter anderem aus Voruntersuchungen. In der hier dargestellten Situation beschränkt sich der Patient auf die verbale Beschreibung eines „L4/5-Problems“. Der Chefarzt akzeptiert diese Antwort des Patienten nicht und zwingt ihn durch mehrmaliges Rückfragen und Zurückweisen der verbalen Antworten dazu, sich umzudrehen und eine Geste am Rücken zu zeigen.

Stukenbrock (2008) schließt aus ihrer Untersuchungen, dass Deiktik im Kontext des jeweiligen Gesprächsinhaltes gesehen werden muss. Sprache und Gestik beeinflussen sich gegenseitig, treten jedoch nicht immer simultan auf.

### **3.4. Gestik und Körpersprache in der Osteopathie**

Das Faziendistorsionsmodell wird von Osteopathen in der Praxis angewendet und dabei die visuelle Diagnose der Körpersprache eingesetzt. Im Allgemeinen spie-

len jedoch Gestik und Körpersprache im osteopathischen Befund keine Rolle. In einer qualitativen Studie zum Thema Basisuntersuchungen in der Osteopathie zeigt Biberschick (2011), dass klassische visuelle Diagnosemethoden wie die Begutachtung der Statik bei einigen Autoren osteopathischer Standardliteratur zur Untersuchung des Patienten gehören, eine gezielte Analyse der Patientengestik oder Körpersprache im Sinne des Faszienmodells spielen dabei bis dato keine Rolle. Hier könnten durchaus wichtige Informationen für die Befunderhebung verloren gehen. Es ist jedoch anzumerken, dass nonverbales Verhalten bisweilen generell als medizinisches Diagnosekriterium kaum untersucht ist (Schmid Mast 2007).

### **3.5. Gestik und Körpersprache als reliables Diagnosekriterium**

Die Diagnose von Gestik und Körpersprache erfolgt in einem visuellen Verfahren. Visuelle Diagnostik ist der Schwerpunkt der Radiologie, die sich mit der bildlichen Darstellung und Diagnose von Körperstrukturen und Funktionen beschäftigt. Im Gegensatz zu manualmedizinischen und osteopathischen Reliabilitätsuntersuchungen bieten sich durch klar definierte Beurteilungskriterien und einer weitgehenden Standardisierung des Untersuchungsablaufs gute Voraussetzungen, Reliabilitätsuntersuchungen mit hoher methodologischer Qualität hervorzubringen. In radiologischen Untersuchungen wird darauf Wert gelegt, dass Entscheidungskriterien für oder gegen eine Diagnose einfach, klar und eindeutig formuliert und anwendbar sind. Die Studien von Peterson et al. (2007) und Schiphof et al. (2008) sind Beispiele für diesen Ansatz, der offensichtlich in Untersuchungen der Reliabilität manueller Diagnoseverfahren schwer umzusetzen ist.

Reliabilitätsuntersuchungen visueller Diagnostik im Bereich von Gestik und Körpersprache sind im Gegensatz dazu nur im geringen Ausmaß in medizinischen Datenbanken zu finden, obwohl sich Gestik sehr gut und erfolgreich auf ihre Reliabilität hin testen lässt. Kappa-Werte von 0,80 und höher werden in vielen Studien im Bereich der (nicht-medizinischen) Gestikforschung erreicht und können durch gezieltes Training der Tester weiter verbessert werden (Harrigan 2008). Schmid Mast (2007) betont weiters die hohe Genauigkeit der Testpersonen in der Beurteilung nonverbalen Verhaltens eines Gegenübers.

Die größte Anzahl an miteinander vergleichbaren medizinischen Reliabilitätsstudien im Bereich Gestik findet sich in der Notfallmedizin mit Konzentration auf die Diagnose von Herzinfarkten. Bei diesen Studien werden die Reliabilität und Validität von Handgestik als sichtbarer Hinweis auf ein ischämisches Geschehen un-

tersucht. Auch im Bereich der Neurologie, Psychiatrie, Pädiatrie und Gerontologie gibt es Ansätze, die nonverbale Kommunikation bei Patienten mit verbalen Kommunikationsstörungen unter verschiedensten Aspekten zu erforschen.

Die drei aufgeführten Studien aus dem kardiologischen Bereich berühren die Thematik dieser Masterthese, weil hier explizit Patientengestik für die Diagnosestellung herangezogen wird. Ausserdem behandeln diese Studien miteinander vergleichbare Fragestellungen und geben einen Einblick in die Ergebnisse und Möglichkeiten von Reliabilitätsuntersuchungen bei diagnoseunterstützender Handgestik. Interessant ist die Thematik auch deshalb, weil Typaldos (2002) in der vierten Auflage seines FDM-Buches „FDM - Clinical and Theoretical Application of the Facial Distortion Model within the Practice of Medicine and Surgery“, die Kardiologie als zukünftiges Anwendungsgebiet des FDM hervorstreicht.

### **3.5.1. Untersuchungen aus dem Bereich der Kardiologie**

Edmondstone (1995) beschreibt, dass Patienten mit Brustschmerzen kardialen Ursprungs während der Anamnese meist drei typische Handbewegungen zeigen. Dabei handelt es sich um das Zeichen der flachen Hand auf dem Brustbein, der geschlossenen Faust auf dem Brustbein (*Levine's sign*) und das Zeichen von beiden Händen auf der Brust, die entlang der Rippen nach aussen ziehen und ein „Zuschnüren“ symbolisieren. Im Gegensatz dazu deuten laut Edmondstone Patienten mit gastro-ösophagealen Schmerzen eher eine Linie mit den Fingerspitzen entlang des Brustbeins an.

In seiner Studie werden 203 Patienten mit Brustschmerzen während eines Zeitraums von sechs Monaten begleitend beobachtet, um die Sensitivität, die Spezifität und den klinischen Vorhersagewert dieser Gesten in Bezug auf ischämische Brustschmerzen zu untersuchen. Die Patienten werden in der initialen Arztkonsultation gebeten, den Schmerz am Körper zu zeigen und zu beschreiben, wie dieser sich anfühlt. Die untersuchenden Ärzte ordnen die beobachteten Patientengesten den oben genannten drei Kategorien zu oder vermerken andere Gesten. Diese Aufzeichnungen werden für mindestens ein Jahr aufbewahrt, um in dieser Zeit, die Ursache der Brustschmerzen medizinisch abzuklären. Von den teilnehmenden Patienten haben 138 kardiale Herzprobleme, 21 Patienten haben Schmerzen nicht-kardialen Ursprungs und bei 44 Patienten ist die Ursache der Schmerzen unklar. Es zeigt sich, dass 80% (n=110) der Herzpatienten, aber nur 51% (n=33) der Patienten ohne Herzprobleme die typischen Handbewegungen zeigen. 46% (n=64) der Herzpatienten zeigen das Zeichen der flachen Hand auf dem



Brustbein, 14% (n=19) das *Levine's sign* und 20% (n=27) das Zeichen der beiden ziehenden Hände auf dem Brustbein ( $p < 0,01$ ). Edmondstone folgert aus seiner Studie eine hohe Wahrscheinlichkeit von 77%, dass oben beschriebene Gesten bei der Erkennung von Brustschmerzen ischämischen Ursprungs weiterhelfen, verbunden mit einer hohen Sensitivität (80%) und einer geringen Spezifität (49%).

Diese Studie beschreibt wenige Details ihrer Methodologie, der Studiendaten und deren statistischer Auswertung. Ihre Aussagekraft ist somit kritisch zu hinterfragen. Trotzdem zeigen die Beobachtungen von Edmondstone (1995), dass Patientengestik in der klinischen Diagnose eine Rolle spielen könnte.

Albarran et al. (2000) adaptieren und erweitern die Untersuchungen von Edmondstone (1995). Sie befragen 267 Patienten unter anderem, wo ihr Schmerz war, als sie im Brustbereich begannen. In dieser Studie werden die Patientengesten in vier Gruppen eingeteilt: die flache Hand am Brustbein, eine streichende Bewegung mit der Hand über das Brustbein, geschlossene Fäuste über dem Brustbein und ein lokalisierendes Zeigen mit einem Finger auf dem Brustbein. Von den 267 Patienten haben 118 Patienten nachweislich einen Herzinfarkt erlitten, bestätigt durch ein EKG und/oder einen Anstieg von Herzenzymen im Blut. 149 haben keinen Herzinfarkt, ihre Beschwerden werden Angina pectoris, Arrhythmien, Pericarditis oder muskuloskeletalen bzw. ösophagealen Ursachen zugeschrieben. Die Patienten werden spätestens 24 Stunden nach ihrer Einlieferung ins Krankenhaus befragt, nachdem sie schmerzfrei sind. In der Studie werden neben der Gestik der Patienten andere Parameter wie das Ausmaß der Ausstrahlung der Schmerzen, die verwendeten Worte für die Schmerzbeschreibung oder die subjektiv formulierten Ursachen der Beschwerden untersucht.

In Hinblick auf die Gestik zeigt sich kein statistisch signifikantes Ergebnis, was sowohl die Form als auch die Häufigkeit der verwendeten Gesten in den beiden Gruppen betrifft, wobei sich die Gruppe der Herzinfarktpatienten bei der gestischen Beschreibung tendenziell auf zwei Handbewegungen konzentriert (Geste der flachen Hand auf der Brust und Geste der Faust auf dem Brustbein) und in der Gruppe der Patienten ohne Herzinfarkt die Streuung der verwendeten Gesten größer ist. Wie schon zuvor in der Studie von Edmondstone (1995) ist die Geste der flachen Hand jedoch das von den Herzinfarkt-Patienten am häufigsten verwendete Handzeichen.

Die Untersuchung von Albarran et al. (2000) beschäftigt sich in einem Teilbereich mit der Reliabilität von Gestik bei Herzinfarktpatienten. Durch die Abwandlung des Gestenvokabulars ist ein Vergleich mit der Studie von Edmondstone (1995) nur be-



dingt möglich. Im Gegensatz zu Edmondstone (1995) werden die Patienten in einem subakuten Zustand befragt, was Auswirkungen auf die Gestenproduktion und deren Prägnanz haben könnte (vgl. Stukenbrock 2008). Andererseits entspricht es der Untersuchungssituation, in der McNeill (1992) Gestik anhand der Erinnerung untersucht (vgl. Kap. 3.2.2.).

Marcus et al. (2007) untersuchen 202 Patienten mit Brustschmerzen, Einschlusskriterium ist ein Alter über 18 Jahren, Ausschlusskriterien sind schwerwiegende kardio-pulmonale Erkrankungen wie pulmonale Embolie oder Aorta-Dissection. In einem standardisierten Interview werden die Patienten durchschnittlich einen Tag nach dem letzten Auftreten der Schmerzen in einer blindierten Untersuchung simultan von mehreren Untersuchern befragt bzw. beobachtet. Angenommen wird, dass Patienten mit einer myokardialen Ischämie unter anderem signifikant häufiger das *Levine's sign*, das *palm sign* (flache Hand am Sternum) oder das *arm sign* (kontralaterale Hand fasst den linken Arm) zeigen und bei Patienten, die keine myokardiale Ischämie haben, das *pointing sign* (zeigen mit dem Finger auf die Brust) häufiger ist. Die Laborwerte (Troponin I) und andere medizinische Untersuchungen und Eingriffe führen bis zu vier Monate nach Einlieferung der Patienten zur Klärung der Ursachen der Brustschmerzen und der definitiven Einteilung der Patienten in die beiden Gruppen ischämische und nicht-ischämische Brustschmerzen. Die Ergebnisse der Untersuchung sind vielschichtig. 53% (n=107) der Patienten zeigen eine der Gesten, wobei das *palm sign* wiederum die am häufigsten (35%, n=72) codierte Geste ist. Keine der Gesten erreicht eine höhere Sensitivität als 38%. Der negative Vorhersagewert erreichte Werte von unter 68%. Die Autoren schließen daraus, dass das Fehlen von Gesten bei Patienten eine ischämische Ursache für die Beschwerden nicht ausschließen kann. Hingegen scheint das *pointing sign* mit einer Spezifität von 98% bei nicht-ischämischen Brustschmerzen einen positiven Vorhersagewert in Bezug auf einen nicht erhöhten Troponin-I-Wert zu geben, wobei die Autoren durch den geringen Anteil dieser Geste an der Gesamtkohorte dieses Ergebnis relativieren. Die Interrater-Reliabilität wurde anhand einer simultanen Untersuchung von 22 Patienten getestet. Es zeigt sich eine 100%ige Übereinstimmung ( $k=1.000$ ) bei der Diagnose des *Levine's sign* und des *arm sign*, eine 100%ige Übereinstimmung beim *pointing sign* und eine 95%ige Übereinstimmung ( $k=0,9014$ ) hinsichtlich des *palm sign*. Interessant ist, dass die Handgesten bei Fragen nach der Lokalisation der Schmerzen häufiger gezeigt werden als bei Fragen nach der Qualität der Schmerzen.

Diese drei Studien zeigen die Möglichkeit, wie Gestik und deren Codierung in die Diagnosestellung eingebaut werden können. Durch den vorhandenen Referenz-

Standard in der Diagnostik ischämischer Brustschmerzen können Aussagen über die Validität von Gestik getroffen werden. Damit erfüllen diese Untersuchungen die Kriterien, die in Kapitel zwei beschrieben werden. Die unterschiedlichen Studiendesigns mit teilweise stark aufgefächerten Fragestellungen und verschiedenen statistischen Ansätzen machen die Ergebnisse schwer zu vergleichen. Im Allgemeinen zeigt sich jedoch keine akzeptable Reliabilität der Patientengesten.

### **3.6. Zusammenfassung und Konsequenz für diese Masterthese**

Die Körpersprache im FDM besteht hauptsächlich aus händischem Zeigen des Patienten am eigenen Körper, Mienenspiel oder Körperhaltungen spielen keine Rolle. Der Begriff Deiktik beschreibt folglich per Definition am besten die Körpersprache im FDM, zumindest in der Osteopath-Patient-Kommunikation. Adaptoren sind unbewusster und spiegeln ebenfalls das körpernahe Verhalten der Patienten wider. Dies ist von Interesse, weil im FDM Gestik einerseits während der Behandlung beobachtet und beurteilt wird. In der FDM-Ausbildung wird andererseits darauf hingewiesen, dass bei Patienten, die sich unbeobachtet fühlen, ebenfalls Körpersprache präsent ist, die als Indiz für eine bestimmte Faszienstörung gewertet werden könnte. Generell kann FDM-Körpersprache wissenschaftlich eingeordnet werden und Hintergrundinformationen aus der Gestikforschung für die FDM-Diagnostik herangezogen werden.

Unklar bleibt allerdings, ob sich die FDM-Diagnostik von sich aus auf Gesten konzentriert, denn der temporale Charakter der Gestik im FDM ist nicht eindeutig definiert. Führt man sich die Beschreibung der Patientengesten und deren Illustration in den Büchern von Typaldos (1999, 2002) vor Augen (vgl. Kap. 4.3.) bzw. analysiert man Videos, die in Zusammenhang mit der Lehre des FDM präsentiert werden, können diese temporalen Elemente durchaus beobachtet werden. Sie sind jedoch keine Voraussetzung, um Körperbewegungen des Patienten als FDM-Körpersprache bzw. Gestik zu klassifizieren. Damit kann nicht eindeutig festgestellt werden, ob im FDM ausschließlich Gestik codiert wird. Prinzipiell wäre eine genauere Definition des Begriffs Körpersprache im FDM vorteilhaft. Hinsichtlich der Reliabilität und der darauf aufbauenden Diagnostik haben so exakt als möglich definierte Kriterien und Standards Vorteile bzw. sind Voraussetzung für hohe Messgenauigkeit (vgl. Kap. 2.2.3.).

Folglich wird der Begriff Körpersprache als FDM-typisch in dieser Masterthese weiterverwendet. Der Begriff Gestik bzw. nonverbale Kommunikation wird in Zusammenhang mit einschlägigem wissenschaftlichen Material verwendet.

Reliabilitätsuntersuchungen im Bereich der visuellen Diagnostik von Gestik und Körpersprache sind ein Randthema in medizinischen Datenbanken. Die präsentierten Studien aus dem Bereich der Akut-Kardiologie zeigen, dass Gestik offensichtlich ein gewisser Informationsgehalt bzw. Hinweischarakter innelegt. In den angeführten Untersuchungen können jedoch im Allgemeinen keine ausreichenden Belege für akzeptable Reliabilität präsentiert werden. Die visuelle Diagnostik im FDM kann im Gegensatz zu diesen Untersuchungen keinen Referenzstandard vorweisen. Faziendistorsionen und deren Diagnose durch FDM-Körpersprache sind Teil der modellhaften Vorstellung des FDM. Und spielen ausserhalb dieses Systems in der Medizin derzeit keine Rolle. Abgesehen vom Konzept des Hernierten Triggerpunkts, der teils als Pathologie der Hernie in der Medizin ein anerkannte Diagnose darstellt<sup>1</sup>, gibt es derzeit keine Möglichkeit, das Auftreten von Faziendistorsionen alternativ zu verifizieren und damit prinzipielle Voraussetzungen für einen Referenzstandard zu schaffen. Dadurch müssen jegliche Untersuchungsergebnisse relativiert werden, da eine Aussage betreffend der Validität nicht möglich ist.

## 4. Das Faziendistorsionsmodell (FDM)

Im Gegensatz zur Osteopathie und der Schulmedizin spielt Körpersprache als Diagnose-Element im FDM eine entscheidende Rolle. In diesem Kapitel wird die zentrale Position der Körpersprache im FDM dargestellt, sowohl aus einem geschichtlichen Blickwinkel als auch in der praktischen Umsetzung des FDM. Basis dieses Kapitels sind die beiden letzten Auflagen des FDM-Buchs von Typaldos (1999, 2002) sowie die in der Online-Datenbank OSTMED-DR gelisteten Artikel aus den Anfängen des FDM. Weil Typaldos sein Modell fortlaufend erweiterte und überarbeitete, beziehen sich die Referenzen dieser Masterthese großteils auf seine letzte Publikation, der vierten Auflage seines Buches (Typaldos 2002).

### 4.1. Entwicklungsgeschichte

Das Faziendistorsionsmodell (FDM) wurde von Dr. Stephen Typaldos, einem U.S. amerikanischen Arzt und Osteopathen im Lauf mehrerer Jahre entwickelt. Vor allem durch empirische Studien im Kontakt mit Patienten, durch persönliche Er-

---

<sup>1</sup> Im FDM gibt es die mögliche Pathologie eines Hernierten Triggerpunkts an der Leiste. Dieser entspricht weitgehend der schulmedizinischen Diagnose eines Leistenbruchs bzw. einer Leistenhernie. Ähnlich ist auch das therapeutische Vorgehen. Sowohl in der Schulmedizin als auch in der manualtherapeutischen Anwendung des FDM wird die Hernie (mit den jeweiligen Mitteln) reponiert und damit die Einklemmungssymptomatik an der Durchtrittspforte beseitigt.

lebnisse mit mangelnden Behandlungserfolgen und Dissektionen am Präparat entstand ein immer detaillierteres System zur Diagnose und Korrektur von Faszien-distorsionen (European FDM Association 2011).

Als Faszien-distorsionen beschrieb Typaldos Verformungen der Faszie, die aus seiner Sicht Ursache für die meisten körperlichen Beschwerden waren, mit welchen Patienten ihn in der Praxis konfrontierten. Er entwickelte ein spezielles Diagnoseverfahren, um diese Faszien-distorsionen am Patienten festzustellen. Besonders wichtig erschien ihm die Körpersprache des Patienten. Er entdeckte, dass Patienten, unabhängig von ihrer schulmedizinischen Diagnose ihre Beschwerden auf ähnliche Art und Weise, verbal und vor allem auch nonverbal äußerten. Diese Beobachtung war Anlass dafür, dass er im Lauf der Entwicklung des FDM fortlaufend, für eine Faszien-distorsion typische Körpersprache klassifizierte und in sein Diagnosemodell integrierte (European FDM Association 2011).

Die erste Faszien-distorsion, das Triggerband, gründete auf der modellhaften Vorstellung, dass sich bandartige Faszie pathologisch verformen kann und dadurch ursächlich für Schmerzen und Mobilitätseinschränkungen wird. Im Gegenzug kann sie aber wieder in ihren Ursprungszustand zurückgeformt werden (Typaldos 1994). Nach und nach kategorisierte er drei weitere Faszien-distorsionen (Hernierte Triggerpunkte, Continuum-distorsionen, Faltdistorsionen), welche aus Sicht von Patienten auf typische Art und Weise erlebt, durch Körpersprache präsentiert sowie vom Therapeuten spezifisch korrigiert werden konnten. Als Osteopath verwendete er besonders in den Anfangsjahren verschiedene osteopathische Korrekturtechniken, deren Wirkung er anhand des FDM erklärte und spezifisch für die Korrektur einzelner Faszien-distorsionen einsetzte. Beispielsweise nutzte er *myofascial release*, um bestimmte Triggerbänder und Faltdistorsionen zu korrigieren oder sah in kranialen Techniken einen Behandlungsansatz zur Korrektur von anatomisch entfernt liegenden Faszien-distorsionen (Typaldos 1994). Diese Ansätze spielten in seinen späteren Publikationen jedoch keine wichtige Rolle mehr (Typaldos 1999, 2002).

1994 bestand das Faszien-distorsionsmodell aus vier Faszien-distorsionen (Triggerbänder, Hernierte Triggerpunkte, Continuum-distorsionen, Faltdistorsionen). 1995 beinhaltet das FDM alle sechs heute bekannten Faszien-distorsionen<sup>1</sup>. Sein erstes 3-Tages Seminar über das FDM hielt Typaldos 1997 in Österreich an der Wiener Schule für Osteopathie (WSO) (Kasten 2010). Von diesem Zeitpunkt an gab Stephen

---

<sup>1</sup> Die sechs heute im Faszien-distorsionsmodell (FDM) enthaltenen Faszien-distorsionen sind Triggerbänder, Hernierte Triggerpunkte, Continuum-distorsionen, Faltdistorsionen, Zylinderdistorsionen und Tektonische Fixierungen.

Typaldos sein Wissen in Lehrveranstaltungen u.a. in Amerika, Frankreich, Portugal, Deutschland und in Österreich weiter. Heute wird das Faszien-Distorsionsmodell an Osteopathieschulen in Europa und in für Ärzte, Heilpraktiker und Physiotherapeuten zugänglichen Kursen gelehrt (European FDM Association 2011).

#### **4.1.1. Publikationen**

Stephen Typaldos publizierte fortlaufend seine Entwicklungen im Bereich des FDM unter anderem in der *Journal of the American Academy of Osteopathy*. Er schrieb 1996 ein erstes Lehrbuch zum Thema (European FDM Association 2011). Die erste veröffentlichte Auflage seines Buchs „*Orthopathic Medicine: The Unification of Orthopedics with Osteopathy through the Fascial Distortion Model*“ wurde 1997 präsentiert, 1998 folgte die zweite Auflage (Kasten 2010). Die dritte Auflage aus dem Jahr 1999 ist in Deutsch erhältlich. Die 2002 erschienene vierte Auflage mit dem Titel „*FDM - Clinical and Theoretical Application of the Fascial Distortion Model within the Practice of Medicine and Surgery*“ zeichnet das vollständigste Bild seiner Arbeit.

Im Laufe der Zeit entstanden mehrere Diplomarbeiten und eine Doktorarbeit, die sich mit dem FDM und seiner klinischen Anwendung beschäftigten (Rossmly 2005, Geiger 2007, Stein 2008, Engel 2009). In der medizinischen Datenbank PubMed finden sich derzeit (Stand 1/2011) keine Einträge zum Thema FDM, die Datenbank Osteopathic Research Web listet drei der oben angeführten Diplomarbeiten auf. In OSTMED-DR sind teilweise die veröffentlichten und oben erwähnten Artikel von Typaldos ersichtlich.

#### **4.2. Modellvorstellung im FDM**

In der 4. Auflage seines Buches „*FDM - Clinical and Theoretical Application of the Fascial Distortion Model within the Practice of Medicine and Surgery*“ beschreibt Stephen Typaldos (2002) den Inhalt seines Modells wie folgt:

*"The fascial distortion model (FDM) is an anatomical perspective in which underlying etiology of virtually every musculoskeletal injury (and many neurological and medical conditions as well) is considered to be comprised of one or more of six specific alterations of the body's connecting tissue (fascial bands, ligaments, tendons, retinacula, etc.). This model not only allows for strikingly effective manipulative treatments for diverse afflictions*

*such as pulled muscles, fractures, and frozen shoulders, but the results are objective, obvious, measurable, and immediate." (Typaldos 2002:3)*

Das FDM ist ein anatomisches Modell, welches in sechs potentiellen Formveränderungen von Faszie die primäre Ursache von körperlichen Beschwerden und Funktionseinschränkungen sieht. Über das Konzept der Fasziendistorsionen sind aus der Sicht Typaldos eine Reihe von Erkrankungen im Bereich der Orthopädie, der inneren Medizin oder der Neurologie erklärbar. Das ist hilfreich bei der Wahl der optimalen Therapie oder führt zu neuartigen innovativen Behandlungsansätzen, wenn bisherige Behandlungsmethoden nur unbefriedigende Resultate liefern.

Typaldos (2002) sieht im FDM einen Paradigmenwechsel in der Medizin, welcher statt Gewebeschädigung, Gewebsveränderungen sowie daraus resultierenden physiologischen Reparaturprozessen, die anatomische Verformbarkeit der Faszie in den Vordergrund stellt.

Das Fasziendistorsionsmodell ist keine Methode, sondern an sich methodenneutral. Es wird zwischen dem FDM als modellhafte Vorstellung und dessen Anwendung in der manualtherapeutischen Praxis als Typaldos-Methode unterschieden (Harrer 2009). Der Einsatz des FDM ist in vielen Fächern der Medizin möglich. Bekanntermaßen wird es vor allem in der Osteopathie und Physiotherapie angewendet. Vor diesem Hintergrund prägte Typaldos den Begriff „Orthopathie“, der sich aus den Worten Orthopädie und Osteopathie zusammensetzt<sup>1</sup>.

Der Osteopathie liegt ein Vorstellung über Krankheit und Gesundheit zugrunde. Basis dieser Sichtweise ist Stills These von der triune nature des Menschen (Körper, Geist und Seele), deren Harmonie zu Gesundheit führt. Auf struktureller Ebene geht es um ein Gleichgewicht auf faszialer, biomechanischer, nervaler, zirkulatorischer und endokriner Ebene. Manuelle Diagnose- und Therapiemethoden sind die Mittel der Wahl, dieses Gleichgewicht im Erkrankungsfall wiederherzustellen (Liem 2005). Im Gegensatz dazu ist das FDM ein Ansatz, klinische Probleme und Patientenbeschwerden aus anatomischer Sicht zu ergründen:

*„It [The FDM] is developed in an attempt to improve current treatments by basing them on a more anatomical approach. Many of the most commonly*

---

<sup>1</sup> Da nach Ansicht von Typaldos (2002) der Einsatz des FDM im chirurgischen, internen, neurologischen bzw. pharmakologischen Bereich ebenfalls denkbar ist, distanzierte er sich im Lauf der Entwicklung wieder von diesem Begriff.



*seen musculoskeletal injuries are vaguely defined and often respond poorly to conventional treatments. One example of this is a pulled muscle. It is difficult to visualize what a pulled muscle is, and therefore most treatments are not specially designed to correct the underlying dysfunction. In the fascial distortion model a pulled muscle is defined as a muscle that has a triggerband [...]. This definition allows us to conceptualize the pathology and to speculate on how our treatment choices might affect the dysfunction.“ (Typaldos 1994: 14)*

Im Gegensatz zur osteopathischen Basisuntersuchung (vgl. Biberschick 2011) liegt der Schwerpunkt der FDM-Diagnostik nicht auf den palpatorischen Fähigkeiten und Empfindungen des Therapeuten, sondern orientiert sich an den Aussagen und Empfindungen des Patienten. Typaldos (2002) sieht das Faszien-system als ein sensorisches Organ, durch welches periphere Informationen zentral verarbeitet werden können. Dadurch ist es für jeden möglich, genaue Angaben über etwaige Verformungen „seines“ Faszien-systems zu machen.

#### **4.2.1. Faszie aus Sicht des FDM**

Der Ort der Verformung, der Distorsion, ist laut Typaldos die Faszie. Welche Gewebe der Begriff Faszie beinhaltet, darüber gibt es jedoch Auffassungsunterschiede in verschiedenen medizinischen Feldern.

Faszie ist wie andere anatomische Begriffe auch, in ihrer Definition stark von kulturellen und sprachlichen Hintergründen abhängig und es existieren verschiedene Einteilungsschemata, sowohl anatomischer, mikro- und makroskopischer als auch funktioneller Natur (Mirkin 2009). In diesem Sinn sehen Huijing und Langevin (2009) im *Terminus fascia* einen Sammelbegriff für Gewebe von unterschiedlicher Anatomie und Typologie. Sie schlagen 12 Begriffe für die genauere Definition von Faszien-gewebe vor zB *dense connective tissue* und *areolar/non-dense connective tissue*, *superficial fascia* und *deep fascia*. Die Autoren warnen jedoch davor, diese Begrifflichkeiten mit althergebrachten Vorstellungen von Faszie zu vermischen und von diesen Bezeichnungen auf die mikroskopische Struktur der Gewebe, deren Lokalisation im Körper oder ihrer mechanischen Qualitäten rückzuschließen. Zum Beispiel besteht *deep fascia* keineswegs nur aus *dense connective tissue* mit einem hohen Faseranteil, wohingegen *loose connective tissue* (besser: *areolar connective tissue*) sehr wohl Fasern besitzen kann. Ihrer Ansicht nach sind Bänder und Sehnen keine Faszie, können jedoch von ihr durchsetzt sein.

Am 1. Internationalen Fascia Research Congress 2007 wurde folgende Definition präsentiert.

*„Fascia is the soft tissue component of the connective tissue system that permeates the human body. It forms a whole-body continuous three-dimensional matrix of structural support. Fascia interpenetrates and surrounds all organs, muscles, bones and nerve fibres, creating a unique environment for body systems functioning. The scope of our definition of and interest in fascia extends to all fibrous connective tissues, including aponeuroses, ligaments, tendons, retinaculae, joint capsules, organ and vessel tunics, the epineurium, the meninges, the periosteum, and all the endomysial and intermuscular fibres of the myofasciae“ (Findley und Schleip 2007: 2).*

Die Fasziendefinition von Typaldos (2002) ist ähnlich umfassend und Basis für seine grundlegende Überlegung eines *continuity models*, nach welchem einzelne Fasziendfasern den gesamten Organismus durchziehen und so eine einzige strukturelle Einheit bilden. Sie bilden die Basis für Ligamente, Weichgewebe aber auch Knochen.

*„Fascia is found throughout the body and constitutes a tremendous amount of sheer weight and bulk. As the primary connective tissue, it presents in many well-known forms such as tendons, ligaments, retinacula, fascial bands, aponeuroses, adhesions, pericardial sac, pleura, meninges and the perimysium and epimysium of muscles, as well as many other structures. In addition to connecting, fascia surrounds, engulfs, encases, separates, compartmentalizes, divides, protects, insulates, and buffers bones, nerves, muscles and other tissues. In fact, each individual muscle fiber is sheathed with fascia, as is each and every individual muscle bundle, and each and every muscle, as well as every group of muscles (Typaldos 2002: 9)*

Typaldos (2002) unterscheidet Faszie nach anatomischen (die Lokalisation und Form betreffenden) Gesichtspunkten und schreibt diesen einzelnen Kategorien bestimmte Eigenschaften, Funktionen und Möglichkeiten der Distorsion (Verformung) zu. Beispielsweise sieht er bandartiger Faszie als primär auf mechanischen Zug ausgerichtetes Gewebe. Es besteht die Möglichkeit, dass sich die parallel laufenden Fasern durch Scherkräfte voneinander trennen und sich verdrehen bzw. zerknittern. Diese pathologische Konstitution bezeichnet er als Triggerband, eine der sechs Fasziendistorsionen.



Basis dieser Einteilung und Vorstellungen waren anatomische Literatur, Studien am Präparat (Typaldos 1994) oder die Sichtweise von Zeichnern, die Typaldos um Skizzen von Geweben aus dem Seziersaal bat (Kasten 2010). Diese makroskopische Vorgangsweise steht im Gegensatz zur heutigen Fasziensforschung, die sich stark im mikroskopischen und biomechanischen Bereich abspielt, wie die präsentierten Vorträge am ersten und zweiten Internationalen Fascia Research Congress belegen (Findley und Schleip 2007, Huijing et al. 2009).

### **4.3. Typologie der Fasziendistorsionen**

Intuitive Gesten die von Typaldos als Körpersprache bezeichnet werden und subjektive verbale Beschreibungen, sind unter anderem der Schlüssel zur Diagnose der sechs Fasziendistorsionen:

- Triggerbänder (TB, internationale FDM-Abkürzungen jeweils in Klammer)
- Hernierte Triggerpunkte (HTP)
- Continuumdistorsionen (CD)
- Faltdistorsionen (FD)
- Zylinderdistorsionen (CyD)
- Tektonische Fixierungen (TF)

Einige dieser Fasziendistorsionen werden zusätzlich in Subtypen unterteilt und sind teilweise auch in ihrer Körpersprache differenzierbar. Das führt theoretisch zu einem einheitlichen FDM-spezifischen Körpersprache-Vokabular. Körpersprache kann anhand dieses Vokabulars von Osteopathen mit einer Ausbildung im FDM codiert werden. In der Gestikforschung ist das Erstellen eines Gestenvokabulars eine effektive Methode, Bias in der Bewertung von Gesten zu vermeiden (Harrigan 2008).

Im folgenden werden die sechs Fasziendistorsionen, auszugweise ihre wichtigsten Charakteristika und die von Typaldos beschriebene typische Körpersprache vorgestellt. Typaldos bezieht sich dabei hauptsächlich auf Bewegungen und Positionen der Hände. Wenn für die visuelle Diagnose relevant, werden die Subtypen

der einzelnen Distorsionen angeführt. Unter den jeweiligen Beschreibungen sind exemplarisch Fotos von FDM-Körpersprache eingefügt.

#### 4.3.1. Triggerbänder (TB)

Definition, Beschwerden und verbale Beschreibung	Körpersprache
Verformung von bandartiger Faszie, die durch Verdrehen, Trennen, Reissen und Verknittern von Fasern entsteht; Patienten beschreiben meist <b>brennenden oder ziehenden Schmerz</b> entlang eines subjektiv definierten anatomischen Verlaufs; Bewegungseinschränkungen, Kraftverlust oder Schwellungen sind weitere mögliche Indizien für ein Triggerband; sechs Subtypen von Triggerbändern in Lokalisation und Palpation unterscheidbar (Typaldos 1999, 2002)	<b>Streichen mit mehreren Fingern entlang einer druckschmerzhaften Linie</b> (Typaldos 1999, 2002); gestische Demonstration einer Verbindung zwischen zwei Punkten; Subtypen werden mit derselben Körpersprache angezeigt (Typaldos 1999, 2002)



Abb. 2: Beispiel für Triggerband-Körpersprache



Abb. 3: Beispiel für Triggerband-Körpersprache



Abb. 4: Beispiel für Triggerband-Körpersprache

#### 4.3.2. Hernierte Triggerpunkte (HTP)

Definition, Beschwerden und verbale Beschreibung	Körpersprache
Protrusion eines tiefer liegenden Gewebes durch eine darüberliegende Faszienebene; findet sich fast ausschließlich am Körperstamm; kann <b>dumpfe Schmerzen</b> , auch im Umkreis der Herniation verursachen; möglich sind auch Bewegungseinschränkungen der benachbarten Gelenke (Typaldos 1999, 2002)	<b>Drücken mit Fingern, Daumen oder Knöcheln</b> in das Protrusions-Areal (Typaldos 1999, 2002)



Abb. 5: Beispiel für Körpersprache bei Herniertem Triggerpunkt



Abb. 6: Beispiel für Körpersprache bei Herniertem Triggerpunkt

### 4.3.3. Continuumdistorsionen (CD)

Definition, Beschwerden und verbale Beschreibung	Körpersprache
Distorsion der Übergangszone zwischen Knochen und Ligament, Sehne oder anderer Faszienstruktur; <b>punktueller Schmerzen</b> oder Bewegungseinschränkungen in einer Achse als Folge; es werden zwei Subtypen (invertierte und evertierte Continuumdistorsionen) unterschieden (Typaldos 1999, 2002)	<b>Zeigen mit einem Finger</b> auf einen oder mehrere knochen nahe Schmerzpunkte; keine Unterschied in der Körpersprache der Subtypen (Typaldos 1999, 2002)



Abb. 7: Beispiel für Körpersprache bei Continuumdistorsion





Abb. 8: Beispiel für Körpersprache bei Continuumdistorsion

#### 4.3.4. Faltdistorsionen (FD)

Definition, Beschwerden und verbale Beschreibung	Körpersprache
<p>3-Dimensionale Verformung der gelenkumhüllenden Faszie, der intermuskulären Septen oder der interossären Membranen; Leitsymptome sind <b>Bewegungsschmerzen, Schmerzen in der Tiefe, Instabilitätsgefühl</b>; verursachen kaum Bewegungseinschränkungen; Unterscheidung zwischen <b>Einfalt- und Entfaltdistorsionen</b> (Typaldos 1999, 2002)</p>	<p><b>Halten, Kneten oder Ziehen des Gelenks mit der Hand oder den Händen</b>; im Bereich des Rückens werden (je nach Erreichbarkeit) der Handrücken auf die betroffene Stelle gelegt, Fäuste in den Rücken gedrückt, der Rücken gestützt oder mit Hilfe der Hände entlastet;</p> <p><b>Einfaltdistorsion: queres Streichen mit den Fingern über den Gelenkspalt;</b></p> <p><b>Interossäre Membranen und intermuskuläre Septen: Drücken mehrerer Fingerspitzen zwischen Muskelbäuche oder Knochen, Ziehen oder Zerren an den Weichteilen</b>; Faltdistorsionen werden oft dynamisch mit repetierter Bewegung des betroffenen Körperteils angezeigt (Typaldos 1999, 2002)</p>



Abb.9: Beispiel für Körpersprache bei Faltdistorsion



Abb. 10: Beispiel für Körpersprache bei Faltdistorsion



Abb. 11: Beispiel für Körpersprache bei Faltdistorsion



Abb.12: Beispiel für Körpersprache bei Faltdistorsion (intermuskuläres Septum)

#### 4.3.5. Zylinderdistorsionen (CyD)

Definition, Beschwerden und verbale Beschreibung	Körpersprache
Verformung der spiralförmigen Windungen der zirkulären Faszie; <b>möglich sind Symtome wie Parästhesien, Krämpfe, Zittern, Lähmungserscheinungen usw.</b> ; oft besteht eine große Diskrepanz zwischen Auslöser und den vorhanden Beschwerden; die <b>Schmerzen</b> werden „ <b>in der Tiefe</b> “ empfunden, sind von aussen meist nicht palpabel (Typaldos 1999, 2002)	<b>Kneten oder flächiges Streichen der Weichteile</b> ; wischende Handbewegungen, Zwicken, Drücken der Weichteile u. a. (Typaldos 1999, 2002); gelegentlich wird die Körpersprache nur mit Abstand zum Körper dargestellt und der betroffene Körperteil gar nicht berührt



Abb. 13: Beispiel für Körpersprache bei Zylinderdistorsion





Abb. 14: Beispiel für Körpersprache bei Zylinderdistorsion

#### 4.3.6. Tektonische Fixierungen (TF)

Definition, Beschwerden und verbale Beschreibung	Körpersprache
Verlust der Gleitfähigkeit faszialer Oberflächen; verursachen in erster Linie kaum Schmerz, hingegen <b>Mobilitätseinschränkungen</b> in allen Ebenen und das Gefühl von <b>Steifheit</b> (Typaldos 1999, 2002)	Patienten versuchen sich mit Kraft selbst zu mobilisieren; im Bereich der Schulter Abduktion nur mit anteriorer Elevation möglich, in Bauchlage kann der Arm nicht horizontal nach vorn gestreckt werden, ohne dass sich der Oberkörper von der Unterlage abhebt; Umklammern beider Cristae iliaca bei Hüftsteife; <b>Halten und Kneten des Gelenks</b> (Typaldos 1999, 2002)



Abb. 15: Beispiel für Körpersprache bei Tektonischer Fixierung

#### 4.4. Die FDM-Diagnose

Als Erweiterung der schulmedizinischen Diagnose ist nach Typaldos (1999) die FDM-Diagnose zu betrachten, um erfolgreich nach dem FDM behandeln zu können. Die FDM-Diagnose beruht neben einer schulmedizinischen Abklärung und der Einbeziehung von relevanten Laboruntersuchungen oder bildgebenden Verfahren auf folgenden Komponenten:

- Körpersprache des Patienten:

Die Körpersprache wird von Patienten in der Therapiesituation während der Anamnese oder im Lauf der Behandlung gezeigt. Diese „Zeichen“ müssen vom Therapeuten als solche interpretiert und anhand des FDM-Körpersprache-Vokabulars den passenden Faszienstörungen zugeordnet werden.

- Subjektive verbale Beschreibung der Beschwerden (u.a. Schmerzqualität, Ursache, Verlauf):

Prinzipiell kann ein Großteil der subjektiven, verbalen Patientenäußerungen ebenfalls einzelnen Faszienstörungen zugeordnet werden. Die Verbalisation kann ein entscheidender Hinweis für die Diagnose sein, die inhaltliche Aussage der Körpersprache unterstützen oder ihr entgegenstehen. Bestimmte Worte gelten als direkter Hinweis auf eine bestimmte Faszienstörung<sup>1</sup>.

---

<sup>1</sup>Beispielsweise deutet die verbale Aussage von Schmerzen „im Gelenk“ auf eine Faltdistorsion hin.

- Körpertests: Bewegungs- und Belastungstests

Sie ergänzen die FDM-Diagnostik in der Weise, dass dadurch ein Patientenstatus vor der Behandlung erhoben wird, der nach der Behandlung auf Verbesserung überprüft wird. Zudem sind bestimmte Ergebnisse dieser Tests Hinweise auf spezielle Faszienstörungen.

- Palpation:

Sie ist wichtig für die Differentialdiagnose, da einige Faszienstörungen druckschmerzhaft sind und andere nicht. Ausserdem wird sie zur genauen Lokalisation von einzelnen Faszienstörungen verwendet, was vor allem in der Behandlung wichtig ist.

Diese Informationen werden in der Befundung miteinander vernetzt und führen zu einer FDM-Diagnose, bestehend aus einer oder mehreren ursächlichen Faszienstörungen, die nacheinander spezifisch korrigiert werden (Typaldos 1999, 2002). Die FDM-Körpersprache wird inhaltlich dahingehend interpretiert, dass in ihr ein therapeutischer Ansatz zur Lösung der Patientenbeschwerden erkennbar ist. Beispielsweise ist eine streichende Bewegung der Finger entlang einer Linie am eigenen Körper, Hinweis auf ein Triggerband. In ähnlicher Art und Weise, wie der Patient dies gestisch darstellt, erfolgt die Korrektur durch ein kräftiges Ausstreichen des Triggerbands entlang dieser Linie.

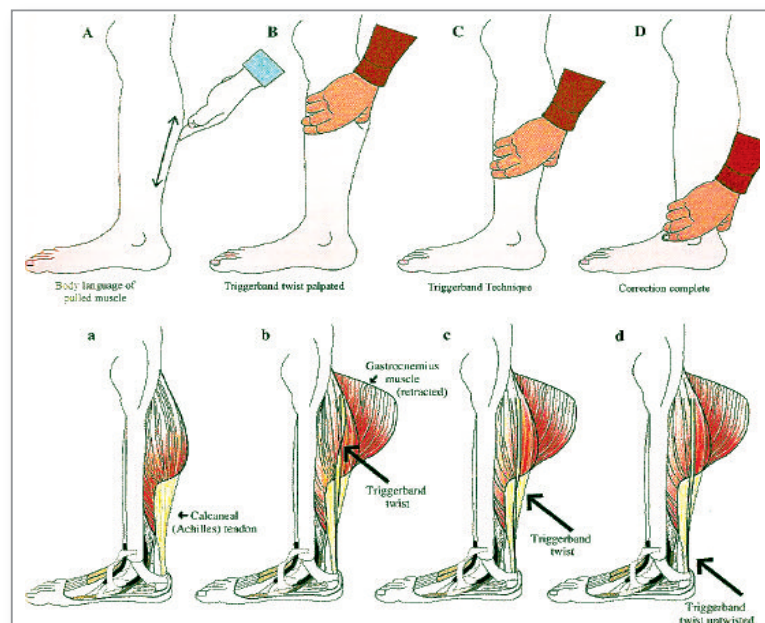


Abb. 16: Der Zusammenhang zwischen FDM-Körpersprache und der spezifischen Behandlung der jeweiligen Faszienstörung anhand des Triggerbandes (Typaldos 2002)

## 5. Fragestellung

Die Fragestellung zielt darauf ab, durch eine wissenschaftlich methodologische Untersuchung, Information über die Interrater-Reliabilität der visuellen FDM-Diagnostik zu erhalten.

Das FDM definiert Kriterien für die visuelle Diagnostik von bestimmten Faszien-distorsionen. In der Praxis liegt die Herausforderung darin, diese spezifische Körpersprache innerhalb der Befunderhebung ausfindig zu machen und entsprechend zu codieren.

### Fragestellung

Wie verlässlich beurteilen Osteopathen mit einer Grundausbildung im FDM die Körpersprache im Sinne des Faszien-distorsionsmodells? Wie hoch ist die Interrater-Reliabilität?

### Zusätzlich sollen weitere Fragen beantwortet werden

- Beurteilen Osteopathen mit verschiedenen Quellberufen (Arzt, Heilpraktiker, Physiotherapeut) die FDM-Körpersprache an Patienten gleich übereinstimmend?
- Beurteilen Osteopathen mit unterschiedlich langer Erfahrung mit dem FDM die Körpersprache an Patienten gleich übereinstimmend?
- Beurteilen Osteopathen die FDM-Körpersprache, die gleichzeitig mit verbalen Aussagen des Patienten auftritt, im Gegensatz zu FDM-Körpersprache, die ohne verbale Kommunikation seitens des Patienten abläuft, gleich übereinstimmend?

Die Ergebnisse der Interrater-Reliabilität im Zusammenhang mit dem Quellberuf der teilnehmenden Osteopathen und deren Erfahrungsdauer mit dem FDM sind von Interesse, weil die in Kapitel 2 präsentierten Reviews keine Korrelation dahingehend finden können. Da diese Studie statt eines manuellen einen visuellen Diagnoseansatz verwendet, ist zu klären, ob die Ergebnisse in diesen Punkten abweichen.

Der Einfluss von verbalen Aussagen des Patienten auf die Reliabilität visueller Diagnostik soll zusätzlich untersucht werden, um Information über die Beeinflussung

der visuellen Diagnostik durch verbale Äußerungen des Patienten zu erlangen. In der Praxis schildern Patienten in vielen Fällen ihre Beschwerden gleichzeitig sowohl verbal als auch nonverbal. Deshalb werden aus Gründen der Praxisrelevanz die Videos im Online-Fragebogen mit dem Originalton verarbeitet. Abgesehen davon werden drei Videopaarungen mit und ohne Ton in den Fragebogen eingearbeitet, um explizit mögliche Unterschiede in der Interater-Reliabilität zu erforschen.

### **Arbeitshypothese**

Wenn die visuelle Diagnostik des Faszienmodells

- ein reliabler Parameter für die Diagnose von Faszienstörungen darstellt,
- und die Kriterien für das Vorliegen einer Faszienstörung ausreichend definiert sind,

können Osteopathen mit Kenntnis der FDM-Diagnostik die Körpersprache nach den Kriterien des Faszienmodells an mehreren Patienten übereinstimmend bewerten.

### **Nullhypothesen**

Bewerten Osteopathen mit einer FDM-Grundausbildung Körpersprache im Sinne des Modells an mehreren Patienten, erreicht deren Verlässlichkeit keine akzeptablen Werte ( $k < 0,40$ ).

Der Quellberuf beeinflusst die Verlässlichkeit, mit der Osteopathen FDM-Körpersprache an Patienten beurteilen, nicht.

Die Dauer der Erfahrung mit dem FDM beeinflusst die Verlässlichkeit, mit der Osteopathen FDM-Körpersprache am Patienten beurteilen, nicht.

Verbale Kommunikation während der Produktion von Körpersprache seitens des Patienten beeinflusst die Verlässlichkeit, mit der Osteopathen FDM-Körpersprache am Patienten beurteilen, nicht.

## 6. Methodologie

Diese Masterthese untersucht mit Hilfe eines videounterstützten Online-Fragebogens die Interrater-Reliabilität der Körpersprache-Diagnostik nach den Kriterien des FDM. Die in Kapitel 2 dieser Arbeit beschriebenen Kritikpunkte an Reliabilitätsstudien im Bereich manueller Palpationstests sind Leitlinie für die Frage, inwieweit die vorliegende Studie die Kriterien für Qualität erfüllen kann. Das bietet den Hintergrund für die kritische Betrachtung der Ergebnisse und wird im Diskussionsteil nochmals reflektiert.

Die Kritikpunkte (vgl. Kap. 2.3.1.) werden im Folgenden in ihrer Konsequenz für das vorliegende Studiendesign erläutert:

- Ein unabhängiger Referenzstandard ist für die Diagnostik im FDM derzeit nicht verfügbar. Abgesehen vom Konzept des Hernierten Triggerpunkts, der in der Schulmedizin in bestimmten Fällen als Hernie eine beschriebene Pathologie darstellt (vgl. Kap. 3.6.), können Faszienstörungen nicht durch andere Messmethoden alternativ verifiziert werden.
- Die visuelle FDM-Diagnostik ist durch die relativ klaren Beurteilungskriterien prinzipiell gut standardisierbar. Ziel und Intention der Körpersprache-Diagnostik sind konstante Grundinhalte der FDM-Ausbildung und können von den ausgewählten Testern mit abgeschlossener Ausbildung vorausgesetzt werden. Die prinzipiellen Kriterien für die Diagnose der einzelnen Faszienstörungen sind im FDM gleichbleibend und ebenfalls Basiswissen. Der Untersuchungsablauf an sich ist durch das Studiendesign als videounterstützter Onlinefragebogen weiter standardisiert. Jeder Tester hat dadurch dieselben Ausgangsbedingungen gegenüber den gezeigten Patienten und den anderen Testern. Zudem kann damit der Ablauf der Untersuchung vorgegeben werden.
- Die Studienpopulation umfasst symptomatische Patienten in dem Sinne, dass alle gefilmten Patienten (vor der randomisierten Auswahl der Videos für den Online-Fragebogen) körperliche Beschwerden haben und Körpersprache zeigen. Ob diese Körperbewegungen den Kriterien für eine der sechs Faszienstörungen entsprechen, wird nicht in der Auswahl der Patienten berücksichtigt.

## **6.1. Erhebungsmaterial**

In der vorliegenden Masterthese werden Videoaufnahmen von Patienten verwendet, um mit deren Hilfe die Interrater-Reliabilität der FDM-Körpersprache zu untersuchen. Der Ablauf der Testung erfolgt mittels eines Online-Fragebogens.

### **6.1.1. Videoaufnahmen**

Videos als Basis für Reliabilitätsstudien werden vor allem im Bereich des Sports und der Physiotherapie verwendet. Hier wird Video beispielsweise bei der Analyse von Bewegungsabläufen eingesetzt. Nach einer Suche in Online-Datenbanken (vgl. Tab.1) mit den Begriffen *video* OR *video analysis* AND *reliability* AND *sport* fanden sich 47 Einträge, bei der Kombination mit *physical therapy* 37 Einträge. In Kombination mit *osteopath\** konnten keine relevanten Treffer gelandet werden.

Video ermöglicht wiederholtes Messen unter stabilen Bedingungen. Zudem haben Videoaufnahmen den Vorteil, dass in kurzer Zeit aufeinander folgende Bewegungsabläufe gefilmt und anschließend in Einzelabschnitte geteilt werden können (Haidet et al. 2009). Dieser Punkt ist für diese Masterthese im Hinblick auf die interne Validität und Vermeidung von internen Fehlerquellen interessant, weil die Videos in der Nachbearbeitung teilweise in Sequenzen geteilt werden. Nachteile von Videoaufnahmen können einerseits die Beeinflussung des gefilmten Patienten durch die ungewohnte Aufnahmesituation sein, andererseits der teilweise durch die Fokussierung der Kamera verloren gehende Aspekt der Interaktion des Patienten mit der Umwelt (Haidet et al. 2009).

#### **6.1.1.1. Aufnahmesituation**

Um einerseits die Osteopath-Patient-Kommunikation während der Videoaufnahmen so praxisnah wie möglich zu gestalten, andererseits die Aufnahmen in ihrem Ablauf zu standardisieren, werden die Videos vom Untersuchungsleiter in der eigenen osteopathischen Praxis in Wien selbst gefilmt. Für die Entscheidung, keine zusätzliche Person für die Aufnahmetätigkeit heranzuziehen, spricht folgendes Argument:

Durch das Fehlen einer dritten Person wird eine klar definierte Osteopath-Patient-Gesprächssituation geschaffen und der Patient hat einen klaren Ansprechpartner für die Präsentation seiner Beschwerden und der Körpersprache. Diese Vor-



gangsweise wird durch die Erkenntnisse von Harrigan (2008), Stukenbrock (2008) und Haidet et al. (2009) unterstützt.

Es wird eine handelsübliche Mini-DV-Kamera verwendet. Zuerst werden die Patienten über die Art und Weise der Untersuchung und die Rahmenbedingungen der Videoaufnahme informiert, der sie schriftlich zustimmen. Auf die Standardfrage „Zeigen Sie mir, wo es weh tut“ bzw. „Zeigen Sie mir, wo das Problem liegt“ werden die Patienten im folgenden bei der nonverbalen Beschreibung ihrer körperlichen Beschwerden gefilmt. Die Aufnahme dauert so lange, bis der Patient von sich aus mit seiner Beschreibung endet.

#### **6.1.1.2. Auswahl der gefilmten Patienten**

Die Videoaufnahmen finden an vier aufeinanderfolgenden Arbeitstagen im Juni 2010 in Wien statt. 48 Patienten kommen in dieser Zeit in die Praxis, davon erklären sich 35 bereit, an der Aufzeichnung teilzunehmen. Die Auswahlkriterien für die gefilmten Patienten sind weiters die schriftliche Zustimmung für die Videoaufzeichnung am Beginn der osteopathischen Behandlung. Die Videos werden den Patienten zusätzlich nach der Aufnahme nochmals gezeigt und ihre Zustimmung für die Verwendung eingeholt.

#### **6.1.1.3. Nachbearbeitung der Videos**

Von Beginn der Arbeit an dieser Masterthese, steht die Frage im Raum, ob und inwieweit die Videoaufnahmen bearbeitet werden dürfen, damit sie einerseits ein realistisches Szenario der Körpersprache-Diagnose in der Praxis widerspiegeln (externe Validität), andererseits interne Fehlerquellen für die Studie minimiert werden können (interne Validität). In diesem Spannungsverhältnis werden die 35 aufgezeichneten Videos nach folgenden Kriterien bearbeitet und aus dem Basisvideomaterial 50 Kurzvideos geschnitten.

- Die Videos werden teilweise in kürzere Sequenzen geteilt. Der Zeitpunkt des Schnitts wird vom Untersuchungsleiter dann gewählt, wenn die Patienten a) Körpersprache an einem Körperteil beenden und zu einem anderen Körperteil übergehen, b) ihre Position im Raum bzw. zum Untersuchungsleiter (Videofilmer) verändern, c) verschiedenartige Sequenzen von Körpersprache in sehr kurzer Zeit präsentieren, die in der Beurteilung durch die Tester nicht voneinander trennbar und damit eindeutig bewertbar sind. Diese

Möglichkeit der Sequenzierung von Videos beschreibt auch Haidet et al. (2009) als Vorteil in Reliabilitätsuntersuchungen.

- Die Videos werden im Fragebogen langsam ein- und ausgeblendet, um den Testern die Möglichkeit zu bieten, den Höhepunkt der Gestik (*peak / stroke*) erkennen zu können und damit den relevanten Teil der Körpersprache zu bewerten (vgl. Kapitel 3.2.2.)
- Der Videoton wird, wenn die Patienten relevante akustische Informationen mitteilen (abgesehen von Räuspern, Husten, u.a.) nachbearbeitet. Dabei werden die Lautstärke und das Hintergrundgeräusch für eine bessere akustische Verständlichkeit der verbalen Aussagen angepasst.
- Bei den Videos (v7, v16, v30) wird der Ton komplett unterdrückt, um zu einem späteren Zeitpunkt dieselben Videos (v8, v22, v33) im Fragebogen nochmals mit dem Originalton zu präsentieren. Damit soll die Auswirkung von verbalen Äußerungen auf die Bewertung der FDM-Körpersprache untersucht werden (vgl. Kapitel 6.1.1.4.).

#### **6.1.1.4. Auswahl der verwendeten Videos für den Fragebogen**

Von den 50 durchnummerierten Patienten-Kurz-Videos werden randomisiert 23 Videos für den Online-Fragebogen ausgewählt. Die Randomisierung erfolgt durch zufällige Nennung von Zahlen zwischen 1 und 50 durch eine dritte, unbeteiligte Person. Zusätzlich werden von diesen 23 Videos drei Videos randomisiert ausgewählt, um sie mit und ohne Originalton im Fragebogen zu präsentieren. Zusätzliches Auswahlkriterium ist hier, dass sich auf den Originalvideos die gefilmten Patienten verbal äußern. Weiters werden zwei Videos, die keine nach den FDM-Kriterien bewertbare Körpersprache enthalten, in den Fragebogen aufgenommen. Damit soll erforscht werden, ob die Tester jegliche Körperbewegungen als FDM-Körpersprache erkennen oder sich auf das vorhandene Körpersprache-Vokabular des FDM beschränken (vgl. Kap. 6.1.2.2.). Aus den restlichen Videos werden weiters zwei Videos für die Übungsphase des Online-Fragebogens ausgewählt. Sie werden nicht in die statistische Auswertung der Ergebnisse miteinbezogen.

#### **6.1.2. Onlinefragebogen**

Für diese Masterthese wird ein Online-Fragebogen verwendet. Die Vorteile dieses Ansatzes für diese Studie sind:

- Standardisierung: Alle Tester sehen dieselben Patienten, in derselben Reihenfolge und aus demselben Blickwinkel.
- Messwiederholung: Um exaktes Messen zu ermöglichen, können die Tester die Videos mehrmals unter denselben Bedingungen ansehen.
- Herkunftsort der Tester: Durch den Online-Fragebogen ist es möglich, Osteopathen als Tester in verschiedenen Ländern anzusprechen. Es sind keine Gruppentermine für die Untersuchung notwendig, da jeder Tester zu dem für ihn optimalen Zeitpunkt die Umfrage durchführen kann.

#### **6.1.2.1. Abwicklung der Befragung**

Die Firma „Onlineumfragen“ mit Hauptsitz in CH-6055 Alpnach ist ein 1999 gegründetes Unternehmen mit Spezialisierung auf die Durchführung von Onlineumfragen (Onlineumfragen 2011). Für diese Masterthese bietet sich somit die Möglichkeit, einen Online-Fragebogen zu entwickeln, an die Testpersonen weiterzuleiten, deren Testergebnisse zu sammeln und anschließend statistisch auszuwerten. Dies erfolgt weitgehend selbständig über eine Online-Maske der Firma. Hilfestellung leistet die Firma bei folgenden technischen Fragen:

- Die Videos sollen für den Tester nur einmal beurteilbar sein und somit ein nachträgliches Verändern der Ergebnis-Auswahl verhindert werden. Damit wird einerseits der Vorteil von Videoaufnahmen in Reliabilitätsstudien schlagend, dass wiederholtes Ansehen der Videos möglich ist (Haidet et al. 2009) und es entspricht auch den Anforderungen an exaktes Messen (Sommerfeld 2006). Andererseits ist dieser Punkt Voraussetzung für die Befragung nach dem Einfluss von Verbalisation auf die Beurteilung der FDM-Körpersprache in der vorliegenden Form (vgl. Kap. 6.1.2.2.).
- Es soll im Verlauf des Fragebogens (ab Frage v6) für die Tester nicht möglich sein, eine Frage unbeantwortet zu überspringen. Damit soll die Rücklaufquote an bewerteten Videos erhöht werden und interne Fehlerquellen durch Überspringen einzelner Videos vermieden werden.

#### **6.1.2.2. Aufbau des Online-Fragebogens**

Zu Beginn des Fragebogens steht eine Begrüßung und Erklärung des Ziels dieser Untersuchung. Der Zeitaufwand für die Umfrage wird mit 20 Minuten beziffert.

Danach folgt die Beschreibung des Untersuchungsschemas. Die Tester werden aufgefordert, sich Zeit zu lassen und die Lautstärke ihres Computers auf ein angenehmes Niveau zu stellen. Es wird erklärt, dass manche Videos doppelt vorkommen, in manchen Videos verbale Äußerungen von Patienten zu hören sind, in anderen nicht. Der Detailablauf einer Frage wird erklärt:

1. Das Video wird geladen
2. Das Video wird vom Tester angesehen, eine Wiederholung ist möglich.
3. Der Tester entscheidet sich für **eine** Antwortmöglichkeit (*single choice*).

Zudem wird darauf hingewiesen, dass die Abgabe einer Antwort Voraussetzung für das Fortfahren in der Umfrage ist, und dass angeklickte Antworten fixiert sind und nachträglich nicht verändert werden können.

Im Anschluss startet die Untersuchung.

Der Online-Fragebogen besteht aus 33 Fragen.

*Frage v1: Welchen Beruf üben Sie aus? Bei mehreren Antwortmöglichkeiten bitte ich Sie, die höchste universitäre Ausbildung anzugeben.*

Nominale Antwortmöglichkeiten sind ÄrztIn, PhysiotherapeutIn und HeilpraktikerIn. Die Antwortmöglichkeiten ergeben sich aus den Zugangsvoraussetzung für die Osteopathie-Ausbildung in Deutschland, Österreich und der Schweiz. In Österreich sind das Ärzte und Physiotherapeuten (Wiener Schule für Osteopathie 2011). Da in Deutschland und der Schweiz auch Heilpraktiker zur Osteopathie-Ausbildung zugelassen sind, wird dieser Titel in die Auswahlmöglichkeiten aufgenommen. Die Ausbildung in einem dieser drei Quellberufe ist Voraussetzung für die FDM-Ausbildung (European FDM Association 2010).

Hintergrund dieser Frage ist es, eventuelle Unterschiede in der Interrater-Reliabilität bei Osteopathen mit unterschiedlichen Quellberufen zu erforschen. Die in dieser Arbeit präsentierten Reviews können derartige Zusammenhänge generell nicht feststellen (Seffinger et al. 2004, Stockendahl et al. 2006).

*Frage v2: Sind Sie Osteopath?*

Diese Frage kann nominal mit JA oder NEIN beantwortet werden. Einschlusskriterium für die Auswahl der Tester ist es, Osteopath zu sein. Nur die Ergebnisse der Osteopathen werden statistisch verarbeitet. Nicht-Osteopathen werden aufgefor-

dert, die Untersuchung fortzusetzen, da ihre Ergebnisse für eine über diese Studie hinaus gehende statistische Analyse der Daten von Interesse sein könnten.

*Frage v3: Wie viele Jahre arbeiten Sie mit dem FDM? – Bitte runden Sie auf ganze Jahre auf oder ab.*

Eine Teilfrage dieser Masterthese ist es, ob es einen Zusammenhang zwischen der Interrater-Reliabilität und der Erfahrung der Tester gibt. Anhand eines metrischen Skalenniveaus in Jahresschritten geben die Teilnehmer der Studie ihre Erfahrungsdauer an. Die Erfahrung soll vom Tester auf- bzw. abgerundet werden, wenn keine genauen Jahresschritte möglich sind.

Seffinger et al. (2004) und Stochkendahl et al. (2006) können in ihren Reviews einen derartigen Zusammenhang nicht feststellen.

*Frage v4 und v5: Probevideos – So funktioniert der Test. Welche Faszien­distorsion ist auf diesem Video erkennbar?*

Diese Fragen sind als Übungsfragen für die folgenden, gewerteten Patientenvideos geplant. Die Fragen v4 und v5 werden nicht in die statistische Auswertung mit­einbezogen.

Sowohl in der Probesituation als auch in der Testsituation sollen sich die Tester nach dem Ansehen des Videos für eine der folgenden Antwortmöglichkeiten entscheiden. Wird keine Antwort angeklickt, ist eine Weiterführung der Untersuchung nicht möglich.

Antwortmöglichkeiten sind:

- Triggerband TB
- Hernierter Triggerpunkt HTP
- Continuumdistorsion CD
- Faltdistorsion FD
- Zylinderdistorsion CyD
- Tektonische Fixation TF
- Keine Faszien­distorsion
- Weiß nicht

Zur Auswahl stehen die sechs Faszienstörungen, die Typaldos (1999, 2002) beschreibt. Auf die Subtypen einzelner Störungen wie zB Einfalt- oder Entfaltdstörungen wird aus Gründen der internen Validität verzichtet. Zudem ist die genaue Bestimmung von diesen Subtypen in der Praxis oft nur in Zusammenhang mit der verbalen Anamnese oder Bewegungs- und Belastungstests möglich.

„Keine Faszienstörung“ soll angekreuzt werden, wenn der Tester keine diagnostizierbare Faszienstörung erkennt. Diese Antwortmöglichkeit bezieht sich auf die Erkenntnisse aus Kapitel 3.2. über die charakteristischen Elemente von Gesten. Obwohl keine eindeutige Definition von FDM-Körpersprache als Gestik existiert, soll diese Frage erforschen, inwieweit die Tester Patientenbewegungen als Körpersprache bzw. Geste erkennen und bewerten und sie das abgegrenzte Körpersprache-Vokabular des FDM in der Praxis anwenden.

„Weiß nicht“ sollte angekreuzt werden, wenn der Tester die Frage nicht beantworten kann.

Nach den Probevideos werden die Tester nochmals auf den Untersuchungsablauf und die Bewertungskriterien aufmerksam gemacht (siehe oben). Mit der Frage v6 beginnt die Testphase der Untersuchung.

Frage v6 bis v33: Welche Faszienstörung ist auf diesem Video erkennbar?

Wie im oben beschriebenen Probelauf beurteilen die Tester die Videos v6 bis v33. Diese Fragen werden in die Statistik der Interrater-Reliabilitätsuntersuchung miteinbezogen.

Frage v7/v16/v30: Welche Faszienstörung ist auf diesem Video erkennbar?

Diese Fragen unterscheiden sich in ihrem formalen und technischen Ablauf nicht von den restlichen Fragen. Die entsprechenden Videos werden an dieser Stelle ohne Ton gezeigt und von den Testern beurteilt. Zu einem jeweils späteren Zeitpunkt werden im Fragebogen dieselben Videos mit Originalton gezeigt (v8, v22, v33).

Diese Art der Frage zielt darauf ab, den Zusatzeffekt der Verbalisation auf die Beurteilung der Körpersprache zu untersuchen. Die dadurch entstehenden Fragenpaare werden in zufälliger Reihenfolge in den Online-Fragebogen eingearbeitet. Im Fragebogen ist es nicht möglich, bereits beantwortete Fragen zu verändern. Damit sind die Voraussetzungen für diese Art der Fragestellung geschaffen.

## 6.2. Tester

Die Tester, Osteopathen aus Deutschland, Österreich und der Schweiz erfüllen folgende Auswahlkriterien:

### 6.2.1. Auswahl der Tester

- Abgeschlossene FDM-Ausbildung in drei Teilen nach den Richtlinien der European FDM Association<sup>1</sup> (European FDM Association 2010) mit in der Folge positiv absolvierter Zertifizierungsprüfung bzw. eine absolvierte FDM-Instruktorenprüfung:

Dieses Kriterium dient dazu, einen ausreichenden und einheitlichen Kenntnisstand in der visuellen Diagnostik von Faziendistorsionen sicherzustellen. Basis für die Auswahl der Tester ist die Therapeutenliste der European FDM Association (European FDM Association 2010), die zum Zeitpunkt dieser Studie 79 zertifizierte FDM-Therapeuten auflistet. Voraussetzung für die Eintragung auf dieser Liste ist eine positiv absolvierte Zertifizierungsprüfung bzw. FDM-Instruktorenprüfung. Zudem ist die Mitgliedschaft in der EFDMA notwendig. Zum Zeitpunkt des Zugriffs sind in der Liste Therapeuten aus Deutschland, Österreich und der Schweiz vertreten.

- Ausbildung zum Osteopathen:

Osteopathen sind die relevante Tester-Zielgruppe für diese Masterthese. Anhand der Therapeutenliste kann nicht geklärt werden, welche der angeführten Personen in Osteopathie ausgebildet sind, da nur deren Quellberuf in der Liste aufgeführt werden. Deshalb wird explizit im Fragebogen nach dieser Ausbildung gefragt (vgl. Kapitel 6.1.2.2.).

- Der deutschen Sprache mächtig:

Der Online-Fragebogen wird in Deutsch verfasst, die gefilmten Patienten sprechen Deutsch und teilweise sind deren verbale Äußerungen für die Codierung der Videos wichtig.

---

<sup>1</sup> Die FDM-Grundausbildung besteht aus drei Modulen. Inhalte sind die Modellvorstellung des FDM, die Anwendung und Integration des FDM in der klinischen Praxis sowie die Untersuchung und Behandlung aller Körperabschnitte (European FDM Association 2010).



## 6.2.2. Demografische Angaben

In den vorgestellten Reviews zum Thema Reliabilität manueller Tests an der Wirbelsäule kann keine Korrelation zwischen beruflichen Hintergrund der Tester und deren Erfahrung in Bezug auf die Reliabilität hergestellt werden.

Die Tester dieser Untersuchung werden hinsichtlich dieser beiden Kriterien untersucht und zugeordnet. Damit soll erforscht werden, ob sich im Bereich der visuellen FDM-Diagnostik die Erkenntnisse von Seffinger et al. (2004) und Stochkendahl et al. (2006) bestätigen lassen.

### 6.2.2.1. Geschlecht

Der Anteil der männlichen Tester ist bei den 26 Osteopathen, die den Online-Fragebogen starten, überproportional hoch.

Tester	Männer	Frauen
... welche die Online-Untersuchung starten	21	5
... welche die Online-Untersuchung beenden	20	5

Tab. 2: Das Verhältnis von männlichen und weiblichen Testern in der Studie

Ein männlicher Tester (MvE) beginnt die Untersuchung, beendet diese jedoch nach Frage v3. Er wird in der statistischen Ausarbeitung dieser These nicht berücksichtigt. Damit nehmen vier mal mehr männliche Tester an der Studie teil als weibliche.

### 6.2.2.2. Quellberuf

23 der 26 Tester geben einen Quellberuf gemäß den Auswahlmöglichkeiten der Umfrage an. Unerwarteterweise sind zwei männliche Tester im Quellberuf Masseur (davon ein Tester, der die Umfrage mit der Frage v3 beendet), eine weibliche Testerin ist Hebamme.

Da alle 25 Tester (welche die Online-Umfrage regulär beenden) Osteopathen sind, werden ihre Ergebnisse in die Statistik miteinbezogen. In Bezug auf den Zusammenhang Quellberuf und Reliabilität werden die Ergebnisse dieser drei (respektive zwei) Tester nicht berücksichtigt.

Quellberuf	Arzt (1)	Physiotherapeut (2)	Heilpraktiker (3)	Sonstige
Anzahl der Tester (n=25)	9	10	4	2

Tab. 3: Quellberuf der Tester, welche die Online-Umfrage regulär beendet haben

Ärzte sind überproportional männlichen Geschlechts, Physiotherapeuten überproportional weibliche Tester.

Quellberuf	Arzt (1)	Physiotherapeut (2)	Heilpraktiker (3)	Sonstige
Männer (n=20)	8	7	4	1
Frauen (n=5)	1	3		1

Tab. 4: Quellberuf in Zusammenhang mit dem Geschlecht der Tester

### 6.2.2.3. Erfahrung mit dem FDM

Ein Großteil der Tester hat zwei bis fünf Jahre Erfahrung mit dem FDM. Der Zusammenhang mit dem Geschlecht der Tester und deren Quellberuf wird in Tabelle 5 und Tabelle 6 dargestellt.

Jahre (auf Ganze gerundet)	Männer (n=20)	Frauen (n=5)	Anzahl der Tester gesamt (n=25)
1	0	0	0
2	4	2	6
3	9	1	10
4	3	0	3
5	2	2	4
6	(1) <sup>1</sup>	0	(1) <sup>1</sup>
7	1	0	1
13	1	0	1

Tab. 5: Erfahrung mit dem FDM, Anzahl der Tester und deren Geschlecht im Überblick

Erfahrung in Jahren	Anzahl der Tester	Arzt (1)	Physiotherapeut (2)	Heilpraktiker (3)	Sonstige
1					
2	6	2	2	1	1
3	10	3	5	1	1
4	3	1	1	1	
5	4	2	2		
6	(1) <sup>1</sup>				(1) <sup>1</sup>
7	1			1	
13	1	1			

Tab. 6: Erfahrung mit dem FDM in Zusammenhang mit dem Quellberuf der Tester

<sup>1</sup> Tester (MvE) beendete die Online-Umfrage nach der dritten Frage.

### 6.2.3. Rücklaufquote und Untersuchungsablauf

Die Untersuchung startet am 27. Oktober 2010 und endet am 24. Jänner 2011.

79 potentielle Tester werden für die Untersuchung per Email, sowohl über ein offizielles Einladungsemail der Fa. Onlineumfragen, als auch ein privates Email über den Email-Account des Untersuchungsleiters im Zeitraum zwischen dem 27. Oktober 2010 und dem 29. Dezember 2010 eingeladen. Am 30. November 2011 wird ein zweites Einladungsemail zur Erinnerung versendet. Mit der Einladung erhalten die Tester einen persönlichen Zugangscode für den Einstieg in den Onlinefragebogen.

Bis zum 24. Jänner 2011 öffnen 98,73% (n=78) der Teilnehmer den Online-Fragebogen mit ihrem Code. Davon beantworten 44,30% (n=35) mindestens eine Frage. 32,91% (n=26) dieser Tester erfüllen das Kriterium, Osteopath zu sein. Diese Ergebnisse werden statistisch verarbeitet. Ein Tester bricht nach der dritten Frage die Untersuchung ab. Damit werden 31,65% (n=25) der Fragebogen in die Statistik dieser Masterthese eingerechnet.

Innerhalb des Fragebogens ist es ab der Frage v4 für den Tester nicht mehr möglich, ein Video nicht zu bewerten bzw. keine Antwort abzugeben. Dadurch ist es notwendig, eine Antwortkategorie „Weiß nicht“ einzuführen, um den Fortschritt im Fragebogen auf jeden Fall zu gewährleisten. Auf diese Art und Weise werden die Fragen im Fragebogen von den 25 Testern zu 100% beantwortet. Das entspricht 700 Antworten, die in die Statistik miteinbezogen werden. Antworten der Kategorie „Weiß nicht“ werden statistisch jedoch als fehlend interpretiert. Von insgesamt 700 abgegebenen Antworten entfallen 28 auf diese Antwortmöglichkeit. Das entspricht 4% der abgegebenen Antworten. Das beeinflusst die Auswahl der statistischen Auswertungsmethode (vgl. Kap. 6.3.1.1.)

### 6.3. Statistik

Jeder praktizierende Arzt (und Osteopath) weiß aus Erfahrung, dass es in seinem Fach keine Entscheidung ohne eine gewisse Unsicherheit gibt. Daran kann auch die medizinische Statistik nichts ändern (Weiß 2008).

Statistik kann jedoch helfen, reliable klinische Tests zu bestimmen, die im Sinne der *evidence based practice* einen Grundpfeiler für eine wissenschaftlich fundierte und nachvollziehbare Diagnose darstellen (Sackett et al. 1996).

„Die Statistik als die Wissenschaft des Zufalls stellt [...] Methoden zur Verfügung, die es ermöglichen, trotz der Unberechenbarkeit der Einzelfälle allgemein gültige Aussagen herzuleiten.“ (Weiß 2008: 4)

### 6.3.1. Statistische Auswertung der Daten

Die Aussagekraft von Reliabilitätsstudien hängt entscheidend von den angewandten statistischen Methoden ab. Bei nominalen bzw. ordinalen Bewertungsschemata können durch die Berechnung des Kappa-Koeffizienten fundierte Aussagen über die Interrater- und Intrarater-Reliabilität getroffen werden. Der Kappa-Koeffizient bezieht im Gegensatz zur Berechnung des Prozentsatzes der Übereinstimmung „Zufallstreffer“ aus dem Testablauf mit ein (Sim und Wright 2005).

Der Kappa-Koeffizient wird von den Faktoren Prävalenz, Bias<sup>1</sup> und abhängigen Ergebnissen beeinflusst. Während die ersten beiden Faktoren an die Validität eines Tests gekoppelt sind (Sim und Wright 2005) und durch die fehlende Validisierung der Diagnose von Faszienstörungen damit nicht beurteilt werden können, sind abhängige Ergebnisse durch das vorliegende Studiendesign kontrollierbar. Der videounterstützte Online-Fragebogen bietet die Möglichkeit von unabhängigen Patienten als Untersuchungssubjekt und unabhängigen Testern, da diese zu unterschiedlichen Zeitpunkten und voneinander unabhängig den Online-Fragebogen durchführen.

#### 6.3.1.1. Der Kappa-Index (k-Index) nach Siegel und Castellan (1988)

Für die Bewertung der Reliabilität wird der Kappa-Index (k-Index) nach Siegel und Castellan (1988) verwendet. Im Gegensatz zu *Cohen's Kappa*, welches auf den Fall von zwei Untersuchern beschränkt ist, handelt es sich dabei um eine generalisierte Form des Kappa-Index. Das heißt, dass die Übereinstimmung aller Untersucher in einem Index ausgedrückt werden kann. Ein weiterer Vorteil ist die Möglichkeit, auch im Fall von teilweise fehlenden Daten den Index berechnen zu können, was im Fall dieser Studie notwendig ist (vgl. Kap. 6.2.3.)

Die Kappa-Indices von Siegel und Castellan (1988) sind bei der Bewertung von zwei Untersuchern nur geringfügig niedriger, so dass die von Landis und Koch

---

<sup>1</sup> Bias wird definiert als „durch falsche Untersuchungsmethoden (z. B. Suggestivfragen) verursachte Verzerrung des Ergebnisses einer Repräsentativerhebung“ (Duden online 2011)

(1977) vorgeschlagene textliche Interpretation der Indices angewandt werden kann.

$k < 0,20$	poor	schwach
$0,20 < k < 0,40$	fair	leidlich
$0,40 < k < 0,60$	moderate	mittelmäßig
$0,60 < k < 0,80$	substantial	beträchtlich
$0,80 < k < 1,00$	almost perfect	fast ideal

Tab. 7: Interpretation von Reliabilitätsergebnissen nach Landis und Koch (1977)

Laut Fjellner et al. (1999) werden Werte von zumindest 0,4 als Indikator für eine akzeptable interindividuelle Verlässlichkeit angesehen.

Der Wert von 0,4 als Indikator für akzeptable Reliabilität ist jedoch kein ausreichendes Kriterium für die Güte und damit verbunden der Aussagekraft eines Tests in der Praxis (vgl. Kap. 2.1.)

Neben dem generalisierten Kappa-Index werden die Kappa-Indices von sämtlichen Untersucher-Paaren (nach Siegel und Castellan 1988) berechnet, um damit einen Überblick über die Streuung der Kappa-Indices zu bekommen. Diese Daten bilden auch die Grundlage für die weiterführenden Vergleiche mehrerer Untergruppen.

Als Datengrundlage werden die Bewertungen der Untersucher zu den Videos v6 bis v33 herangezogen. Die Bewertungen zu den Videos v4 und v5, die lediglich zu Übungszwecken eingefügt sind, finden keine Berücksichtigung.

### 6.3.1.2. Prüfung der teststatistischen Voraussetzungen

Vor der Entscheidung, welcher Signifikanztest zur Anwendung kommt, werden die multivariate Normalität und Homogenität mit statistischen Tests untersucht.

Als Test auf Normalverteilung wird der Shapiro-Wilk Test verwendet, wobei das Signifikanzniveau mit  $\alpha=0,10$  gewählt wurde (Sachs 2004). Im Fall von p-Werten  $<0,10$  werden verteilungsunabhängige Tests durchgeführt. Bei Werten zwischen  $p=0,10$  und  $p=0,20$  werden sowohl verteilungsunabhängige, als auch verteilungsabhängige Tests durchgeführt. Eine Ausnahme bildet die Varianzanalyse, die gegen Verletzungen der Normalitätsbedingung als robust gilt und die auch in diesen Fällen durchgeführt wird. Für die Untersuchung der multivariaten Homogenität der Varianzen wird der Bartlett Test verwendet.

Im Fall von zwei Ausprägungen der unabhängigen Variable (zB hohe und niedrige Erfahrung) werden je nach Ergebnis der Vortests, für die Daten bei denen multivariate Normalität und Varianzhomogenität vorliegt, Independent Samples t-Tests verwendet, im anderen Fall Wilcoxon-Tests.

Bei unabhängigen Variablen, die mehr als drei Ausprägungen aufweisen (zB Quellberuf) wird eine einfaktorielle Varianzanalyse (ANOVA...Analysis of Variance) durchgeführt. Im Anschluss daran wird, im Fall eines signifikanten Ergebnisses der ANOVA, mit Independent Samples t-Tests, bzw. Wilcoxon-Tests untersucht, zwischen welchen der durch die unabhängige Variable gebildeten Gruppen ein signifikanter Unterschied besteht.

### 6.3.1.3. Zur Betrachtung der Untergruppen bei den Testern

Neben der Gesamtstichprobe werden zusätzlich die in Tabelle 8 ersichtlichen Untergruppen untersucht. Gegliedert wird nach Quellberuf und Erfahrung.

<b>Beruf</b>
A. Osteopathen (n=25)
<b>Quellberuf<sup>1</sup></b>
A1. Ärzte (n=9)
A2. Physiotherapeuten (n=10)
A3. Heilpraktiker (n=4)
<b>Erfahrung</b>
<b>Gliederung nach Median</b>
A4. langjährige Erfahrung (>3 Jahre) (n=9)
A5. kurze Erfahrung (1-3 Jahre) (n=16)
A6. langjährige Erfahrung (>2 Jahre) (n=19)
A7. kurze Erfahrung (1 und 2 Jahre) (n=6)

Tab. 8: Untersuchte Untergruppen der Stichprobe

Anschließend wird mittels Varianzanalyse bzw. Signifikanztests (Independent Samples t-Tests bzw. Wilcoxon-Tests) untersucht, ob hinsichtlich des Erwartungswerts signifikante Unterschiede zwischen verschiedenen Untergruppen (vgl. Tab 9) bestehen.

<sup>1</sup> Durch die Teilnahme von zwei Testern mit keinem der angeführten Quellberufe kommt ein Gesamtanzahl von 23 Testern zustande, an denen ein möglicher Zusammenhang zwischen Quellberuf und Reliabilität untersucht wird (vgl. Kapitel 6.2.2.2.).

Osteopathen		
A1. Ärzte	A2. Physiotherapeuten	ANOVA
A1. Ärzte	A3. Heilpraktiker	
A2. Physiotherapeuten	A3. Heilpraktiker	
A4. langjährige Erfahrung (>3 Jahre)	A5. kurze Erfahrung (1-3 Jahre)	t-Test
A6. langjährige Erfahrung (>2 Jahre)	A7. kurze Erfahrung (1 und 2 J)	t-Test

Tab. 9: Mittels Varianzanalyse (ANOVA) bzw. Signifikanztests (t-Test...Independent Samples t-Test, Wilcoxon...Wilcoxon Test) hinsichtlich des Erwartungswerts des k-Index verglichene Untergruppen.

Zur Überprüfung, ob mündliche Erklärungen des Patienten einen Einfluss auf die Ergebnisse haben, werden den Untersuchern drei Videos jeweils mit und ohne Ton zur Bewertung vorgespielt (vgl. Tab. 10). Die Bewertungen zu diesen Videos bilden die Datengrundlage für die Berechnung der k-Indices für die paarweise Übereinstimmung der Untersucher. Diese werden wiederum mittels Signifikanztests (unabhängige Variable „Ton“) miteinander verglichen, wobei die Kappa-Indices nach Substitution der negativen Werte durch Null verwendet werden, da deren Mittelwert weitgehend dem generalisierten Kappa-Index entspricht.

Mit Ton	Ohne Ton
v8	v7
v22	v16
v33	v30

Tab. 10: Datengrundlage für die Überprüfung, ob mündliche Erklärungen der Patienten einen Einfluss auf die Ergebnisse haben.

#### 6.3.1.4. Zur Betrachtung der Ergebnisse anhand der einzelnen Videos

Um ein genaueres Bild über mögliche Zusammenhänge zwischen der Interrater-Reliabilität bei einzelnen Videos und deren Beschaffenheit und Inhalt zu erhalten, werden die Ergebnisse zu den Videos nach folgenden Kriterien analysiert:

- Liegen alle Kriterien für eine Geste vor?

Hintergrund dieser Frage ist die Erkenntnis der Gestikforschung über den temporalen Ablauf von Gestik (vgl. Kapitel 3.2.). Besonderen Wert wird hier darauf gelegt, ob der dreiteilige temporale Ablauf mit Anfangs- *peak*- und Endsequenz ersichtlich ist. Diese Kriterien könnten einen Einfluss auf die Codierung von Körperbewegungen als FDM-Körpersprache haben und die Bewertung der FDM-Körpersprache erleichtern.



- Sind ein oder mehrer gestische Höhepunkte (*peaks*) ersichtlich?  
Damit soll geklärt werden, ob die Interrater-Reliabilität mit eventuell mehreren am Video ersichtlichen *peaks* in Zusammenhang steht, sodass die Diagnose bei einem single choice design beeinflusst werden könnte (vgl. Kap. 6.1.2.2.).
- Verläuft der Gestenhöhepunkt (*peak*) symmetrisch?  
Dieses Kriterium nach Kendon (1996) könnte Gestik von anderen Körperbewegungen weiter unterscheidbar machen und damit die Interrater-Reliabilität beeinflussen.
- Haben die Gesten einen dynamischen oder statischen *peak*?  
Könnte die Beurteilung von Körpersprache mit dynamischem oder statischem *peak* unterschiedlich übereinstimmend sein?
- Werden die Gesten mit einem oder mehreren Fingern oder mit der ganzen Hand ausgeführt?  
Beeinflusst die Durchführung der Körpersprache mit Finger(n) oder Hand die Interrater-Reliabilität?
- Befinden sich die gezeigten Gesten im Blickfeld des Patienten?  
Patienten kontrollieren ihre eigene Gestik vorzugsweise haptisch und/oder visuell (vgl. Kapitel 3.3.). Gibt es einen Zusammenhang zwischen der Interrater-Reliabilität der Körpersprache und der Durchführung der Körpersprache des Patienten ausserhalb seines Blickfelds?
- Welche Streuung zeigt sich bei den Einzelergebnissen der Videos?  
Für welche Antworten entscheiden sich die Tester abseits der Antwortmöglichkeit mit der höchsten Übereinstimmung?

Dazu werden die Videos randomisiert nach den beschriebenen Kriterien vom Untersuchungsleiter begutachtet und diese Informationen in eine Tabelle (Tab. 30, siehe Seite 83) eingearbeitet. In einem zweiten Schritt werden die Videos anhand der Statistik in Gruppen hoher und niederer Übereinstimmung (vgl. Tab. 11) geordnet und den Ergebnissen der oben beschriebenen Analyse gegenübergestellt.

## 7. Ergebnisse

Die statistischen Ergebnisse der an der Untersuchung teilnehmenden Osteopathen werden im folgenden dargestellt:

### 7.1. Interrater-Reliabilität der Osteopathen

In Tabelle 11 ist die Häufigkeit der Diagnosen durch die 25 Osteopathen für jeden einzelnen Patienten zusammengefasst.

Der Gesamtindex beträgt  $k$  (Siegel und Castellan, 1988) = **0,51 (mittelmäßig)**, der  $k$ -Index weicht signifikant von Null ab ( $Z = 53,42$ ,  $p < 0,0001$ ).

Die Übereinstimmung zwischen jeweils zwei Untersuchern sind anhand der  $k$ -Indizes in Tabelle 12 dargestellt.

Fasziendistortion	1 - Triggerband TB	2 - Hernierter Triggerpunkt HTP	3 - Continuumdistorsion CD	4 - Faltdistorsion FD	5 - Zylinderdistorsion CyD	6 - Tektonische Fixation TF	7 - keine Fasziendistorsion		gültige Antworten (n)	Weiß nicht - Fehlende Antworten
V6	1	2	<b>11</b>	<b>11</b>	0	0	0		25	
V7	0	5	<b>15</b>	3	0	0	0		23	2
V8	2	5	<b>13</b>	2	0	0	0		22	3
V9	<b>25</b>	0	0	0	0	0	0		25	
V10	<b>24</b>	0	0	0	0	0	0		24	1
V11	4	0	0	4	<b>17</b>	0	0		25	
V12	3	0	0	<b>18</b>	3	1	0		25	
V13	0	7	0	0	0	<b>13</b>	3		23	2
V14	0	0	0	0	<b>25</b>	0	0		25	
V15	1	0	<b>23</b>	0	1	0	0		25	
V16	5	0	1	<b>16</b>	0	0	0		22	3
V17	1	0	2	<b>21</b>	0	0	0		24	1
V18	<b>22</b>	0	1	0	1	0	0		24	1
V19	5	<b>12</b>	2	3	1	1	0		24	1
V20	<b>25</b>	0	0	0	0	0	0		25	
V21	<b>16</b>	0	0	0	9	0	0		25	
V22	2	0	1	<b>21</b>	1	0	0		25	
V23	<b>17</b>	2	1	3	2	0	0		25	
V24	<b>23</b>	0	0	0	2	0	0		25	
V25	0	<b>24</b>	0	0	0	0	0		24	1
V26	0	0	<b>25</b>	0	0	0	0		25	
V27	0	0	0	<b>13</b>	10	2	0		25	
V28	<b>22</b>	0	0	0	3	0	0		25	
V29	1	0	0	4	<b>19</b>	0	0		24	1
V30	6	7	0	<b>8</b>	3	1	0		25	
V31	0	0	0	<b>10</b>	5	0	1		16	9
V32	1	0	0	0	<b>18</b>	0	3		22	3
V33	8	<b>9</b>	0	6	1	1	0		25	

Tab. 11: Häufigkeit der Diagnosen durch die 25 Osteopathen für jeden einzelnen Patienten. (Fett & Kursiv = Wahrscheinlichste Diagnose, dunkelgrüne Markierung: >90% ( $n > 22$ ) Übereinstimmung, hellgrün...>80% ( $n > 20$ ), hellgelb...>70% ( $n > 17$ ) und orange...>60% ( $n > 15$ )).

U1/U2	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	
1	0,24	0,57	0,36	0,44	0,50	0,33	0,40	0,59	0,38	0,39	0,49	0,47	0,29	0,50	0,45	0,43	0,44	0,58	0,54	0,63	0,62	0,45	0,54	0,47	
2		0,30	0,35	0,13	0,35	0,24	0,34	0,31	0,19	0,30	0,31	0,25	0,30	0,40	0,43	0,34	0,26	0,48	0,35	0,40	0,26	0,31	0,40	0,30	
3			0,46	0,36	0,55	0,34	0,54	0,59	0,43	0,58	0,55	0,44	0,54	0,64	0,54	0,53	0,59	0,50	0,50	0,55	0,49	0,37	0,50	0,58	
4				0,29	0,60	0,34	0,53	0,55	0,42	0,36	0,55	0,32	0,37	0,68	0,63	0,54	0,46	0,46	0,59	0,51	0,28	0,29	0,38	0,37	
5					0,55	0,39	0,37	0,47	0,35	0,41	0,60	0,40	0,49	0,51	0,33	0,40	0,50	0,38	0,42	0,51	0,36	0,55	0,42	0,50	
6						0,44	0,73	0,65	0,39	0,45	0,87	0,64	0,59	0,91	0,73	0,54	0,64	0,73	0,73	0,73	0,55	0,65	0,69	0,50	
7							0,25	0,39	0,14	0,25	0,39	0,25	0,43	0,39	0,25	0,38	0,39	0,43	0,39	0,52	0,34	0,35	0,44	0,47	
8								0,55	0,42	0,44	0,68	0,54	0,45	0,73	0,72	0,48	0,59	0,55	0,63	0,64	0,41	0,46	0,51	0,45	
9									0,48	0,50	0,65	0,50	0,46	0,73	0,68	0,59	0,51	0,47	0,60	0,60	0,64	0,56	0,60	0,36	
10										0,47	0,48	0,38	0,34	0,47	0,55	0,55	0,43	0,39	0,52	0,43	0,43	0,44	0,39	0,34	
11											0,45	0,44	0,63	0,50	0,58	0,67	0,59	0,36	0,45	0,46	0,44	0,59	0,46	0,54	
12												0,73	0,63	0,87	0,68	0,54	0,64	0,64	0,73	0,69	0,46	0,69	0,69	0,50	
13													0,49	0,59	0,54	0,53	0,54	0,59	0,50	0,64	0,44	0,64	0,59	0,39	
14														0,55	0,41	0,63	0,72	0,46	0,45	0,55	0,49	0,59	0,50	0,58	
15															0,82	0,59	0,60	0,69	0,73	0,69	0,46	0,60	0,69	0,46	
16																0,62	0,55	0,64	0,68	0,60	0,50	0,60	0,60	0,41	
17																	0,63	0,54	0,49	0,63	0,53	0,55	0,45	0,58	
18																		0,51	0,59	0,69	0,45	0,60	0,51	0,59	
19																			0,64	0,73	0,55	0,56	0,56	0,54	
20																					0,73	0,59	0,60	0,46	
21																						0,59	0,56	0,65	0,68
22																							0,46	0,50	0,44
23																								0,51	0,42
24																									0,41

Tab. 12: Übereinstimmung zwischen jeweils zwei Untersuchern (Osteopathen) anhand der k-Indices nach Siegel und Castellan (1988). Hellgrüne Zellen:  $k > 0,6$ , hellgelbe Zellen:  $k \leq 0,4$

Auffallend sind die Untersucher 2 und 7, die hauptsächlich  $k \leq 0,4$  und kein einziges Mal  $k > 0,6$  aufweisen.

Dabei handelt es sich um zwei männliche Tester mit Quellberuf Physiotherapeut und drei bzw. vier Jahren an Erfahrung mit dem FDM.

Der Mittelwert aller in Tabelle 12 ersichtlichen Indices beträgt  $k = 0,50 \pm 0,13$  (95%-Konfidenzintervall: 0,49-0,51), ohne die Berücksichtigung der beiden Untersucher 2 und 7 beträgt er  $k = 0,53 \pm 0,11$  (95%-Konfidenzintervall: 0,52-0,54).

### 7.1.1. Interrater-Reliabilität und Quellberuf

Die Ergebnisse der Osteopathen unterscheiden sich in Bezug auf deren Quellberuf wie folgt:

### 7.1.1.1. Interrater-Reliabilität der Ärzte

Der Gesamtindex der Ärzte beträgt  $k$  (Siegel und Castellan, 1988) = **0,58 (mittelmäßig)**, der  $k$ -Index weicht signifikant von Null ab ( $Z = 30,547$ ,  $p < 0,0001$ ). Die Übereinstimmung zwischen jeweils zwei Untersuchern sind anhand der  $k$ -Indices in Tabelle 13 dargestellt.

U1/U2	4	8	9	10	12	15	16	21	25
4		0,53	0,55	0,42	0,55	0,68	0,63	0,51	0,37
8			0,55	0,42	0,68	0,73	0,72	0,64	0,45
9				0,48	0,65	0,73	0,68	0,60	0,36
10					0,48	0,47	0,55	0,43	0,34
12						0,87	0,68	0,69	0,50
15							0,82	0,69	0,46
16								0,60	0,41
21									0,68

Tab. 13: Übereinstimmung zwischen jeweils zwei Untersuchern (Osteopathen Quellberuf: Arzt) anhand der  $k$ -Indices nach Siegel und Castellan (1988). Hellgrüne Zellen:  $k > 0,6$ , hellgelbe Zellen:  $k \leq 0,4$

Der Mittelwert aller in Tabelle 13 ersichtlichen Indices beträgt  $k = 0,57 \pm 0,13$  (95%-Konfidenzintervall: 0,53-0,61).

### 7.1.1.2. Interrater-Reliabilität der Physiotherapeuten

Der Gesamtindex der Physiotherapeuten beträgt  $k$  (Siegel und Castellan, 1988) = **0,43 (mittelmäßig)**, der  $k$ -Index weicht signifikant von Null ab ( $Z = 22,953$ ,  $p < 0,0001$ ). Die Übereinstimmung zwischen jeweils zwei Untersuchern sind anhand der  $k$ -Indices in Tabelle 14 dargestellt.

U1/U2	3	5	7	11	18	19	22	23	24
2	0,30	0,13	0,24	0,30	0,26	0,48	0,26	0,31	0,40
3		0,36	0,34	0,58	0,59	0,50	0,49	0,37	0,50
5			0,39	0,41	0,50	0,38	0,36	0,55	0,42
7				0,25	0,39	0,43	0,34	0,35	0,44
11					0,59	0,36	0,44	0,59	0,46
18						0,51	0,45	0,60	0,51
19							0,55	0,56	0,56
22								0,46	0,50
23									0,51

Tabelle 14: Übereinstimmung zwischen jeweils zwei Untersuchern (Osteopathen, Quellberuf: Physiotherapeut) anhand der  $k$ -Indices nach Siegel und Castellan (1988). Hellgrüne Zellen:  $k > 0,6$ , hellgelbe Zellen:  $k \leq 0,4$

Der Mittelwert aller in Tabelle 14 ersichtlichen Indices beträgt  $k=0,43\pm 0,11$  (95%-Konfidenzintervall: 0,40-0,46), ohne Berücksichtigung der beiden Untersucher zwei und sieben ist er jedoch deutlich besser mit  $k=0,49\pm 0,08$  (95%-Konfidenzintervall: 0,46-0,52).

### 7.1.1.3. Interrater-Reliabilität der Heilpraktiker

Der Gesamtindex der Heilpraktiker beträgt  $k$  (Siegel und Castellan, 1988) = **0,57 (mittelmäßig)**, der  $k$ -Index weicht signifikant von Null ab ( $Z = 12,582$ ,  $p < 0,0001$ ). Die Übereinstimmung zwischen jeweils zwei Untersuchern sind anhand der  $k$ -Indices in Tabelle 15 dargestellt.

U1/U2	6	13	20
1	0,50	0,47	0,54
6		0,64	0,73
13			0,50

Tab. 15: Übereinstimmung zwischen jeweils zwei Untersuchern (Osteopathen, Quellberuf: Heilpraktiker) anhand der  $k$ -Indices nach Siegel und Castellan (1988). Hellgrüne Zellen:  $k > 0,6$ , hellgelbe Zellen:  $k \leq 0,4$

Der Mittelwert aller in 15 ersichtlichen Indices beträgt  $k=0,56\pm 0,10$  (95%-Konfidenzintervall: 0,48-0,64).

### 7.1.1.4. Vergleich der Interrater-Reliabilität bei unterschiedlichen Quellberufen

Vergleicht man den Erwartungswert der  $k$ -Indices der Osteopathen in Abhängigkeit vom Faktor Ursprungsberuf, weisen die Ergebnisse der ANOVA auf signifikante Unterschiede zwischen zumindest zwei Gruppen hin ( $F(2,84) = 15,187$ ,  $p < 0,0001$ ). In Abbildung 17 sind die Mittelwerte und 95%-Konfidenzintervalle für die  $k$ -Indices der einzelnen Gruppen ersichtlich.

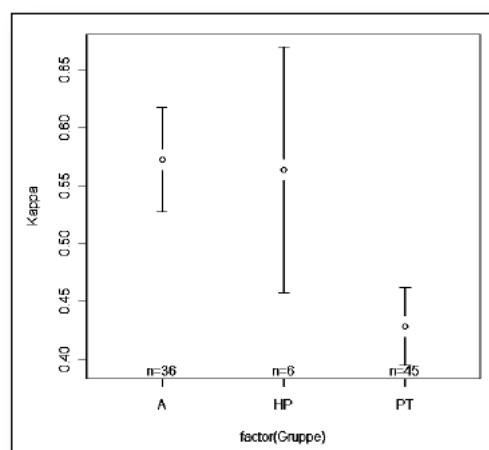


Abb. 17: Mittelwerte und 95%-Konfidenzintervalle für die  $k$ -Indices der durch die unabhängige Variable „Quellberufe“ gebildeten Gruppen (A...Ärzte, HP... Heilpraktiker, PT...Physiotherapeuten) ersichtlich.

Es liegt sowohl multivariate Normalverteilung (vgl. Tab. 16), als auch Varianzhomogenität (Barlett  $K^2= 1,4711$ ,  $df = 2$ ,  $p = 0,48$ ) der Daten vor. Die Ergebnisse der Independent Samples t-Test sind in Tabelle 17 zusammengefasst.

Shapiro-Wilk Test auf Normalverteilung		
gesamt	W = 0,9929	p = 0,92
Ärzte(OA)	W = 0,968	p = 0,37
Physiotherapeuten (OP)	W = 0,967	p = 0,22
Heilpraktiker (OHP)	W = 0,8659	p = 0,21

Tab. 16: Ergebnisse der Shapiro-Wilk Tests auf Normalverteilung der k-Indices.

Gruppe 1	Gruppe 2	t	df	p	Differenz
OA	OP	5,1795	68,2	<b>&lt;0,0001</b>	0,14
OA	OHP	0,1898	8,2	0,85	0,01
OP	OHP	-3,0396	7,8	<b>0,02</b>	0,13

Tab. 17: Ergebnisse der Independent Samples t-Test auf Gruppenunterschiede im Erwartungswert des k-Index (OA... Ärzte, OP... Physiotherapeuten, OHP...Heilpraktiker).

Der mittlere k-Index der Physiotherapeuten der Stichprobe unterscheidet sich signifikant sowohl von jenem der Ärzte ( $t=5,180$ ,  $df=68,2$ ,  $p<0,0001$ ) und jenem der Heilpraktiker ( $t=3,040$ ,  $df=7,8$ ,  $p=0,02$ ). Der mittlere k-Wert der Physiotherapeuten ist um 0,14 bzw. 0,13 niedriger als der der Vergleichsgruppen. Zwischen der Verlässlichkeit der Bewertungen der Ärzte und jener der Heilpraktiker besteht kein signifikanter Unterschied (Differenz:  $k=0,01$ ,  $t=0,1900$ ,  $df=8,2$ ,  $p=0,85$ ).

Nach Ausschluss der Ergebnisse der beiden Physiotherapeuten 2 und 7 unter der Annahme, dass diese Ergebnisse Ausreißer sind, bzw. dass das Fehlen ähnlich niedriger Ergebnisse in den anderen Ursprungsberufsgruppen auf eine nicht repräsentative Stichprobe zurückzuführen ist, sind anhand der ANOVA dennoch signifikante Unterschiede zwischen den drei Gruppen zu erkennen ( $F(2, 67)= 4,656$ ,  $p=0,01$ ).

Es liegt weder multivariate Normalverteilung (vgl. Tab. 18), als auch Varianzhomogenität (Barlett  $K^2= 7,935$ ,  $df = 2$ ,  $p = 0,02$ ) der Daten vor. Die Ergebnisse der Wilcoxon-Tests sind in Tabelle 19 zusammengefasst.

Shapiro-Wilk Test auf Normalverteilung		
gesamt	W = 0,968	<b>p = 0,07</b>
Ärzte(OA)	W = 0,968	p = 0,37
Physiotherapeuten(OP1)	W = 0,929	p = 0,06
Heilpraktiker (OHP)	W = 0,8659	p = 0,21

Tab. 18: Ergebnisse der Shapiro-Wilk Tests auf Normalverteilung der k-Indices.

Gruppe 1	Gruppe 2	Wilcoxon-W	p	Differenz
OA	OP1	687,5	<b>0,01</b>	0,08
OA	OHP	111	0,93	0,01
OP	OHP	54	0,18	0,07

Tab. 19: Ergebnisse der Wilcoxon-Tests auf Gruppenunterschiede im Erwartungswert des k-Index (OA... Ärzte, OP1... Physiotherapeuten ohne case 2 und 7, OHP...Heilpraktiker).

Der mittlere k-Index der Physiotherapeuten der Stichprobe unterscheidet sich signifikant von jenem der Ärzte (Wilcoxon-W=687,5, p=0,01) und deutlich von jenem der Heilpraktiker (Wilcoxon-W=54, p=0,07). Der mittlere k-Wert der Physiotherapeuten ist um 0,08 bzw. 0,07 niedriger als der der Vergleichsgruppen.

### 7.1.2. Interrater-Reliabilität und Erfahrung mit dem FDM

Wie in Tabelle 8 dargestellt, wird der Zusammenhang zwischen der Interrater-Reliabilität und der Erfahrung mit dem FDM in mehreren Schritten analysiert.

#### 7.1.2.1. Interrater-Reliabilität bei mindestens drei Jahren an Erfahrung

Der Gesamtindex der 19 Untersucher (Osteopathen mit mehr als zwei Jahren Erfahrung mit dem FDM) beträgt  $k$  (Siegel und Castellan, 1988) = **0,49 (mittelmäßig)**, der k-Index weicht signifikant von Null ab ( $Z = 42,968$ ,  $p < 0,0001$ ). Die Übereinstimmung zwischen jeweils zwei Untersuchern sind anhand der k-Indices in Tabelle 20 dargestellt.

U1/U2	2	4	5	7	8	9	11	12	13	15	16	17	18	19	20	22	23	25
1	0,24	0,36	0,44	0,33	0,40	0,59	0,39	0,49	0,47	0,50	0,45	0,43	0,44	0,58	0,54	0,62	0,45	0,47
2		0,35	0,13	0,24	0,34	0,31	0,30	0,31	0,25	0,40	0,43	0,34	0,26	0,48	0,35	0,26	0,31	0,30
4			0,29	0,34	0,53	0,55	0,36	0,55	0,32	0,68	0,63	0,54	0,46	0,46	0,59	0,28	0,29	0,37
5				0,39	0,37	0,47	0,41	0,60	0,40	0,51	0,33	0,40	0,50	0,38	0,42	0,36	0,55	0,50
7					0,25	0,39	0,25	0,39	0,25	0,39	0,25	0,38	0,39	0,43	0,39	0,34	0,35	0,47
8						0,55	0,44	0,68	0,54	0,73	0,72	0,48	0,59	0,55	0,63	0,41	0,46	0,45
9							0,50	0,65	0,50	0,73	0,68	0,59	0,51	0,47	0,60	0,64	0,56	0,36
11								0,45	0,44	0,50	0,58	0,67	0,59	0,36	0,45	0,44	0,59	0,54
12									0,73	0,87	0,68	0,54	0,64	0,64	0,73	0,46	0,69	0,50
13										0,59	0,54	0,53	0,54	0,59	0,50	0,44	0,64	0,39
15											0,82	0,59	0,60	0,69	0,73	0,46	0,60	0,46
16												0,62	0,55	0,64	0,68	0,50	0,60	0,41
17													0,63	0,54	0,49	0,53	0,55	0,58
18														0,51	0,59	0,45	0,60	0,59
19															0,64	0,55	0,56	0,54
20																0,59	0,60	0,46
22																	0,46	0,44
23																		0,42

Tab. 20: Übereinstimmung zwischen jeweils zwei Untersuchern (Osteopathen mit mehr als zwei Jahren Erfahrung mit dem FDM) anhand der k-Indices nach Siegel und Castellan (1988). Hellgrüne Zellen:  $k > 0,6$ , hellgelbe Zellen:  $k \leq 0,4$



Der Mittelwert aller in Tabelle 20 ersichtlichen Indices beträgt  $k=0,50\pm0,12$  (95%-Konfidenzintervall: 0,48-0,52), ohne Berücksichtigung der Untersucher 2 und 7 beträgt er  $k=0,53\pm0,11$  (95%-Konfidenzintervall: 0,51-0,55).

### 7.1.2.2. Interrater-Reliabilität bei unter drei Jahren an Erfahrung

Der Gesamtindex der sechs Untersucher (Osteopath mit weniger als drei Jahren Erfahrung mit dem FDM) beträgt  $k$  (Siegel und Castellan, 1988) = **0,52 (mittelmäßig)**, der  $k$ -Index weicht signifikant von Null ab ( $Z = 18,860$ ,  $p < 0,0001$ ). Die Übereinstimmung zwischen jeweils zwei Untersuchern sind anhand der  $k$ -Indices in Tabelle 21 dargestellt.

U1/U2	6	10	14	21	24
3	0,55	0,43	0,54	0,55	0,50
6		0,39	0,59	0,73	0,69
10			0,34	0,43	0,39
14				0,55	0,50
21					0,65

Tab. 21: Übereinstimmung zwischen jeweils zwei Untersuchern (Osteopathen mit weniger als drei Jahren Erfahrung mit dem FDM) anhand der  $k$ -Indices nach Siegel und Castellan (1988). Hellgrüne Zellen:  $k > 0,6$ , hellgelbe Zellen:  $k \leq 0,4$

Der Mittelwert aller in Tabelle 21 ersichtlichen Indices beträgt  $k=0,52\pm0,11$  (95%-Konfidenzintervall: 0,46-0,58).

### 7.1.2.3. Interrater-Reliabilität im Vergleich bei Osteopathen mit weniger als drei und mit mehr als zwei Jahren an Erfahrung mit dem FDM

Vergleicht man den Erwartungswert der  $k$ -Indices der Untersucher mit weniger als drei Jahren Erfahrung mit dem FDM mit jenem der Untersucher, die mehr als zwei Jahre an Erfahrung aufweisen, zeigt sich, dass keine signifikanten Unterschiede in der Verlässlichkeit zwischen diesen beiden Gruppen bestehen. In Abbildung 18 sind die Mittelwerte und 95%-Konfidenzintervalle für die  $k$ -Indices der einzelnen Gruppen ersichtlich.

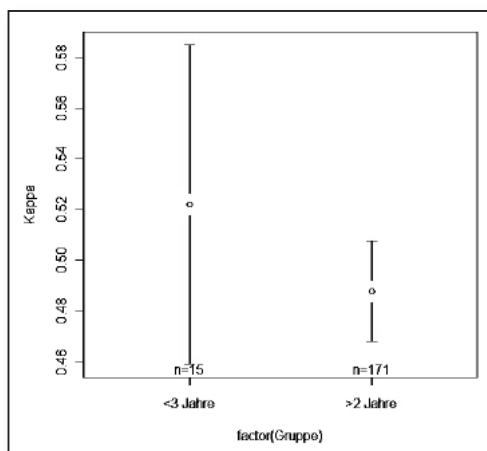


Abb. 18: Mittelwerte und 95%-Konfidenzintervalle für die  $k$ -Indices der anhand der Jahre an Erfahrung mit dem FDM gebildeten Gruppen ersichtlich.

Es liegt sowohl multivariate Normalverteilung (vgl. Tabelle 22), als auch Varianzhomogenität (Barlett  $K^2= 0,485$ ,  $df = 1$ ,  $p = 0,49$ ) der Daten vor. Die Ergebnisse des Independent Samples t-Tests sind in Tabelle 23 zusammengefasst.

Shapiro-Wilk Test auf Normalverteilung		
gesamt	W = 0,9932	p = 0,55
Niedrige Erfahrung (OEN3)	W=0,9639	p=0,76
Hohe Erfahrung (OEH3)	W = 0,9929	p=0,57

Tab. 22: Ergebnisse der Shapiro-Wilk Tests auf Normalverteilung der k-Indices

Gruppe 1	Gruppe 2	t	df	p	Differenz
OEN3	OEH3	-1,0971	17,5	0,28	0,03

Tab. 23: Ergebnisse des Independent Samples t-Tests auf Gruppenunterschiede im Erwartungswert des k-Index (OEN3... Osteopathen mit weniger als drei Jahren Erfahrung mit dem FDM, OEH3... Osteopathen mit mehr als zwei Jahren Erfahrung mit dem FDM).

Es besteht kein signifikanter Unterschied in der mittleren Verlässlichkeit jener Osteopathen mit mindestens drei Jahren Erfahrung mit dem FDM und jenen mit weniger als drei Jahren Erfahrung (Differenz:  $k=0,03$ ,  $t=1,097$ ,  $df=17,5$ ,  $p=0,28$ ).

#### 7.1.2.4. Interrater-Reliabilität bei mehr als drei Jahre an Erfahrung

Der Gesamtindex der neun Untersucher (Osteopathen mit mehr als drei Jahren an Erfahrung mit dem FDM) beträgt  $k$  (Siegel und Castellan, 1988) = **0,50 (mittelmäßig)**, der  $k$ -Index weicht signifikant von Null ab ( $Z = 39,912$ ,  $p < 0,0001$ ). Die Übereinstimmung zwischen jeweils zwei Untersuchern sind anhand der  $k$ -Indices in Tabelle 24 dargestellt.

U1/U2	9	11	12	13	16	18	20	25
2	0,31	0,30	0,31	0,25	0,43	0,26	0,35	0,30
9		0,50	0,65	0,50	0,68	0,51	0,60	0,36
11			0,45	0,44	0,58	0,59	0,45	0,54
12				0,73	0,68	0,64	0,73	0,50
13					0,54	0,54	0,50	0,39
16						0,55	0,68	0,41
18							0,59	0,59
20								0,46

Tab. 24: Übereinstimmung zwischen jeweils zwei Untersuchern (Osteopathen mit mehr als drei Jahren Erfahrung mit dem FDM) anhand der  $k$ -Indices nach Siegel und Castellan (1988). Hellgrüne Zellen:  $k > 0,6$ , hellgelbe Zellen:  $k \leq 0,4$

Der Mittelwert aller in Tabelle 24 ersichtlichen Indices beträgt  $k=0,50 \pm 0,14$  (95%-Konfidenzintervall: 0,45-0,55), ohne Berücksichtigung des Untersuchers 2 beträgt er  $k=0,55 \pm 0,10$  (95%-Konfidenzintervall: 0,51-0,59).

### 7.1.2.5. Interrater-Reliabilität bei drei und weniger Jahren an Erfahrung

Der Gesamtindex der 16 Untersucher (Osteopathen mit weniger als vier Jahren Erfahrung mit dem FDM) beträgt  $k$  (Siegel und Castellan, 1988) = **0,50 (mittelmäßig)**, der  $k$ -Index weicht signifikant von Null ab ( $Z = 40,312$ ,  $p < 0,0001$ ). Die Übereinstimmung zwischen jeweils zwei Untersuchern sind anhand der  $k$ -Indices in Tabelle 25 dargestellt.

U1/U2	3	4	5	6	7	8	10	14	15	17	19	21	22	23	24
1	0,57	0,36	0,44	0,50	0,33	0,40	0,38	0,29	0,50	0,43	0,58	0,63	0,62	0,45	0,54
3		0,46	0,36	0,55	0,34	0,54	0,43	0,54	0,64	0,53	0,50	0,55	0,49	0,37	0,50
4			0,29	0,60	0,34	0,53	0,42	0,37	0,68	0,54	0,46	0,51	0,28	0,29	0,38
5				0,55	0,39	0,37	0,35	0,49	0,51	0,40	0,38	0,51	0,36	0,55	0,42
6					0,44	0,73	0,39	0,59	0,91	0,54	0,73	0,73	0,55	0,65	0,69
7						0,25	0,14	0,43	0,39	0,38	0,43	0,52	0,34	0,35	0,44
8							0,42	0,45	0,73	0,48	0,55	0,64	0,41	0,46	0,51
10								0,34	0,47	0,55	0,39	0,43	0,43	0,44	0,39
14									0,55	0,63	0,46	0,55	0,49	0,59	0,50
15										0,59	0,69	0,69	0,46	0,60	0,69
17											0,54	0,63	0,53	0,55	0,45
19												0,73	0,55	0,56	0,56
21													0,59	0,56	0,65
22														0,46	0,50
23															0,51

Tab. 25: Übereinstimmung zwischen jeweils zwei Untersuchern (Osteopathen mit weniger als vier Jahren Erfahrung mit dem FDM) anhand der  $k$ -Indices nach Siegel und Castellan (1988). Hellgrüne Zellen:  $k > 0,6$ , hellgelbe Zellen:  $k \leq 0,4$

Der Mittelwert aller in Tabelle 25 ersichtlichen Indices beträgt  $k = 0,49 \pm 0,12$  (95%-Konfidenzintervall: 0,47-0,51), ohne Berücksichtigung des Untersuchers 7 beträgt er  $k = 0,51 \pm 0,11$  (95%-Konfidenzintervall: 0,49-0,53).

### 7.1.2.6. Interrater-Reliabilität im Vergleich bei Osteopathen mit weniger als vier und mit mehr als drei Jahren an Erfahrung mit dem FDM

Vergleicht man den Erwartungswert der  $k$ -Indices der Untersucher mit weniger als vier Jahren Erfahrung mit dem FDM mit jenem der Untersucher, die mehr als drei Jahre an Erfahrung aufweisen, zeigt sich, dass keine signifikanten Unterschiede in der Verlässlichkeit zwischen diesen beiden Gruppen bestehen.

In Abbildung 19 sind die Mittelwerte und 95%-Konfidenzintervalle für die  $k$ -Indices der einzelnen Gruppen ersichtlich.

Es liegt sowohl multivariate Normalverteilung (vgl. Tab. 26), als auch Varianzhomogenität (Barlett  $K^2=0,6605$ ,  $df = 1$ ,  $p = 0,42$ ) der Daten vor. Die Ergebnisse des Independent Samples t-Tests sind in Tabelle 27 zusammengefasst.

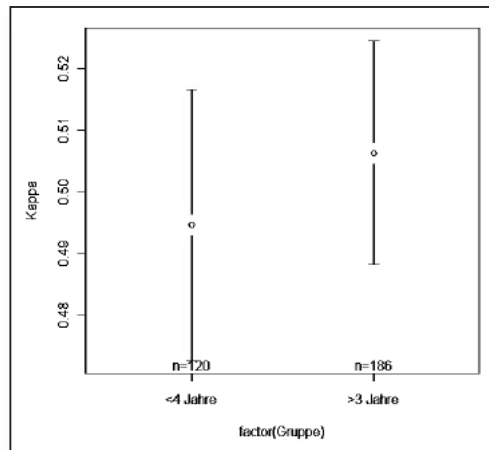


Abb. 19: Mittelwerte und 95%-Konfidenzintervalle für die k-Indices der anhand der Jahre an Erfahrung mit dem FDM gebildeten Gruppen ersichtlich.

Shapiro-Wilk Test auf Normalverteilung		
gesamt	W = 0,9909	p = 0,42
Niedrige Erfahrung (OENM)	W = 0,9866	p = 0,28
Hohe Erfahrung (OEHM)	W = 0,9666	p = 0,34

Tab. 26: Ergebnisse der Shapiro-Wilk Tests auf Normalverteilung der k-Indices

Gruppe 1	Gruppe 2	t	df	p	Differenz
OENM	OEHM	0,0905	53,0	0,93	0,01

Tab. 27: Ergebnisse des Independent Samples t-Tests auf Gruppenunterschiede im Erwartungswert des k-Index (OEN3... Osteopathen mit weniger als vier Jahren Erfahrung mit dem FDM, OEHM3... Osteopathen mit mehr als drei Jahren Erfahrung mit dem FDM).

Es besteht kein signifikanter Unterschied in der mittleren Verlässlichkeit jener Osteopathen mit mehr als drei Jahren Erfahrung mit dem FDM und jenen mit weniger als vier Jahren Erfahrung (Differenz:  $k=0,01$ ,  $t=0,091$ ,  $df=53,0$ ,  $p=0,93$ ).

### 7.1.3. Einfluss der verbalen Aussagen der Patienten auf die Interrater-Reliabilität

Drei Paarungen von Testvideos wurden für diese Untersuchung herangezogen (vgl. Kap. 6.3.1.3.).

#### 7.1.3.1. Interrater-Reliabilität bei Videomaterial ohne Ton

Der Gesamtindex der 25 Untersucher (Osteopathen) bei den drei Videos ohne Ton ( $v7$ ,  $v16$ ,  $v30$ ) beträgt  $k$  (Siegel und Castellano, 1988) = **0,16 (schwach)**, der k-Index weicht dennoch signifikant von Null ab ( $Z = 4,4589$ ,  $p < 0,0001$ ).

Die Übereinstimmung zwischen jeweils zwei Untersuchern sind anhand der k-Indices in Tabelle 28 dargestellt.

U1\U2	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	0,54	0,08	-0,50	0,14	0,08	0,08	0,14	0,54	-0,20	-0,38	0,08	0,08	-0,38	0,08	0,08	-0,38	-0,38	0,54	0,08	0,54	0,54	0,08	0,14	0,45
2		0,14	-0,50	-0,20	0,08	-0,29	0,14	0,14	-0,20	-0,29	0,08	0,08	-0,29	0,08	0,08	-0,38	-0,29	0,54	0,08	0,14	0,14	0,08	0,14	0,08
3			-0,50	-0,29	-0,29	-0,29	0,14	-0,29	0,14	0,08	-0,29	-0,29	0,08	-0,29	-0,29	-0,38	0,08	0,14	-0,29	-0,29	-0,29	-0,29	-0,20	0,45
4				-0,33	-0,20	-0,33	-0,33	-0,33	-0,33	-0,33	-0,20	-0,50	-0,33	-0,20	-0,20	-0,20	-0,33	-0,33	-0,20	-0,50	-0,33	-0,33	-0,33	-0,64
5					0,00	0,08	0,14	0,54	0,54	0,08	0,00	-0,29	0,08	0,00	0,00	0,00	0,08	-0,29	0,00	0,08	0,54	0,14	-0,38	0,00
6						-0,64	0,45	0,45	0,00	0,00	1,00	0,08	0,00	1,00	1,00	-0,33	0,00	0,45	1,00	0,08	0,45	0,45	0,45	-0,38
7							-0,38	0,00	-0,29	-0,38	-0,64	-0,29	-0,38	-0,64	-0,64	0,00	-0,38	-0,38	-0,64	0,08	0,00	-0,38	-0,38	0,00
8								0,54	0,54	0,14	0,45	0,14	0,14	0,45	0,45	0,00	0,14	0,08	0,45	0,14	0,54	0,54	0,08	-0,29
9									0,14	0,08	0,45	0,14	0,08	0,45	0,45	0,00	0,08	0,08	0,45	0,54	1,00	0,54	0,08	0,00
10										0,14	0,00	-0,29	0,14	0,00	0,00	0,00	0,14	-0,29	0,00	-0,29	0,14	0,14	-0,38	-0,29
11											0,00	0,14	1,00	0,00	0,00	0,45	1,00	-0,38	0,00	-0,29	0,08	0,54	-0,29	-0,09
12												0,08	0,00	1,00	1,00	-0,33	0,00	0,45	1,00	0,08	0,45	0,45	0,45	-0,38
13													0,14	0,08	0,08	0,08	0,14	0,08	0,08	0,54	0,14	0,54	0,54	-0,38
14														0,00	0,00	0,45	1,00	-0,38	0,00	-0,29	0,08	0,54	-0,29	-0,09
15															1,00	-0,33	0,00	0,45	1,00	0,08	0,45	0,45	0,45	-0,38
16																-0,33	0,00	0,45	1,00	0,08	0,45	0,45	0,45	-0,38
17																	0,45	-0,64	-0,33	-0,29	0,00	0,45	-0,50	-0,50
18																		-0,38	0,00	-0,29	0,08	0,54	-0,29	-0,09
19																			0,45	0,14	0,08	0,00	0,54	0,08
20																				0,08	0,45	0,45	0,45	-0,38
21																					0,54	0,14	0,54	0,00
22																						0,54	0,08	0,00
23																							0,08	-0,38
24																								-0,29

Tab. 28: Übereinstimmung bei den drei Videos ohne Ton zwischen jeweils zwei Untersuchern (Osteopathen) anhand der k-Indices nach Siegel und Castellan (1988). Hellgrüne Zellen:  $k > 0,6$ , hellgelbe Zellen:  $k \leq 0,4$ .

Der Mittelwert aller in Tabelle 28 ersichtlichen Indices beträgt  $0,06 \pm 0,37$  (95%-Konfidenzintervall: 0,02-0,10), ersetzt man die negativen Werte durch Null, beträgt der Mittelwert  $k = 0,17 \pm 0,26$  (95%-Konfidenzintervall: 0,14-0,20).

### 7.1.3.2. Interrater-Reliabilität bei Videomaterial mit Ton

Der Gesamtindex der 25 Untersucher (Osteopathen) bei den drei Videos mit Ton ( $v_8, v_{22}, v_{33}$ ) beträgt  $k$  (Siegel und Castellan, 1988) = **0,24 (leidlich)**, der  $k$ -Index weicht signifikant von Null ab ( $Z = 5,8944, p < 0,0001$ ).

Die Übereinstimmung zwischen jeweils zwei Untersuchern sind anhand der k-Indices in Tabelle 29 dargestellt.

U1\U2	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	-0,29	0,45	0,00	0,08	0,00	0,08	0,00	0,54	0,14	0,00	0,08	0,08	0,00	0,00	0,00	-0,09	1,00	0,00	0,54	0,54	0,45	0,08	0,08	-0,38
2		-0,29	0,08	-0,29	0,08	-0,29	0,08	0,14	-0,20	-0,29	0,08	0,08	-0,29	0,08	0,08	-0,38	-0,29	0,08	0,14	0,14	-0,29	0,08	0,14	-0,29
3			-0,09	0,08	-0,09	-0,09	-0,09	0,45	0,08	0,45	0,08	0,08	0,45	-0,09	-0,09	-0,09	0,45	-0,09	0,45	0,45	1,00	0,08	0,00	0,00
4				0,00	1,00	-0,64	1,00	0,45	0,00	0,00	0,45	0,45	0,00	1,00	1,00	-0,33	0,00	1,00	0,45	0,45	-0,09	0,45	-0,29	-0,29
5					0,00	-0,29	0,00	0,14	0,54	0,54	0,54	0,54	0,54	0,00	0,00	0,45	0,08	0,00	0,14	0,14	0,08	0,54	0,08	0,14
6						-0,64	1,00	0,45	0,00	0,00	0,45	0,45	0,00	1,00	1,00	-0,33	0,00	1,00	0,45	0,45	-0,09	0,45	-0,29	-0,29
7							-0,64	0,00	-0,29	-0,38	-0,38	-0,38	-0,38	-0,64	-0,64	0,00	0,08	-0,64	0,00	0,00	-0,09	-0,38	0,14	-0,29
8								0,45	0,00	0,00	0,45	0,45	0,00	1,00	1,00	-0,33	0,00	1,00	0,45	0,45	-0,09	0,45	-0,29	-0,29
9									0,14	0,08	0,54	0,54	0,08	0,45	0,45	0,00	0,54	0,45	1,00	1,00	0,45	0,54	0,14	-0,29
10										0,14	0,14	0,14	0,14	0,00	0,00	0,00	0,14	0,00	0,14	0,14	0,08	0,14	0,14	-0,29
11											0,54	0,54	1,00	0,00	0,00	0,45	0,00	0,00	0,08	0,08	0,45	0,54	-0,38	0,54
12												1,00	0,54	0,45	0,45	0,45	0,08	0,45	0,54	0,54	0,08	1,00	-0,29	0,14
13													0,54	0,45	0,45	0,45	0,08	0,45	0,54	0,54	0,08	1,00	-0,29	0,14
14														0,00	0,00	0,45	0,00	0,00	0,08	0,08	0,45	0,54	-0,38	0,54
15															1,00	-0,33	0,00	1,00	0,45	0,45	-0,09	0,45	-0,29	-0,29
16																-0,33	0,00	1,00	0,45	0,45	-0,09	0,45	-0,29	-0,29
17																	-0,09	-0,33	0,00	0,00	-0,09	0,45	-0,38	0,08
18																		0,00	0,54	0,54	0,45	0,08	0,08	-0,38
19																			0,45	0,45	-0,09	0,45	-0,29	-0,29
20																				1,00	0,45	0,54	0,14	-0,29
21																					0,45	0,54	0,14	-0,29
22																						0,08	0,00	0,00
23																							-0,29	0,14
24																								-0,38

Tab. 29: Übereinstimmung bei den drei Videos mit Ton zwischen jeweils zwei Untersuchern (Osteopathen) anhand der k-Indices nach Siegel und Castellán (1988). Hellgrüne Zellen:  $k > 0,6$ , hellgelbe Zellen:  $k \leq 0,4$ ).

Der Mittelwert aller in Tabelle 29 ersichtlichen Indices beträgt  $0,17 \pm 0,38$  (95%-Konfidenzintervall: 0,13-0,21), ersetzt man die negativen Werte durch Null, beträgt der Mittelwert  $k = 0,24 \pm 0,30$  (95%-Konfidenzintervall: 0,20-0,27).

### 7.1.3.3. Interrater-Reliabilität im Vergleich bei Videomaterial mit und ohne Ton

Für den Vergleich der Verlässlichkeit der Untersuchungen mit Ton mit jenen ohne Ton wurden die Kappa-Indices nach Substitution der negativen Werte durch Null herangezogen, da deren Mittelwert weitgehend dem generalisierten Kappa-Index entspricht.

Der Vergleich der Mittelwert der k-Indices, die aus den Bewertungen der Videos mit Ton errechnet wurden, mit jenen, die aus den Bewertungen der Videos ohne Ton errechnet wurden, zeigt einen signifikanten Unterschied (Wilcoxon  $W = 40598,5$ ,  $p = 0,03$ ).

Mit Ton beträgt der Mittelwert  $k=0,24\pm 0,30$  (95%-Konfidenzintervall: 0,20-0,27), ohne Ton  $k=0,17\pm 0,26$  (95%-Konfidenzintervall: 0,14-0,20).

In Abbildung 20 sind die Mittelwerte und 95%-Konfidenzintervalle für die k-Indices der einzelnen Gruppen ersichtlich.

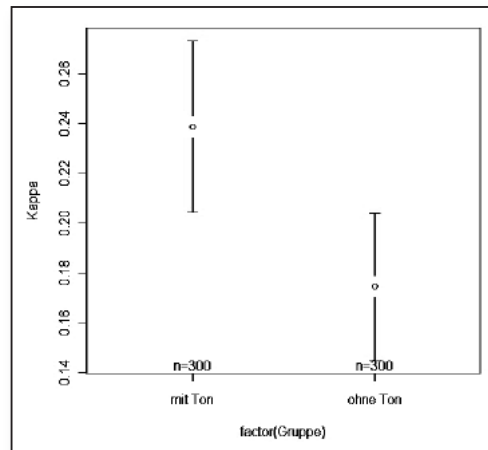


Abb. 20: Mittelwerte und 95%-Konfidenzintervalle für die k-Indices die anhand der Beurteilung der drei Videos der selben drei Patienten mit und ohne Ton errechnet wurden.

#### 7.1.4. Qualitative Analyse der Ergebnisse anhand der einzelnen Videos

Basierend auf den Fragestellung aus Kapitel 6.3.1.4. werden Ergebnisse zu den einzelnen Videos anhand der in Tabelle 30 (siehe Seite 83) ersichtlichen Kriterien weiter untersucht, um eventuelle Zusammenhänge zwischen der Reliabilität und den Kriterien der Gestikforschung zu erlangen. Die Ergebnisse der Tabelle 30 werden im Diskussionsteil aufgegriffen und analysiert.

## 8. Diskussion

Aus Gründen der Übersicht wird dieses Kapitel in eine Zusammenfassung der Ergebnisse, Diskussion der Ergebnisse und Methodenreflexion unterteilt.

### 8.1. Zusammenfassung der Ergebnisse

Diese Masterthese untersucht die Interrater-Reliabilität der visuellen FDM-Körpersprache-Diagnostik mit Hilfe eines videounterstützten Online-Fragebogens.

Weiters werden folgende Zusammenhänge erforscht:

- Interrater-Reliabilität im Zusammenhang mit dem Quellberuf der Tester
- Interrater-Reliabilität im Zusammenhang mit der Erfahrung der Tester mit der FDM-Diagnostik
- Einfluss von verbaler Kommunikation auf die Interrater-Reliabilität der nonverbalen FDM-Körpersprache



(1) Video + max.Ü	(2) A+E	(3) PEAK /n	(4) SYM	(5) S/D	(6) F/H	(7) Im Blickfeld	(8) Streuung
V9: 25 TB			-	D	F		-
V10: 24 TB				D	F		1WN
V14: 25 CyD				D	H		-
V15: 23 CD	E			S	F		1TB, 1CyD
V20: 25 TB				D	F	N	-
V24: 23 TB				D	F	N	2CyD
V25: 24 HTP	E			D	F		1WN
V26: 25 CD				S	F		-
V17: 21 FD	E	2		D	F		1TB, 2CD, 1WN
V18: 22 TB	E			D	F		1CD, 1CyD, 1WN
V22: 21 FD		2		D	F		2TB, 1CD
V28: 22 TB				D	F	N	3CyD
V12: 18 FD				S	H	N	3TB, 3CyD, 1TF
V29: 19 CyD				D	H		1TB, 4FD, 1WN
V32: 18 CyD	E			D	H		1TB, 3KFD
V7: 15 CD	E			D	F		5HTP, 3FD, 2WN
V11: 17 CyD	A+E			D	F	N	4TB, 4FD
V16: 16 FD		2		D	F		5TB, 1CD, 3WN
V21: 16 TB				D	F		9CyD
V23: 17 TB	E		-	D	F	N	2HTP, 1CD, 3FD
V6: 11 CD, 11 FD		2	-	D	F		1TB, 2HTP, 11CD, 1FD
V8: 13 CD	E			D	F		2TB, 5HTP, 2FD, 3WN
V13: 13 TF	E			D	H		7HTP, 3KFD, 2WN
V19: 12 HTP	E	3		D	F	N	5TB, 2CD, 3FD, 1CyD, 1TF, 1WN
V27: 13 FD				S	F		10CyD, 2TF
V30: 8 FD	E			D	F	N	6TB, 7HTP, 3CyD, 1TF
V31: 10 FD				D	-		5CyD, 1KFD, 9WN
V33: 9 HTP	E			D	F	N	8TB, 6FD, 1CyD, 1TF

Tab. 30: Untersuchung der in der Statistik eingearbeiteten Videos nach den Kriterien maximale Übereinstimmung (1), temporaler Ablauf (2, 3, 4), Durchführung der Geste (5, 6, 7) und Streuung der Ergebnisse(8). Die Anordnung der Videos in der Tabelle orientiert sich an der Tabelle 11. Die Videos mit hoher Übereinstimmung (>90%) sind in der Tabelle oben angeordnet, darunter die jeweiligen Gruppen mit weniger prozentueller Übereinstimmung.

**(1) Video + maxÜ:** Größte Übereinstimmung bei der angeführten Diagnose: Triggerband ... TB, Hernierter Triggerpunkt ... HTP, Continuumdistorsion ... CD, Faltdistorsion ... FD, Zylinderdistorsion ... CyD, Tektonische Fixierung ... TF, Keine Fasziendistorsion ... KFD, Weiß nicht ... WN

**(2) A + E: Anfangs- und Endsequenz der Geste:** Fehlen einer Anfangssequenz ... A, Fehlen einer Endsequenz ... E, Fehlen beider Teile ... A + E

**(3) PEAK/n:** Anzahl der gestischen peaks: Anzahl der peaks abweichend von 1

**(4) SYM:** Symmetrie der Geste: kein symmetrischer Gestenablauf ... (-)

**(5) S/D:** Statisch/dynamische Geste: dynamischer peak ... D, statischer peak ... S

**(6) F/H:** Fingergeste/Handgeste: Fingergeste ... F, Handgeste ... H, weder Finger noch Hand am Video ersichtlich ... (-)

**(7) Im Blickfeld:** Gesten im Blickfeld der gefilmten Person: Geste nicht im Blickfeld ... N

**(8) Streuung:** Antworten abseits der größten Übereinstimmung: Triggerband ... TB, Hernierter Triggerpunkt ... HTP, Continuumdistorsion ... CD, Faltdistorsion ... FD, Zylinderdistorsion ... CyD, Tektonische Fixierung ... TF, Keine Fasziendistorsion ... KFD, Weiß nicht ... WN

Die Ergebnisse können in Bezug auf diese Studie folgendermaßen zusammengefasst werden:

Der Gesamtindex beträgt  $k = 0,51$  (nach Siegel und Castellan, 1988), der  $k$ -Index weicht signifikant von Null ab ( $Z = 53,42$ ,  $p < 0,0001$ ). Der Mittelwert aller aus paarweisen Vergleichen der Bewertungen der Osteopathen berechneten  $k$ -Indices beträgt  $k = 0,57 \pm 0,08$  (95%-Konfidenzintervall: 0,54-0,60).

**Damit kann man von einer akzeptablen Interrater-Reliabilität (nach Fjellner et al. 1999) der visuellen FDM-Diagnostik in Zusammenhang mit dieser Studie sprechen.**

Der mittlere  $k$ -Index der Physiotherapeuten der Stichprobe unterscheidet sich signifikant sowohl von jenem der Ärzte ( $t = 5,180$ ,  $df = 68,2$ ,  $p < 0,0001$ ) und jenem der Heilpraktiker ( $t = 3,040$ ,  $df = 7,8$ ,  $p = 0,02$ ). Der mittlere  $k$ -Wert der Physiotherapeuten ist um 0,14 bzw. 0,13 niedriger als der der Vergleichsgruppen. Zwischen der Verlässlichkeit der Bewertungen der Ärzte und jener der Heilpraktiker besteht kein signifikanter Unterschied (Differenz:  $k = 0,01$ ,  $t = 0,1900$ ,  $df = 8,2$ ,  $p = 0,85$ ). Ärzte erreichen einen mittleren  $k$ -Index von  $k = 0,57 \pm 0,13$  (95%-Konfidenzintervall: 0,53-0,61), die Heilpraktiker  $k = 0,56 \pm 0,10$  (95%-Konfidenzintervall: 0,48-0,64) und die Physiotherapeuten  $k = 0,43 \pm 0,11$  (95%-Konfidenzintervall: 0,40-0,46).

Auch ohne Berücksichtigung der beiden Untersucher 2 und 7, die deutlich schlechtere Übereinstimmungen mit den anderen Physiotherapeuten aufweisen, liegt die Reliabilität jedoch mit  $k = 0,49 \pm 0,08$  (95%-Konfidenzintervall: 0,46-0,52) signifikant unter jenem der Ärzte (Wilcoxon- $W = 687,5$ ,  $p = 0,01$ ).

**Der Quellberuf der Osteopathen beeinflusst die Interrater-Reliabilität teilweise signifikant.**

Es besteht kein signifikanter Unterschied in der mittleren Verlässlichkeit jener Osteopathen mit mehr als zwei Jahren Erfahrung mit dem FDM und jenen mit weniger als drei Jahren Erfahrung (Differenz:  $k = 0,03$ ,  $t = 1,097$ ,  $df = 17,5$ ,  $p = 0,28$ ). Auch unter Berücksichtigung eines anderen Grenzwerts für hohe und geringe Erfahrung sind keine signifikanten Unterschiede festzustellen. Die Differenz der mittleren Verlässlichkeit jener Osteopathen mit mehr als drei Jahren Erfahrung mit dem FDM und jenen mit weniger als vier Jahren Erfahrung beträgt  $k = 0,01$  ( $t = 0,091$ ,  $df = 53,0$ ,  $p = 0,93$ ).

## **Die Dauer der Erfahrung mit der FDM-Diagnostik beeinflusst die Interrater-Reliabilität nicht.**

Der Vergleich der Mittelwert der k-Indices, die aus den Bewertungen der drei Videos mit Ton errechnet wurden, mit jenen, die aus den Bewertungen der drei Videos ohne Ton errechnet wurden, zeigt einen signifikanten Unterschied (Wilcoxon  $W = 40598,5$ ,  $p = 0,03$ ). Mit Ton beträgt der Mittelwert  $k=0,24\pm 0,30$  (95%-Konfidenzintervall: 0,20-0,27), ohne Ton  $k=0,17\pm 0,26$  (95%-Konfidenzintervall: 0,14-0,20).

## **Verbale Erläuterungen der Patienten erhöhen die Reliabilität signifikant.**

**Damit bestätigen sich die Null-Hypothesen dieser Masterthese, bis auf den Zusammenhang zwischen Erfahrung und Reliabilität, nicht.**

## **8.2. Diskussion der Ergebnisse**

Die visuelle FDM-Diagnostik scheint ein reliabler Test für die Diagnose von Faszienstörungen nach dem FDM zu sein. Sie erreicht auch im Vergleich mit manuellen Testverfahren an der Wirbelsäule akzeptable Reliabilitätswerte. Die Ergebnisse dieser Untersuchung werfen damit ein positives Licht auf diesen diagnostischen Ansatz, der von Osteopathen angewendet wird.

Diese Untersuchung ist die erste, welche die Interrater-Reliabilität der FDM-Körpersprache untersucht und es gibt keinen unabhängigen Referenzstandard für die Diagnose von Faszienstörungen. Diese Umstände sind in die Beurteilung der Ergebnisse miteinzubeziehen.

Die Patientenvideos des Online-Fragebogens werden in unterschiedlicher Weise übereinstimmend bewertet. Nach einer Analyse des in Kapitel 6.3.1.4. beschriebenen Verfahrens zur weiteren qualitativen Untersuchung der Ergebnisse zeigt sich folgendes:

- Anzahl der *peaks*:  
Obwohl die Videos mit der höchsten Übereinstimmung immer einen klar definierten *peak* zeigen, ist im Überblick keine Zusammenhang hinsichtlich geringerer Übereinstimmung und mehreren *peaks* erkennbar.
- Symmetrie des *peaks*:  
Der überwiegende Teil der Gesten zeigt diese Charakteristik, sowohl bei hoher als auch bei niedriger Übereinstimmung.

- Dynamischer oder statischer *peak* :  
Sowohl Videos mit hoher als auch Videos mit niedriger Übereinstimmung zeigen dynamische und statische Gestenhöhepunkte.
- Finger- bzw Handgesten:  
Ein Großteil der Gesten auf den Videos wird mit den Fingern gezeigt, es findet sich kein Unterschied in der Verteilung dieser Gesten im Zusammenhang mit der Übereinstimmung.
- Gesten im Blickfeld des Patienten:  
Auch hier zeigt sich kein Zusammenhang zwischen optischer Kontrolle des Patienten und der Übereinstimmung durch die Tester.

Bei der Durchsicht von Tabelle 30 zeigt sich eine deutlich größere Streuung der Diagnoseergebnisse bei Videos mit geringerer Übereinstimmung. Interessanterweise häufen sich mit der Abnahme der Übereinstimmung auch Testergebnisse, die in keinster Weise mit den spezifischen Kriterien für die Diagnose von einzelnen Faziendistorsionen nach dem FDM in Zusammenhang gebracht werden können.

Dieser Punkt könnte einen Zusammenhang mit folgender Beobachtung haben: Bei den Videos mit hoher Übereinstimmung (>90% bis >80%) zeigt sich fast ausnahmslos eine deutlich ersichtlicher Ablauf der gezeigten Körpersprache im Sinne von Kendon (1996) und McNeill (1992) mit Anfangs- *peak*- und Endsequenz. Selten fehlt die Endsequenz.

Im Gegensatz dazu fehlen bei den Videos mit geringerer Übereinstimmung öfter Anfangs- oder Endsequenz der Körpersprache oder es sind mehrere gestische *peaks* am Video ersichtlich. Beide Abweichungen betreffen nicht immer dieselben Videos. Jedoch sind in 10 von 13 Videos aus der Gruppe mit einer geringeren Übereinstimmung (vgl. Tab. 11) diese Abweichungen ersichtlich. Das könnte für den Tester einerseits dazu führen, dass damit unklar wird, welcher Teil der Körperbewegung für die Testung entscheidend ist. Andererseits könnte eine derartige Verlaufsform von Körpersprache mit einem unklaren Gestenhöhepunkt negativ auf „naives“ Betrachten der Geste auswirken (vgl. Kap. 3.3.) und mehr Spielraum für Interpretationen lassen. Inwieweit diese Beobachtungen mit dem Aufbau des Fragebogens in der vorliegenden Form, dem Videoschnitt bzw. unzureichend formulierten FDM-Kriterien für Körpersprache an sich, zu tun haben, bleibt an dieser Stelle ungeklärt.

Bei den Antwortmöglichkeiten des Fragebogens haben die Tester die Möglichkeit, den Punkt „Keine Faziendistorsion“ anzuklicken, wenn sie keine beurteilbare Fas-

ziendistorsion erkennen. Von dieser Möglichkeit wird nur bei sieben von insgesamt 700 abgegebenen Antworten Gebrauch gemacht (vgl. Tab. 11). Das entspricht 1% der gesamten Antworten. Interessant ist dieses Ergebnis, weil in zwei Videos keine dem Körpersprache-Vokabular des FDM zuordenbare Bewegungen ersichtlich sind. In weiteren 28 Fällen entscheiden sich die Tester für die Antwortmöglichkeit „Weiß nicht“. Das sind 4% der Antworten. In der Arbeitshypothese zu dieser Arbeit wird die Bedingung formuliert, dass die Interrater-Reliabilität mit ausreichenden diagnostischen Kriterien seitens des FDM in Zusammenhang steht. Obwohl die generelle Reliabilität der visuellen Diagnostik akzeptable Ergebnisse liefert, zeigen sich bei einzelnen Videos Unstimmigkeiten seitens der Tester. Das betrifft sowohl die Entscheidung über das prinzipielle Vorhandensein bzw. Nichtvorhandensein einer codierbaren Körpersprache als auch deren Zuordnung zu bestimmten Faszien- und Distorsionen.

In der Frage der Interrater-Reliabilität schneiden die Berufsgruppen Ärzte und Heilpraktiker signifikant besser ab als Physiotherapeuten. Dies steht im Widerspruch zu den Erkenntnissen von Seffinger et al. (2004) und Stockendahl et al. (2006), die keinen Zusammenhang zwischen Beruf der Tester und der Interrater-Reliabilität in ihren Reviews bestätigen können.

Eine Erklärung für diesen statistischen Unterschied könnte eine unterschiedliche Zusammensetzung der einzelnen Testerguppen in Bezug auf Erfahrung oder Geschlecht sein. Diese Kriterien miteinzubeziehen würde die Anwendung einer ANCOVA-Statistik anstatt der hier verwendeten ANOVA-Statistik notwendig machen. Hauptausgangspunkt, warum in dieser Studie nicht die ANCOVA-Statistik berechnet wird, ist die Tatsache, dass der Wertebereich der Erfahrung der Tester mit dem FDM nur eine geringe Varianz aufweist. 16 von 25 Testern haben zwischen zwei und drei Jahre an Erfahrung (vgl. Tab. 5). Zudem hat die Variable Erfahrung nur wenige Abstufungen (insgesamt nur sechs Abstufungen: 2,3,4,5, 7 und 13 Jahren an Erfahrung). Daher ist zu erwarten, dass aus dieser Untersuchung kein linearer Zusammenhang zwischen Erfahrung und Reliabilität herausgearbeitet werden kann (selbst wenn er prinzipiell existieren sollte). Ein Miteinbeziehen der Erfahrung könnte sich demzufolge negativ durch die Extremwerte bei der Erfahrung auf die Aussagekraft der Ergebnisse auswirken. Bei einer homogeneren Verteilung der Erfahrung wäre allerdings ANCOVA von Vorteil.

Die Testerguppe besteht aus zwanzig männlichen und fünf weiblichen Teilnehmern. Der überproportional hohe Anteil an Männern in der Gruppe der Ärzte und der überproportional hohe Anteil an Frauen in der Gruppe der Physiotherapeuten,

könnte die Ergebnisse in diesem Bereich ebenfalls beeinflussen. Wie beschrieben wird keine ANCOVA-Statistik angewendet, in der diese Charakteristik in Zusammenhang mit dem Quellberuf der Tester gestellt werden könnte.

In dieser Studie kann kein Zusammenhang zwischen der Erfahrung der Tester und der Interrater-Reliabilität gefunden werden. Das könnte mit der Auswahl der Tester in Bezug auf ihre Erfahrung zusammenhängen. Eine homogenere Verteilung der Erfahrungsjahre mit dem FDM seitens der Tester könnte die Ergebnisse verändern. Prinzipiell korreliert dieses Ergebnis mit den Reviews von Seffinger et al. (2004), Stockendahl et al. (2006) und Haneline und Young (2009), die keinen generellen Zusammenhang von Reliabilität und Erfahrung feststellen können.

Ob das Geschlecht der Tester in Zusammenhang mit der Erfahrung Auswirkung auf dieses Ergebnis hat, wird nicht durch eigene Berechnungen überprüft. Die Geschlechterverteilung der unterschiedlichen Erfahrungsgruppen sind in Tabelle 5 dargestellt.

Der Einfluss der verbalen Aussage der Patienten kann im Durchschnitt mit einem signifikanten Unterschied in der Interrater-Reliabilität zugunsten der Videos mit Ton beantwortet werden. Prinzipiell scheint die verbale Patientenaussage zu einer höheren mittleren Übereinstimmung in der Diagnose zu verhelfen, wie in Tabelle 28 und 29 ersichtlich ist. Ohne Ton entscheiden sich 14 Testerpaare für dieselbe Diagnose, im Gegensatz zu 28 Paarungen bei Videos mit Ton.

Die Tester werden durch die verbalen Schilderungen signifikant beeinflusst. Bei den drei Patientenvideos, die nicht repräsentativ für die Population sein müssen, kommt es gesamt betrachtet zu einer Vereinheitlichung der Bewertungen bei zusätzlichen verbalen Aussagen. Allerdings kann aus den Ergebnissen geschlossen werden, dass die Verbalisation auch zu stärker divergierenden Untersuchungsergebnissen führen kann, als es ohne verbale Schilderungen der Fall ist. Grundsätzlich können diese Ergebnisse jedoch nur eine Tendenz wiedergeben, da die Anzahl an Videopaarungen mit und ohne Ton auf drei beschränkt ist. Hier wäre eine deutlich größere Stichprobe notwendig, um diesen Ergebnissen mehr Gewicht zu verleihen.

### **8.3. Methodenreflexion**

Die Ergebnisse dieser Studie sind vor dem Hintergrund folgender methodischer Probleme zu sehen:

Eine höhere Anzahl an Videos im Fragebogen könnte die große Vielfalt an FDM-Körpersprache besser repräsentieren und damit die Aussagekraft der Ergebnisse und deren Umlegbarkeit auf die Praxis verstärken. Die gefilmten Patienten sollten sich deutlicher sowohl im Geschlecht, als auch in der Art der Beschwerden, dem Beschwerdestadium (akut, subakut, chronisch), betroffener Körperregion, Alter u.a. unterscheiden. Wie beispielsweise aus dem Vergleich der angeführten Studien im Bereich der Kardiologie (Edmondstone 1995, Albarran et al. 2000) hervorgeht, könnte Gestik im Akutzustand deutlicher präsentiert werden. Das könnte auch die Ergebnisse dieser Studie beeinflussen.

Weiters ist zu überlegen, ob die Produktion der Videos, vor allem die Nachbearbeitung von der Person des Untersuchungsleiters getrennt werden sollte, um eine Blindierung in diesem Zusammenhang zu erreichen.

Die Durchführung dieser Studie als videounterstützter Online-Fragebogen bringt folgende Probleme mit sich:

Die Videos werden teilweise nachbearbeitet und geschnitten. Damit sollen mögliche interne Fehlerquellen reduziert werden. Die Videos werden zudem ein- und ausgeblendet. Diese Eingriffe könnten im Sinne der Gestikforschung einen künstlichen gestischen Höhepunkt des Einzelvideos kreieren, welcher möglicherweise nicht mit dem „natürlichen“ *peak* der Geste konform geht. Das sollte bei zukünftigen Studien durch eine stärkere Berücksichtigung dieser Problematik beim Videoschnitt bzw. länger dauernde Videoausschnitten berücksichtigt werden. Die damit verbundene mögliche Erhöhung interner Fehlerquellen bei *single choice* Studiendesigns, ist jedoch in diesem Zusammenhang zu berücksichtigen.

Durch die Videosituation innerhalb eines Online-Fragebogens ergibt sich, dass eine Provokation bzw. ein Nachfragen bei fehlender Klarheit der Körpersprache nicht möglich ist, was der praktischen Herangehensweise nicht entspricht.

Aus technischen Gründen ist es nicht möglich, ein wiederholtes Ansehen der Videos durch die Tester zu unterbinden. Obwohl diese Reproduzierbarkeit an sich einen Vorteil in der Verwendung von Videoaufnahmen in Reliabilitätstestungen darstellt (Haidet et al. 2009) und Vorteile in der Messgenauigkeit (Sommerfeld 2006) mit sich bringt, ist hier anzumerken, dass diese Situation nicht der Praxis entspricht, in der Körpersprache manchmal nur einmal in derselben Art und Weise gezeigt wird.



Für die Beantwortung der Frage nach dem Einfluss der Verbalisation auf die Interrater-Reliabilität wäre eine größere Anzahl an Videopaarungen mit und ohne Ton notwendig, um den statistischen Aussagen mehr Gewicht zu verleihen.

Am Aufbau des Online-Fragebogens ist weiters zu hinterfragen, inwieweit die Tester zwischen der Antwortmöglichkeit „Keine Faszienstörung“ und „Weiß nicht“ unterscheiden können. Hier wäre eine weitreichendere Aufklärung des theoretischen Hintergrundes dieser Antwortmöglichkeiten und der damit verbundenen Forschungsfrage notwendig, um entsprechende Sensibilität von Seiten der Tester zu generieren.

Die Auswahl der Tester sollte in Bezug auf ihre Erfahrung mit dem FDM ausgeglichener sein, um die Aussagekraft der Ergebnisse in der Frage nach dem Zusammenhang zwischen Erfahrung und Interrater-Reliabilität zu stärken.

Hauptproblem bei dieser Studie ist, wie bei vielen Reliabilitätsstudien im Bereich der Osteopathie und manuellen Medizin, die Diskrepanz zwischen interner und externer Validität. Der reduktionistische Ansatz derartiger wissenschaftlicher Untersuchungen spiegelt nur bedingt den klinischen Alltag wider. In der vorliegenden Studie liegt die Problematik darin, dass Körpersprache selten singulär auftritt und (durch die Beobachtung innerhalb der Osteopath-Patienten-Kommunikation) zumindest im Kontext mit verbalen Aussagen und der Anamnese des Patienten zu beurteilen ist. Dadurch kann für diese Untersuchung auch keine Aussage über die Reliabilität der FDM-Diagnostik an sich getroffen werden, sondern nur über deren visuell sichtbaren Anteil. Wie beschrieben wird dieser Teil mit anderen Informationen aus der Untersuchung vernetzt, um zu einer FDM-Diagnose zu gelangen. Hier könnte ein Studiendesign mit längeren Videos und/oder inkludierter Anamnese bzw. Bewegungs- und Belastungstests praxisorientiertere Ergebnisse liefern.

## **9. Zusammenfassung und Ausblick**

Obwohl in der Osteopathie vorwiegend manuelle Testmethoden bei der Befunderhebung von Patienten verwendet werden, zeigen sich in dieser Studie erste Hinweise für die akzeptable Reliabilität eines visuell orientierten Diagnoseansatzes, der von zahlreichen Osteopathen in der Praxis angewendet wird.

Mit Hilfe einer videounterstützten Onlineumfrage kann in dieser Masterthese eine akzeptable Interrater-Reliabilität der Körpersprache-Diagnose nach dem Faszienstörungsmodell nachgewiesen werden. Da dieser visuelle Diagnoseansatz

derzeit nicht durch einen unabhängigen Referenzstandard überprüft werden kann und damit die Validität der Körpersprache-Diagnose nicht geklärt ist, sind diese Ergebnisse mit Vorbehalt in die Praxis umzulegen. Besonders die Verwendung des Tests als singuläres Diagnosekriterium ist kritisch zu hinterfragen. Im Kontext der gesamten FDM-spezifischen Befundung zur Diagnose von Faszienstörungen, sind die Ergebnisse dieser Studie jedoch bedeutend. Die signifikante Verbesserung der Reliabilität durch verbale Aussagen des Patienten während der nonverbalen Kommunikation unterstützt diese Aussage trotz methodologischer Vorbehalte.

Folgestudien sollten unter Berücksichtigung der in Kapitel 8 angeführten Kritikpunkte weitere Ergebnisse zum Thema liefern. Von Interesse könnten auch Untersuchungen sein, welche das gesamte Prozedere der FDM-Diagnose auf seine Reliabilität hin testen. Damit wären Aussagen über die Reliabilität möglich, die dem Vorgehen in der Praxis entsprechen.

Forschungsbestrebungen sollten auch hinsichtlich eines unabhängigen Referenzstandards für Faszienstörungen unternommen werden. Gemeinsam mit einer weiteren Präzisierung der Diagnosekriterien für die visuelle FDM-Diagnostik könnte die Reliabilität zusätzlich verbessert werden.

## 10. Literaturverzeichnis

Albarran JW, Durham B, Chappel G, Dwight J, Gowers J (2000): Are manual gestures, verbal descriptors and pain radiation as reported by patients reliable indicators of myocardial infarction? Preliminary findings and implications. *Intensive and Critical Care Nursing*, 16(2), 98 – 110

Bannert T (2011): Nonverbale Kommunikation. Definition. <http://www.nonverbale-kommunikation.info/definition.php> (Zugriff: 10. April 2011)

Biberschick M (2011): „Legt Euch ein Schema zurecht!“. die Routineuntersuchung in der Osteopathie. Master Thesis zur Erlangung des akademischen Grades „Master of Science“ in Osteopathie. Krems. Donau Universität Krems

Duden online (2011): <http://www.duden.de/>

European FDM Association (2010): Therapeutenliste. <http://www.fdm-europe.com/patienteninformationen/fdm-therapeuten/index.html> (Zugriff: 10. Oktober 2010)

European FDM Association (2010): FDM-Ausbildung. <http://www.fdm-europe.com/fdm-ausbildung/ausbildung-fdm--curriculum/index.html> (Zugriff: 10. Oktober 2010)

European FDM Association (2011): Geschichte des FDM. <http://www.fdm-europe.com/fdm/geschichte-fdm/index.html> (Zugriff: 10. April 2011)

Edmondstone WM (1995): Cardiac chest pain: Does body language help the diagnosis? *British Medical Journal*, 311 (7021), 1660 – 1661

Ekman P, Friesen WV (1969): The repertoire of nonverbal behavior: categories, origins, usage, and coding. *Semiotica*, 1, 49 – 98

Ekman P (2011): About Ekman. <http://www.paulekman.com/about-ekman/> (Zugriff: 10. April 2011)

Engel R (2009): The treatment of Chronic Low Back Pain following the principles of the Fascial Distortion Modell (FDM). A quasi experimental study. Master Thesis to obtain the degree Master of Science in Osteopathy. Krems. Donau Universität Krems

Findley TW, Schleip R (2007): Introduction. In: Findley TW, Schleip R (Hrsg.): *Fascia Research. Basic science and implications for conventional and complementary health care*. Munich: Elsevier

Findley TW, Schleip R (Hrsg.) (2007): *Fascia Research. Basic science and implications for conventional and complementary health care*. Munich: Elsevier

Fjellner A, Bexander C, Faleji R, Strender LE (1999): Interexaminer reliability in physical examination of the cervical spine. *J. Manipulative Physiol Ther*, 22, 511 – 516

Fritz JM, Wainner RS (2001): Examining Diagnostic Tests: An Evidence-Based Perspective. *Physical Therapy*, Vol. 81, 9, 1546 – 1564

Geiger, A (2007): Fascia – key for stability, sensomotor function and symmetry. The effects of orthopathic treatment according to the fascia distortion model (FDM) on stability, sensomotor function and symmetry in the upright stance. Master Thesis to obtain the degree Master of Science in Osteopathy. Krems. Donau Universität Krems

Goffman E (1963): Behavior in Public Places. Notes on the Social Organization of Gatherings. New York: Free Press

Haidet KK, Tate J, Divirgilio-Thomas D, Kolanowski A, Happ MB (2009): Methods to Improve Reliability of Video Recorded Behavioral Data. Res Nurs Health, 32 (4): 465 - 474

Haneline MT, Young M (2009): A review of intraexaminer and interexaminer reliability of static spinal palpation: a literature synthesis. J Manipulative Physiol Ther. 32, 379 – 386

Harrer, G (2009): Das Faszienmodell. Die Typaldos Methode. Level1. Modul1. Unterrichtsmaterial der Wiener Schule für Osteopathie

Harrigan, JA (2008): Proxemics, kinesics and gaze. In: Harrigan JA, Rosenthal, R, Scherer KR (Hrsg.): The new handbook of Methods in Nonverbal Behavior Research. Oxford: Oxford University Press

Hartman SE, Norton JM (2002): Interexaminer Reliability And Cranial Osteopathy. The Scientific Review of Alternative Medicine, 6 (1), 23 - 34

Hollerwöger D (2006): Methodological quality and outcomes of studies addressing manual cervical spine examinations: a review. Man Ther. 11, 93-98

Huijing PA, Langevin HM (2009): Communication about Fascia: History, pitfalls and recommendations. In: Huijing PA, Hollander P, Findley TW, Schleip R (Hrsg): Fascia Research II. Basic science and implications for conventional and complementary health care. Munich: Elsevier

Huijing PA, Hollander P, Findley TW, Schleip R (Hrsg) (2009): Fascia Research II. Basic science and implications for conventional and complementary health care. Munich: Elsevier

Johnston WL (1982): Interexaminer reliability studies: Spanning a gap in medical research. Louisa Burns Memorial Lecture. Chicago

Kasten, M (2010): Which Way is Up When You are Upside Down. Bangor: Eigenverlag

Kendon A (1996): An Agenda for Gesture Studies. *Semiotic Review of Books*, 7 (3), 8 – 12

Kleibel V, Mayer H (2005): *Literaturrecherche für Gesundheitsberufe*. Wien: Facultas

Knapp ML, Hall JA (2002): *Nonverbal communication in human interaction*. 5th Edition. Fort Worth: Harcourt Brace

Korr, IM (1992): Osteopathic research: The needed paradigm shift. *JAOA*, 91, 2, 156 - 71

Landis, JR, Koch GG (1977): The measurement of observer agreement for categorical data. *Biometrics*, 33, 159 – 174

Lucas N, Bogduk N (2011): Diagnostic reliability in osteopathic medicine. *International Journal of Osteopathic Medicine*, 1, 1 - 5

Liem, T (2005): *Kraniosakrale Osteopathie. Ein praktisches Lehrbuch*. 4., völlig neu bearbeitete und erweiterte Auflage. Stuttgart: Hippokrates

Marcantonio D (2008): *Italiener in Deutschland und Deutsche in Italien: Ihre Gesten im interkulturellen Vergleich*. Dissertation zur Erlangung des akademischen Grades. Berlin. Fakultät I – Geisteswissenschaften der Technischen Universität Berlin

Marcus GM, Cohen J, Varosy PD, Vessey J, Rose E, Massie BM, Chatterjee K, Waters D (2007): The utility of Gestures in Patients with Chest Discomfort. *The American Journal of Medicine*, 120, 83 – 89

McNeill D (1992): *Hand and mind: What gestures reveal about thought*. Chicago: The University of Chicago Press

McNeill D (2005): *Gesture and Thought*. Chicago: The University of Chicago Press

McNeill D (2011): Professor David McNeill. Members and Friends. McNeill Lab. Center for Gesture and Speech Research. <http://mcneilllab.uchicago.edu/members/McNeill.html> (Zugriff: 12. Mai 2011)

Mirkin S (2009): Exerpts from Sue Mirkin's PhD proposal. New Zealand. University of Otago (Email-Kontakt)

Navarro, J (2008): What every body is saying. An Ex-FBI Agent's Guide to Speed-Reading People. New York: HarperCollins

Nöth W (1999): Handbuch der Semiotik. Stuttgart: J.B. Metzler

Onlineumfragen (2011): Über uns. [http://www.onlineumfragen.com/3ind.ex\\_20-10\\_ueberuns.cfm](http://www.onlineumfragen.com/3ind.ex_20-10_ueberuns.cfm) (Zugriff: 3. April 2011)

Österreichisches Wörterbuch – Schulausgabe. (2003): Wien: ÖBV und HPT

Peterson CK, Gatterman B, Carter JC, Humphreys BK, Weibel A (2007): Inter- and intraexaminer reliability in identifying and classifying degenerative marrow (Modic) changes on lumbar spine magnetic resonance scans. J Manipulative Physiol Ther. ,30 (2), 85-90

Rossmly C (2005): Der Effekt des Fasziendistorsionsmodells (FDM) auf die schmerzhaft eingeschränkte Abduktion der Schulter. Wissenschaftliche Arbeit zur Erlangung des „D.O.-DROM“ des Deutschen Registers Osteopathischer Medizin. College für angewandte Osteopathie

Sachs, L (2004): Angewandte Statistik. 11. überarbeitete und ergänzte Auflage. Berlin. Springer

Sackett DL, Rosenberg WMC, Gray Muir JA, Haynes RB, Scott Richardson W (1996): Evidence based medicine: what it is and what it isn't. British Medical Journal, 312, 71

Schiphof D, de Klerk BM, Koes BW, Zeinstra-Bierma S (2008): Good reliability, questionable validity of 25 different classification criteria of knee osteoarthritis: a systematic appraisal. Journal of Clinical Epidemiology, 61(12):1205-1215

Schmid Mast, M (2007): On the importance of nonverbal communication in the physician–patient interaction. Patient Education and Counseling, 67, 315–318

Seffinger MA, Najm WI, Mishra SI, Adams A, Dickerson VM, Murphy LS, Reinsch S (2004): Reliability of spinal palpation for diagnosis of back and neck pain: a systematic review of the literature. Spine, 1, 29, 413 - 25

Siegel S, Castellan NJ Jr (1988): Nonparametric statistics for the behavioural sciences. Boston: McGraw-Hill

Sim J, Wright CC (2005): The Kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Physical therapy*, 85, 3, 257 - 268

Sommerfeld P, Kaider A, Klein P (2004): Inter- und Intratester-Reliabilität der Palpation des „Primären Respiratorischen Mechanismus“ innerhalb des „Cranialen Konzeptes“. *Osteopathische Medizin*, 4, 4 – 10

Sommerfeld P (2006): Methodologie. Kleine und hoffentlich kurzweilige Einführung in den Sinn und Unsinn wissenschaftlichen Arbeitens auf der Grundlage quantitativer Methoden für OsteopathInnen. 3. Überarbeitete Version. Unterrichtsmaterial der Wiener Schule für Osteopathie

Stein C (2008): Untersuchung der Wirksamkeit einer manuellen Behandlungstechnik nach dem Faszien-Distorsions-Modell bei schmerzhaft eingeschränkter Schulterbeweglichkeit. Eine explorativ-prospektive, randomisierte und kontrollierte klinische Studie. Dissertation zur Erlangung des Doktorgrades der Medizin in der Medizinischen Hochschule Hannover

Stochkendahl MJ, Christensen HW, Hartvigsen J, Vach W, Haas M, Hestbaek L, Adams A, Bronfort G (2006): Manual Examination of the Spine: A Systematic Critical Literature Review of Reproducibility. *Journal of Manipulative and Physiological Therapeutics*, 29, 6, 475 – 85

Stukenbrock A (2008): „Wo ist der Hauptschmerz?“ – Zeigen am eigenen Körper in der medizinischen Kommunikation. *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion*, 9, 1 - 33

Typaldos S (1994): Triggerband Technique. *Journal of the American Academy of Osteopathy* , 4 (4), 15 – 33

Typaldos S (1994): Introducing the Fascial Distortion Model. *Journal of the American Academy of Osteopathy*, 4 (2), 14 – 18, 30 – 36

Typaldos S (1995): Coninum Technique. *Journal of the American Academy of Osteopathy*, 5 (2), 15 – 19

Typaldos S (1999): Orthopathische Medizin. Die Verbindung von Orthopädie und Osteopathie durch das Faszien-Distorsionsmodell. Kötzing/Bayer. Wald: Verlag für Ganzheitliche Medizin Wühr



Typaldos S (2002): FDM – Clinical and Theoretical Application of the Fascial Distortion Model within the Practice of Medicine and Surgery. Bangor: Eigenverlag

Weiß, C (2008): Basiswissen Medizinische Statistik. 4. Auflage. Heidelberg: Springer

Whiting P, Rutjes AWS, Reitsma JB, Bossuyt PMM, Kleijnen J (2003): The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Medical Research Methodology. <http://www.biomedcentral.com/content/pdf/1471-2288-3-25.pdf>

Wiener Schule für Osteopathie (2011): Die Ausbildung im Überblick. <http://www.wso.at/index.php/die-ausbildung/die-ausbildung-im-ueberblick> (Zugriff: 23. Mai 2011)

## 11. Abbildungsverzeichnis

- Abb. 1: Körpersprache bei schmerzhafter Schulter (Typaldos 2002)
- Abb. 2: Beispiel für Triggerband-Körpersprache (© EFDMA 2011)
- Abb. 3: Beispiel für Triggerband-Körpersprache (© EFDMA 2011)
- Abb. 4: Beispiel für Triggerband-Körpersprache (© EFDMA 2011)
- Abb. 5: Beispiel für Körpersprache bei Herniertem Triggerpunkt (© EFDMA 2011)
- Abb. 6: Beispiel für Körpersprache bei Herniertem Triggerpunkt (© EFDMA 2011)
- Abb. 7: Beispiel für Körpersprache bei Continuumdistorsion (© EFDMA 2011)
- Abb. 8: Beispiel für Körpersprache bei Continuumdistorsion (© EFDMA 2011)
- Abb. 9: Beispiel für Körpersprache bei Faltdistorsion (© EFDMA 2011)
- Abb. 10: Beispiel für Körpersprache bei Faltdistorsion (© EFDMA 2011)
- Abb. 11: Beispiel für Körpersprache bei Faltdistorsion (© EFDMA 2011)

- Abb. 12: Beispiel für Körpersprache bei Faltdistorsion (intermuskuläres Septum) (© EFDMA 2011)
- Abb. 13: Beispiel für Körpersprache bei Zylinderdistorsion (© EFDMA 2011)
- Abb. 14: Beispiel für Körpersprache bei Zylinderdistorsion (© EFDMA 2011)
- Abb. 15: Beispiel für Körpersprache bei Tektonischer Fixierung (© EFDMA 2011)
- Abb. 16: Der Zusammenhang zwischen FDM-Körpersprache und der spezifischen Behandlung der jeweiligen Faziendistorsion anhand des Triggerbandes (Typaldos 2002)
- Abb. 17: Mittelwerte und 95%-Konfidenzintervalle für die k-Indices der durch die unabhängige Variable „Ursprungsberuf“ gebildeten Gruppen (A...Ärzte, HP... Heilpraktiker, PT...Physiotherapeuten) ersichtlich.
- Abb. 18: Mittelwerte und 95%-Konfidenzintervalle für die k-Indices der anhand der Jahre an Erfahrung mit dem FDM gebildeten Gruppen ersichtlich.
- Abb. 19: Mittelwerte und 95%-Konfidenzintervalle für die k-Indices der anhand der Jahre an Erfahrung mit dem FDM gebildeten Gruppen ersichtlich.
- Abb. 20: Mittelwerte und 95%-Konfidenzintervalle für die k-Indices die anhand der Beurteilung der drei Videos der selben drei Patienten mit und ohne Ton errechnet wurden.

## 12. Tabellenverzeichnis

- Tab. 1: Homepage-Adressen der verwendeten Online-Datenbanken
- Tab. 2: Das Verhältnis von männlichen und weiblichen Testern in der Studie
- Tab. 3: Quellberuf der Tester, welche die Online-Umfrage regulär beenden
- Tab. 4: Quellberuf in Zusammenhang mit dem Geschlecht der Tester

- Tab. 5: Erfahrung mit dem FDM, Anzahl der Tester und deren Geschlecht im Überblick
- Tab. 6: Erfahrung mit dem FDM in Zusammenhang mit dem Quellberuf der Tester
- Tab. 7: Interpretation von Reliabilitätsergebnissen nach Landis und Koch (1977)
- Tab. 8: Untersuchte Untergruppen der Stichprobe
- Tab. 9: Mittels Varianzanalyse (ANOVA) bzw. Signifikanztests (t-Test... Independent Samples t-Test, Wilcoxon...Wilcoxon Test) hinsichtlich des Erwartungswerts des k-Index verglichene Untergruppen.
- Tab. 10: Datengrundlage für die Überprüfung, ob mündliche Erklärungen der Patienten einen Einfluss auf die Ergebnisse haben.
- Tab. 11: Häufigkeit der Diagnosen durch die 25 Osteopathen für jeden einzelnen Patienten.
- Tab. 12: Übereinstimmung zwischen jeweils zwei Untersuchern (Osteopathen) anhand der k-Indices nach Siegel und Castellán (1988).
- Tab. 13: Übereinstimmung zwischen jeweils zwei Untersuchern (Osteopathen Ursprungsberuf: Arzt) anhand der k-Indices nach Siegel und Castellán (1988).
- Tab. 14: Übereinstimmung zwischen jeweils zwei Untersuchern (Osteopathen, Ursprungsberuf: Physiotherapeut) anhand der k-Indices nach Siegel und Castellán (1988).
- Tab. 15: Übereinstimmung zwischen jeweils zwei Untersuchern (Osteopathen, Ursprungsberuf: Heilpraktiker) anhand der k-Indices nach Siegel und Castellán (1988).
- Tab. 16: Ergebnisse der Shapiro-Wilk Tests auf Normalverteilung der k-Indices.

- Tab. 17: Ergebnisse der Independent Samples t-Test auf Gruppenunterschiede im Erwartungswert des k-Index (OA... Ärzte, OP... Physiotherapeuten, OHP...Heilpraktiker).
- Tab. 18: Ergebnisse der Shapiro-Wilk Tests auf Normalverteilung der k-Indices.
- Tab. 19: Ergebnisse der Wilcoxon-Tests auf Gruppenunterschiede im Erwartungswert des k-Index (OA... Ärzte, OP1... Physiotherapeuten ohne case 2 und 7, OHP...Heilpraktiker).
- Tab. 20: Übereinstimmung zwischen jeweils zwei Untersuchern (OsteopathInnen mit mehr als zwei Jahren Erfahrung mit dem FDM) anhand der k-Indices nach Siegel und Castellan (1988).
- Tab. 21: Übereinstimmung zwischen jeweils zwei Untersuchern (Osteopathen mit weniger als drei Jahren Erfahrung mit dem FDM) anhand der k-Indices nach Siegel und Castellan (1988).
- Tab. 22: Ergebnisse der Shapiro-Wilk Tests auf Normalverteilung der k-Indices.
- Tab. 23: Ergebnisse des Independent Samples t-Tests auf Gruppenunterschiede im Erwartungswert des k-Index (OEN3... Osteopathen mit weniger als drei Jahren Erfahrung mit dem FDM, OEH3... Osteopathen mit mehr als zwei Jahren Erfahrung mit dem FDM).
- Tab. 24: Übereinstimmung zwischen jeweils zwei Untersuchern (Osteopathen mit mehr als drei Jahren Erfahrung mit dem FDM) anhand der k-Indices nach Siegel und Castellan (1988).
- Tab. 25: Übereinstimmung zwischen jeweils zwei Untersuchern (Osteopathen mit weniger als vier Jahren Erfahrung mit dem FDM) anhand der k-Indices nach Siegel und Castellan (1988).
- Tab. 26: Ergebnisse der Shapiro-Wilk Tests auf Normalverteilung der k-Indices.

- Tab. 27: Ergebnisse des Independent Samples t-Tests auf Gruppenunterschiede im Erwartungswert des k-Index (OEN3... Osteopathen mit weniger als vier Jahren Erfahrung mit dem FDM, OEH3... Osteopathen mit mehr als drei Jahren Erfahrung mit dem FDM).
- Tab. 28: Übereinstimmung bei den drei Videos ohne Ton zwischen jeweils zwei Untersuchern (Osteopathen) anhand der k-Indices nach Siegel und Castellan (1988).
- Tab. 29: Übereinstimmung bei den drei Videos mit Ton zwischen jeweils zwei Untersuchern (Osteopathen) anhand der k-Indices nach Siegel und Castellan (1988).
- Tab. 30: Untersuchung der in der Statistik eingearbeiteten Videos nach den Kriterien Übereinstimmung (1), temporaler Ablauf (2, 3, 4), Durchführung (5, 6, 7) und Streuung (8).

## 13. Anhang

### Abkürzungsverzeichnis

A	Arzt
ANCOVA	Analysis of covariance
ANOVA	Analysis of variance
CD	Continuumdistorsion
CyD	Zylinderdistorsion
DV	Digital Video
EFDMA	European Fascial Distortion Model Association
FD	Faltdistorsion
FDM	Fasziendistorsionsmodell
HP	Heilpraktiker
HTP	Hernierter Triggerpunkt
KFD	Keine Fasziendistorsion
OA	Osteopath mit Quellberuf Arzt
OEN3	Osteopath mit weniger als drei Jahren an Erfahrung mit dem FDM
OEH3	Osteopath mit mehr als drei Jahren an Erfahrung mit dem FDM
OEHM	Osteopath mit Erfahrung mit dem FDM höher als der Median der Kohorte
OENM	Osteopath mit Erfahrung mit dem FDM niedriger als der Median der Kohorte
OHP	Osteopath mit Quellberuf Heilpraktiker
OP	Osteopath mit Quellberuf Physiotherapeut
PT	Physiotherapeut
TB	Triggerband
TF	Tektonische Fixierung
U	Untersucher
v4 – v33	Bezeichnung der Videos im Fragebogen
WN	Weiß nicht

## **Einverständniserklärung**

### Einverständniserklärung

Hiermit erkläre ich mich einverstanden, während der osteopathischen Behandlung gefilmt zu werden. Die Aufnahmen werden für die Masterthese von Stefan Anker verwendet. Dafür werden die Videos in einen Online-Fragebogen eingearbeitet und anderen Osteopathen zur Begutachtung vorgeführt. Außerdem werden die Videos eventuell bei der Präsentation der Masterthese gezeigt.

Ort und Datum

Unterschrift

# Variablenspiegel

21876 - Wie übereinstimmend wird die Körpersprache von Patienten nach den Kriterien des Faszienstörungenmodells von ausgebildeten OsteopathInnen bewertet?

onlineumfragen.com - Online Umfragen, Befragungen, Studien

## Variablenspiegel (Variablen-Liste und Codierung)

### v1

Welchen Beruf üben Sie aus? Bei mehreren Antwortmöglichkeiten bitte ich Sie, die höchste universitäre Ausbildung anzugeben.

Skalenniveau: nominal

0 - missing

- 1 - ÄrztIn
- 2 - PhysiotherapeutIn
- 3 - HeilpraktikerIn

### v2

Sind Sie Osteopath?

Skalenniveau: nominal

0 - missing

- 1 - ja
- 2 - nein

### v3

Wie viele Jahre arbeiten Sie mit dem FDM? Bitte runden Sie auf ganze Jahre auf oder ab.

Skalenniveau: metrisch

### v4

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

### v5

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

### v6

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung



21876 - Wie übereinstimmend wird die Körpersprache von Patienten nach den Kriterien des Faszienstörungenmodells von ausgebildeten OsteopathInnen bewertet?

## v7

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

## v8

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

## v9

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

## v10

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

## v11

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

21876 - Wie übereinstimmend wird die Körpersprache von Patienten nach den Kriterien des Faszienstörungenmodells von ausgebildeten OsteopathInnen bewertet?

## v12

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

## v13

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

## v14

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

## v15

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

## v16

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

21876 - Wie übereinstimmend wird die Körpersprache von Patienten nach den Kriterien des Faszienstörungenmodells von ausgebildeten OsteopathInnen bewertet?

## v17

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

## v18

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

## v19

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

## v20

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

## v21

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

21876 - Wie übereinstimmend wird die Körpersprache von Patienten nach den Kriterien des Faszienstörungenmodells von ausgebildeten OsteopathInnen bewertet?

## v22

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

## v23

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

## v24

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

## v25

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

## v26

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

21876 - Wie übereinstimmend wird die Körpersprache von Patienten nach den Kriterien des Faszienstörungenmodells von ausgebildeten OsteopathInnen bewertet?

## v27

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

## v28

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

## v29

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

## v30

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

## v31

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

21876 - Wie übereinstimmend wird die Körpersprache von Patienten nach den Kriterien des Faszienstörungenmodells von ausgebildeten OsteopathInnen bewertet?

## v32

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

## v33

Welche Faszienstörung ist auf diesem Video erkennbar?

Skalenniveau: nominal

0 - missing

- 1 - Triggerband TB
- 2 - Hernierter Triggerpunkt HTP
- 3 - Continuumstörung CD
- 4 - Faltdistorsion FD
- 5 - Zylinderdistorsion CyD
- 6 - Tektonische Fixation TF
- 7 - keine Faszienstörung

# Rohdaten

Case	v1	v2	v3	v4	v5	v6	v7	v8	v9	v10	v11	v12	v13	v14	v15	v16	v17	v18	v19	v20	v21	v22	v23	v24	v25	v26	v27	v28	v29	v30	v31	v32	v33	Datum
1	3	1	3	4	1	4	3	1	1	1	1	1	2	5	3	1	3	1	2	1	1	4	1	1	2	3	5	1	5	2	5	2	12.11.2010 11:13	
2	2	1	4	4	1	3	3	1	1	5	4	7	5	3	3	1	1	0	6	1	5	1	4	5	2	3	6	1	4	6	7	5	6	27.10.2010 13:38
3	2	1	2	4	1	4	2	2	1	1	5	1	6	5	3	1	4	1	1	1	5	4	1	1	2	3	4	1	5	5	5	2	30.11.2010 19:35	
4	1	1	3	4	1	4	4	3	1	1	5	4	7	5	3	4	3	5	3	1	5	4	1	1	2	3	4	5	4	4	5	4	17.01.2011 21:19	
5	2	1	3	4	1	3	0	0	1	1	1	5	6	5	3	4	0	1	2	1	1	4	1	1	0	3	4	1	5	2	0	1	1	27.10.2010 15:58
6	3	1	2	4	1	2	3	3	1	1	5	4	6	5	3	4	4	1	2	1	1	4	1	1	2	3	4	1	5	4	0	5	4	31.10.2010 17:32
7	2	1	3	4	5	3	4	4	1	1	5	5	0	5	3	3	4	3	2	1	1	3	5	1	2	3	4	5	0	2	0	5	2	06.12.2010 9:59
8	1	1	3	4	1	4	3	3	1	1	5	4	6	5	5	4	4	1	4	1	1	4	1	5	2	3	4	1	5	5	4	5	4	13.12.2010 21:15
9	1	1	13	4	1	4	3	3	1	1	1	4	6	5	3	4	4	1	1	1	5	4	1	1	2	3	5	5	2	0	5	2	24.11.2010 23:36	
10	1	1	2	4	1	4	0	0	1	0	4	4	2	5	3	4	4	1	3	1	5	4	3	1	2	3	5	1	5	5	4	7	5	01.11.2010 15:52
11	2	1	5	4	1	3	2	2	1	1	4	1	6	5	3	4	4	1	4	1	5	4	1	1	2	3	5	1	4	1	4	5	1	11.12.2010 22:42
12	1	1	4	4	1	4	3	3	1	1	5	4	6	5	3	4	4	1	2	1	1	4	1	1	2	3	4	1	5	4	0	7	1	07.11.2010 20:10
13	3	1	7	4	1	1	3	3	1	1	5	4	6	5	1	0	4	1	2	1	1	4	1	1	2	3	5	1	5	1	5	7	1	17.11.2010 22:01
14	0	1	2	4	1	3	2	2	1	1	5	4	0	5	3	4	4	1	1	1	1	4	1	1	2	3	4	1	4	1	0	0	1	01.11.2010 18:49
15	1	1	3	4	1	4	3	3	1	1	5	4	6	5	3	4	4	1	2	1	5	4	1	1	2	3	4	1	5	4	0	5	4	28.10.2010 17:34
16	1	1	5	4	1	4	3	3	1	1	5	4	6	5	3	4	4	1	4	1	5	4	4	1	2	3	5	1	5	4	4	5	4	01.12.2010 9:38
17	0	1	3	4	1	3	4	4	1	1	5	4	2	5	3	4	4	1	1	1	5	4	1	1	2	3	5	1	5	1	4	5	1	13.11.2010 17:40
18	2	1	5	4	1	3	2	1	1	1	5	4	6	5	3	4	4	1	0	1	1	4	1	1	2	3	4	1	5	1	4	0	2	29.10.2010 8:36
19	2	1	3	4	1	3	3	3	1	1	5	4	2	5	3	1	4	1	2	1	1	4	5	1	2	3	6	1	5	4	5	5	4	30.11.2010 18:06
20	3	1	4	4	1	4	3	3	1	1	4	4	7	5	3	4	4	1	2	1	1	4	2	1	2	3	4	1	5	4	4	5	2	27.10.2010 20:58
21	1	1	2	4	1	3	3	3	1	1	5	4	2	5	3	0	4	1	2	1	1	4	1	1	2	3	4	1	5	2	4	5	2	29.10.2010 12:39
22	2	1	3	4	1	2	3	2	1	1	1	4	2	5	3	4	4	1	1	1	1	4	2	1	2	3	5	1	1	2	5	5	2	20.11.2010 10:17
23	2	1	3	4	1	3	3	3	1	1	4	6	6	5	3	4	4	1	2	1	1	4	4	1	2	3	5	1	5	1	0	0	1	14.11.2010 18:23
24	0	1	6																															30.10.2010 14:27
25	2	1	2	4	1	4	3	0	1	1	5	4	6	5	3	0	4	1	2	1	1	1	1	1	2	3	5	1	4	4	0	5	2	12.12.2010 20:48
26	1	1	5	4	1	3	2	2	1	1	5	5	2	5	3	1	4	1	5	1	1	5	1	1	2	3	4	1	5	2	4	5	1	14.12.2010 15:08



## 14. Englische Kurzzusammenfassung

# Interrater reliability in evaluating the body language based on the Fascial Distortion Model (FDM)

Stefan Anker

Vienna School of Osteopathy, Danube University Krems –  
Centre for Traditional Chinese Medicine and Complementary Medicine

**Objective:** To investigate the inter-rater reliability of the visual FDM-Body language diagnosis and to observe whether the inter-rater reliability is influenced by profession, experience and verbal statements of patients.

**Methods:** Patient videos were presented using online questionnaires, in which most patients described their physical complaints nonverbally, to twenty-five osteopaths from Germany, Austria and Switzerland. The testers diagnose of fascial distortions is based on the FDM guidelines. The results are statistically analyzed using the kappa index according to Siegel and Castellan.

**Results:** The results are statistically analyzed using the kappa index, according to Siegel and Castellan, and show acceptable levels of reliability ( $k = 0.51$ ). Here, the results of osteopaths with either medical or non-medical backgrounds are significantly better than the results of osteopaths who are physical therapist by trade. Verbal statements that are made by patients parallel to the non-verbal communication do have a significant impact on reliability, in contrast to the duration of the experience of the tester with FDM, which does not.

**Conclusion:** Seen against the background of a lack of information about validity, this study presents an acceptable inter-rater reliability of visual FDM diagnosis. These results concerning the influence of verbal statements can only be seen as an encouragement for further research.

## Introduction

The reliability of diagnostic tests is of great value in osteopathy. Only with a reliable diagnosis the appropriate treatment for a patient can be found or determined whether the patient generally responds to treatment (Lucas and Bogduk 2011). This complies with the requirements of the evidence-based medicine (EBM), which constitute reliable tests as a basic element for comprehensible behaviour (Sackett et al. 1996). Osteopathy predominantly uses manual test methods for diagnosis (Biberschick 2011). Reliability studies in the field of manual test methods show more or less acceptable results, as revealed by review articles (Seffinger et al. 2004, Hollerwöger 2006, Stochkendahl et al. 2006, Haneline and Young 2009). They also show quality problems that undermine the validity of the processed studies. Many reliability studies lack an independent reference standard and neglect therefore statements about the validity of the particular test method (Seffinger et al. 2004, Hollerwöger 2006, Stochkendahl et al. 2006). A further problem is the lack of standardization of the examination process (Stochkendahl et al. 2006, Hollerwöger 2006). In addition measurements in some cases are performed on non-representative study populations (Hollerwöger 2006, Stochkendahl et al. 2006).

In contrast to this manual approach of osteopathy the Fascial Distortion Model (FDM) supports a visual approach to clinical diagnosis, which is applied by a large number of osteopaths in practice.

The FDM is an anatomical model offers the cause of physical symptoms in six potential anatomical changes of the shape of the fascia, called fascial distortions. Through the concept of FDM, developed by osteopathic physician and emergency doctor Stephen Typaldos, a number of diseases in orthopaedics, internal medicine and neurology can be explained. As a result specific treatments can be applied and effects determined. FDM presents a specific approach to diagnosing fascial distortions, which strongly correlates to subjective verbal and nonverbal communication of the patient. Partly the diagnosis in the FDM is based on a type of visual diagnostics, as suggested by Typaldos (1999, 2002). Fascial distortions are presented through patients indicating them on their own body. Typaldos (2002) describes this process as body language, referring to consistent, unconscious movements or postures. From a scientific point of view body language is bound to be associated with the term gesture. Gesture is defined by the predominant movements of the arms and head (Duden online 2011) and by a distinct temporal process (Kendon 1996, McNeill 1992) with an initial starting, peak (stroke) and final end sequence of the gesture. These criteria are of no importance in FDM. Thus,

scientific information from gesture science can only be used with reservations for FDM.

Reliability studies on visual diagnosis of gestures and body language are incidental in medical databases. Some comparable studies in the field of cardiology (Edmondstone 1995, Albarran et al. 2000, Marcus et al. 2007) show that gesture obviously holds certain information and reference matter in diagnosing myocardial infarction. In the cited studies (Edmondstone 1995, Albarran et al. 2000, Marcus et al. 2007) no strong evidence of acceptable reliability could be shown. In contrast to these studies visual diagnosis in FDM lacks independent reference standards. Fascial distortions and their diagnosis through FDM body language are part of the model-like presentation of FDM and are currently of no consequence outside this system in general medicine. Since the FDM-diagnosis was not investigated regarding its inter-rater reliability, no statements concerning the validity of FDM tests are possible.

Generally, FDM distinguishes six fascial distortions. These are distinguished by their formation mechanism, location and quality of complaints, their extend and, above all, through their body language by which they can be encoded in the visual FDM diagnosis (see Table 1).

<b>Fascial distortions</b>	<b>Definition, complaints and verbal description</b>	<b>Body language</b>
<b>Triggerbands</b>	Deformation of banded fascia, which is caused by twisting, cutting, tearing and wrinkling of fibres; patients mostly describe burning or aching pain along a subjectively defined anatomical process; movement restrictions, loss of strength or swelling are other possible indications of a triggerband	Stroking along painful/tender parts with several fingers
<b>Herniated triggerpoints</b>	Protrusion of deeper tissue through overlying fascial layer; is found almost exclusively on the core; causing dull pain, also in the region around the herniation; possible restriction of movement of the adjacent joints	Press with fingers, thumbs or knuckles in the protrusion area
<b>Continuum distortions</b>	Distortion of the transition zone between bone and ligament, tendon or other fascial structures; pain or restricted movement in one axis as a result	Point a finger at one or more painful spots near bones

<b>Fascial distortions</b>	<b>Definition, complaints and verbal description</b>	<b>Body language</b>
<b>Folding distortions</b>	3-D deformation of the joint covering fascia, intramuscular septum or the interosseous membrane; symptoms are painful movement, pain in the lower levels, feeling of instability, causing slight movement restrictions; distinction between Unfolding and Refolding distortions which can be sometimes seen in different body language	Rubbing or pulling of the joint with the hand or hands; hand is put (depending on availability) on the affected area; fists pressed into the back, supporting the back or relieving it with the help of the hands; fingers stroking parallel to the joint line is a reference of a refolding distortion; pressing several fingers between the muscle venter or bones, pulling or tugging on the muscles is presented in folding distortions of interosseous membranes and inter-muscular septa; folding distortions are often displayed dynamically with repetitive movement of the affected body part
<b>Cylinder distortions</b>	Deformation of the spiral turns of the circular fascia; possible symptoms such as paraesthesia, cramps, tremors, paralysis, etc.; there is often a large discrepancy between shutter release and the present complaints; the pain is felt "in depth"; from outside usually not palpable	Massaging or stroking of the soft tissues, wiping hand movements, on occasion the body part is not touched at all
<b>Tectonic fixation</b>	Loss of slippage of fascial surfaces; causes little pain, however, limited mobility on all levels and the feeling of stiffness	Holding and massaging of the joint; patients try to mobilize themselves with own strength; shoulder abduction is only possible with anterior elevation, in the abdominal position the arm cannot be stretched forward horizontally without the upper body lifting off the pad; in case of stiffness in the hip placing palms on both christae iliac; holding and rubbing of the joint

Table 1: Characteristics of the six fascial distortions and their body language (compare Typaldos 1999, 2002)

In FDM diagnosis body language is entwined with verbal description of the complaints of the patients, movement and exercise tests and palpation, which is mainly used in the FDM for differential diagnosis.

The present study examines how reliable osteopaths evaluate body language in terms of the Fascial Distortion Model. Additionally, the influence of the profession of the osteopath (physician, medical practitioner, physiotherapist), experience with FDM and further verbal statements of the patients on the inter-rater reliability should be investigated. Here, the points of criticism of the quality of reliability studies (see above) will be considered in the study design.

## **METHODS**

The study was designed as a video-assisted online questionnaire.

### **Videos**

50 short videos of patients describing their complaints non-verbally are filmed, qualitatively retouched and divided into sequences in the practice of the investigator. First, 23 of those videos are randomly selected. Another three of these videos, which appeared twice in the questionnaire with and without sound are randomly chosen. Further two videos that contain no diagnosable body language are added. In result 28 assessable videos are integrated in the online questionnaire.

### **Online questionnaire**

The online questionnaire allows a high degree of standardization of the examination process. It consists of 33 questions. First, the testers are informed about the investigation process, the conditions for the execution of the questionnaire and criteria of how to answer the questions. It is a single-choice questionnaire, and once the testers have watched the video, they must decide which of the six fascial distortions can be diagnosed on the basis of FDM criteria. Additional answers are possible: "Keine Faszien distortion (No fascial distortion)" if no fascial distortion is visible, or "Weiß nicht (Do not know)" if the question cannot be answered. The submission of a response is necessary to continue the questionnaire. Subsequent alteration of the responses is not possible. The first three questions refer to the testers' original profession, their training in osteopathy and duration of the experience with FDM. Two test videos are shown to the testers to practice the examination process. Their results are not included in the statistics. The 28 questions of the study (v6 to v33) are to be answered according to the presented method. Three videos, first without then with sound, are shown to investigate the impact of verbal statements alongside with non-verbal body language have on reliability.

## Testers

The testers must fulfil following criteria: trained osteopaths, a FDM training, according to the guidelines of the EFDMA (European Association FDM 2010), and sufficient knowledge of the German language. 79 therapists (doctors, physiotherapists, medical practitioners) from Germany, Austria and Switzerland are invited per email, and 25 of those meet all the necessary criteria. There are 5 female and 20 male testers who will answer the questionnaire in full. This data then is included in the statistics.

## Statistics

For the evaluation of reliability the kappa index (k index), according to Siegel and Castellan (1988), is used. In contrast to Cohen's kappa, which is limited to two investigators, this is a generalized form of the kappa index. This means that the compliance of all investigators can be expressed in one index. Another advantage is the possibility in the case of partially absent data to calculate the index. In the case of this study the possible answer "Weiß nicht (Do not know)" is interpreted as lack of response and therefore makes this statistical method of calculation necessary.

The proposed textual interpretation according to Landis and Koch (1977) can be applied for the kappa index of Siegel and Castellan (1988), (see Table 2).

$k < 0,20$	poor	schwach
$0,20 < k < 0,40$	adequate	leidlich
$0,40 < k < 0,60$	moderate	mittelmäßig
$0,60 < k < 0,80$	substantial	beträchtlich
$0,80 < k < 1,00$	almost ideal	fast ideal

Table 2: Interpretation of reliability scores by Landis and Koch (1977)

According to Fjellner et al. (1999) values by at least 0.4 are considered to be an acceptable indicator of inter-individual reliability.

In addition to the generalized kappa index, all of the investigator-pairs (according to Siegel and Castellan 1988) are calculated in order to get an overview of the spread of the kappa indices. This data also forms the basis for further comparisons of the various subgroups.

Subsequently, through variance analysis and significance tests (independent samples t-tests or Wilcoxon tests) is examined to test whether there are significant differences between the various sub-groups (see Table 3) in terms of expectation value.

Osteopaths		
A1. Doctors	A2. Physiotherapists	<b>ANOVA</b>
A1. Doctors	A3. Medical practitioners	
A2. Physiotherapists	A3. Medical practitioners	
A4. Long experience (>3 years)	A5. Short experience (1-3 years)	t-Test
A6. Long experience (>2 years)	A7. Short experience (1 and 2 years)	t-Test

Table 3: Analysis of variance (ANOVA) and significance tests (Independent samples t-Test. .. t-test, Wilcoxon Wilcoxon test ...) as to test the expected value of the k-index compared among sub-groups.

To examine whether verbal statements by the patient influence the results three videos, with and without sound, are presented to the testers. The reviews of these videos present the base of the data for the calculation of the k-indices for the corresponding pairs of the investigators.

These in turn were compared with each other by means of significance tests (independent variable "sound"), where the kappa indices were used for substitution of the negative value of zero, since their results are widely comparable to the generalized kappa index.

Furthermore, the results of the individual questions were specifically investigated using the qualitative criteria of gesture studies in order to obtain possible correlations between the inter-rater reliability for individual videos and their nature and content.

## RESULTS

The results relating to this research can be summarized as follows (see Table 4; page 118):



The overall index is  $k = 0.51$  (according to Siegel and Castellan, 1988) and it is significantly different from zero ( $Z = 53.42$ ,  $p < 0.0001$ ). The mean average of all paired comparisons from the reviews of Osteopaths calculated  $k$  indices adds up to  $k = 0.57 \pm 0.08$  (95% CI: 0.54 to 0.60).

**Consequently one can speak of an acceptable inter-rater reliability (after Fjellner et al. 1999) of the FDM visual diagnosis in the context of this study.**

The mean  $k$ -index of the physical therapists in the sample is significantly different both from that of doctors ( $t = 5.180$ ,  $df = 68.2$ ,  $p < 0.0001$ ) and that of the medical practitioners ( $t = 3.040$ ,  $df = 7.8$ ,  $p = 0.02$ ). The mean  $k$ -value of the physical therapist is lower by 0.14 or 0.13 compared to the one of the control groups. There is no significant difference (difference:  $k = 0.01$ ,  $t = 0.1900$ ,  $df = 8.2$ ,  $p = 0.85$ ) between the reliability of the opinions of doctors and that of the practitioners. Doctors reach an average index of  $k-k = 0.57 \pm 0.13$  (95% CI: 0.53 to 0.61), the Therapy  $k = 0.56 \pm 0.10$

Fascial distortion	1 - Triggerband TB	2 - Herniated triggerpoint HTP	3 - Continuum distortion CD	4 - Folding distortion FD	5 - Cylinder distortion CyD	6 - Tectonic fixations TF	7 - No fascial distortion		Valid answers (n)	Do not know - Missings
V6	1	2	<b>11</b>	<b>11</b>	0	0	0		25	
V7	0	5	<b>15</b>	3	0	0	0		23	2
V8	2	5	<b>13</b>	2	0	0	0		22	3
V9	<b>25</b>	0	0	0	0	0	0		25	
V10	<b>24</b>	0	0	0	0	0	0		24	1
V11	4	0	0	4	<b>17</b>	0	0		25	
V12	3	0	0	<b>18</b>	3	1	0		25	
V13	0	7	0	0	0	<b>13</b>	3		23	2
V14	0	0	0	0	<b>25</b>	0	0		25	
V15	1	0	<b>23</b>	0	1	0	0		25	
V16	5	0	1	<b>16</b>	0	0	0		22	3
V17	1	0	2	<b>21</b>	0	0	0		24	1
V18	<b>22</b>	0	1	0	1	0	0		24	1
V19	5	<b>12</b>	2	3	1	1	0		24	1
V20	<b>25</b>	0	0	0	0	0	0		25	
V21	<b>16</b>	0	0	0	9	0	0		25	
V22	2	0	1	<b>21</b>	1	0	0		25	
V23	<b>17</b>	2	1	3	2	0	0		25	
V24	<b>23</b>	0	0	0	2	0	0		25	
V25	0	<b>24</b>	0	0	0	0	0		24	1
V26	0	0	<b>25</b>	0	0	0	0		25	
V27	0	0	0	<b>13</b>	10	2	0		25	
V28	<b>22</b>	0	0	0	3	0	0		25	
V29	1	0	0	4	<b>19</b>	0	0		24	1
V30	6	7	0	<b>8</b>	3	1	0		25	
V31	0	0	0	<b>10</b>	5	0	1		16	9
V32	1	0	0	0	<b>18</b>	0	3		22	3
V33	8	<b>9</b>	0	6	1	1	0		25	

Table 4: Frequency of diagnoses by the 25 osteopaths for each individual patient. (bold & italic ... most likely diagnosis, dark green mark: > 90% (n > 22) line, light green ... > 80% (n > 20), light yellow ... > 70% (n > 17) and orange ... > 60% (n > 15).

(95% confidence interval: 0.48-0.64) and the physiotherapists  $k = 0.43 \pm 0.11$  (95% CI: 0.40 to 0.46).

Even without considering the examiners 2 and 7, who have significantly poorer agreement with the other physical therapists, the reliability lies, however, at  $k = 0.49 \pm 0.08$  (95% CI: 0.46 to 0.52), which is significantly lower than that of the doctors (Wilcoxon  $W = 687.5$ ,  $p = 0.01$ ).

On average, there was no significant difference in reliability of those osteopaths with more than two years experience with the FDM and those with less than three years experience (difference:  $k = 0.03$ ,  $t = 1.097$ ,  $df = 17.5$ ,  $p = 0.28$ ). Even taking into account a different threshold for high and low experience no significant differences were observed. The difference between the mid-level reliability of those osteopaths with more than three years experience with the FDM and those with less than four years experience is  $k = 0.01$  ( $t = 0.091$ ,  $df = 53.0$ ,  $p = 0.93$ ).

The comparison of the average of the k-indices, which were calculated from the assessments of three videos with sound and those without sound, shows a significant difference (Wilcoxon  $W = 40598.5$ ,  $p = 0.03$ ). With sound, the average value amounts to  $k = 0.24 \pm 0.30$  (95% CI: 0.20 to 0.27), without sound  $k = 0.17 \pm 0.26$  (95% CI: 0.14 - 0.20).

## **DISCUSSION**

The visual diagnostics seems to be a reliable test for the diagnosis of fascial distortions according to the FDM. The acceptable results of this investigation throw a positive light on this diagnostic approach, applied by osteopaths in practice.

This is the first study that examines the inter-rater reliability of the FDM-body language and there is no independent reference standard for the diagnosis of fascial distortions. These factors must be mentioned in the evaluation of the results.

Based on the qualitative analysis of individual results of each video a greater spreading of results of the videos with lower tester agreement was observed. Interestingly, with the decline of the agreements results accumulated which cannot be related to the criteria for the diagnosis of specific distortion types.

In the same videos also a build up of incomplete temporal gesture sequence in the sense of Kendon (1996) and McNeill (1992) is evident. That could make it difficult

for the tester to detect the peak gesture followed by a smaller agreement level. To what extent these observations relate to the structure of the questionnaire in its present form, as well as the video section or the lack of FDM criteria for body language in itself, remains unclear at this point.

Further the results show that from 700 submitted answers only in 35 cases the testers responded with the options "Keine Faszienverzerrung (No fascial distortion)" or "Weiß nicht (Do not know)", although at least two of the videos showed no evidence of encodable body language in terms of FDM. This raises the question of how the criteria of body language vocabulary of FDM are defined.

The different results of inter-rater reliability with respect to the testers' profession challenge the findings of Seffinger et al. (2004) and Stochkendahl et al. (2006), which cannot confirm any connection between the testers and the inter-rater reliability in their reviews.

A possible explanation for this statistical difference might be a different composition of the individual test groups in relation to their experience or gender; these criteria are not included in the statistics. Due to the low variance in the range of volume of experience of the testers, the impacts of these results in this study are unlikely. Through a more consistent distribution of experience the results for the investigation of the relationship between experience and inter-rater reliability could be affected.

In this study no correlation between the experience of the tester and the inter-rater reliability can be found. This may be related to the selection of testers according to their experience. A more consistent distribution of experience with FDM among the testers could alter the results. This result correlates with the reviews of Seffinger et al. (2004), Stochkendahl et al. (2006) and Haneline and Young (2009), which cannot establish a general relationship of reliability and experience.

There is a significant impact of verbal statements on the inter-rater reliability. Though the limiting to three video pairs, these results are to be examined and thus must be seen as a stimulus for further research. Here, a much larger sample group would be necessary to give more weight to these results. In the case of this study the consistent diagnostic results are obtainable through verbalization.

Looking at the results of this study following can be criticised. A higher number of videos in the questionnaire, with significant differences of filmed patients regarding

gender, symptoms, affected body regions and so on, could better represent the wide variety of FDM body language and thus reinforce the significance of the results and their application in practice. The separating of the production, especially the post processing, of the videos from the investigator should be considered to obtain a more blinding effect in this context.

The videos in this study are partially edited to reduce possible internal errors. The videos are also faded in and out, which could create an artificial gestural climax of the individual videos that may not be comparable to the "natural" peak of the gesture. This should be considered in future studies solving the problem through editing or longer video clips. The potential increase of the internal errors in single choice study design should be considered in this context.

The video setting within an online questionnaire shows no possibility of criticism or ability to enquire upon lack of clarity of the seen body language, which doesn't correspond with the practical approach.

Because of technical factors, it is not possible to prevent a repeated viewing of the videos by the testers. Although the reproducibility itself is an advantage, when used for video recording in reliability analysis (Haidet et al. 2009) and in measurement accuracy (Sommerfeld 2006), it should be noted that this is not relevant in this code of practice where body language in the same manner is mostly shown only once. Further critique of the structure of the online questionnaire it is whether the testers can actually distinguish between the possible answer "Keine Faszien distortion (No fascial distortion)" and "Weiß nicht (Do not know)". A wider awareness of the theoretical background of these choices and associated research questions would be necessary to generate corresponding sensitivity on part of testers.

The main problem in this study, as with most reliability studies in the field of osteopathic manual medicine, is the discrepancy between internal and external validity. The reductionist approach in such scientific studies only partially reflects the clinical practice. In the study at hand, the problem is that body language occurs only rarely in isolation, and through the observation within the osteopath-patient communication, it occurs in the context of verbal statements and the patient's medical history. Thus for this study one cannot say anything about the reliability of the FDM-diagnosis in general, but only on their visually evident part that has to be (as described) cross-linked in a particular case with other information from the investigation in order to attain FDM diagnostic approach.

## CONCLUSION

Using a video-assisted online survey, the body language shows acceptable inter-rater reliability in this master thesis. Since this visual diagnostic approach cannot currently be reviewed by an independent reference standard the validity of body language-diagnosis is not clear. These results should be transferred with caution into practice. In particular the use of the test as a singular diagnostic criterion is critical. In the context of the entire FDM-specific findings for the diagnosis of fascial distortions, the results of this study, however, are significant.

## Acknowledgement

Thanks to Mag<sup>a</sup>. Claudia Gamsjäger for the care of this master thesis.

## Literature

Albarran JW, Durham B, Chappel G, Dwight J, Gowers J (2000): Are manual gestures, verbal descriptors and pain radiation as reported by patients reliable indicators of myocardial infarction? Preliminary findings and implications. *Intensive and Critical Care Nursing*, 16(2), 98 – 110

Biberschick M. (2011): "Legt Euch ein Schema zurecht!": die Routineuntersuchung in der Osteopathie. Master Thesis zur Erlangung des akademischen Grades "Master of Science" in Osteopathie. Krems. Donau Universität Krems

Duden online. (2011): <http://www.duden.de/>

European FDM Association EFDMA. (2010): FDM-Ausbildung. <http://www.fdm-europe.com/fdm-ausbildung/ausbildung-fdm--curriculum/index.html> (Access: Oktober 10, 2010)

Edmondstone WM (1995): Cardiac chest pain: Does body language help the diagnosis? *British Medical Journal*, 311(7021), 1660 – 1661

Fjellner A, Bexander C, Faleji R, Strender LE. (1999): Interexaminer reliability in physical examination of the cervical spine. *J. Manipulative Physiol Ther.* 22, 511 – 516

Haidet KK, Tate J, Divirgilio-Thomas D, Kolanowski A, Happ MB. (2009): Methods to Improve Reliability of Video Recorded Behavioral Data. *Res Nurs Health*, 32 (4): 465 - 474

Haneline MT, Young M. (2009): A review of intraexaminer and interexaminer reliability of static spinal palpation: a literature synthesis. *J Manipulative Physiol Ther.* 32, 379 – 386

Hollerwöger D. (2006): Methodological quality and outcomes of studies addressing manual cervical spine examinations: a review. *Man Ther.* ,11, 93-98

Kendon A. (1996): An Agenda for Gesture Studies. *Semiotic Review of Books.* 7 (3), 8 – 12

Landis, JR, Koch GG. (1977): The measurement of observer agreement for categorical data. *Biometrics.* 33, 159 – 174

Lucas N, Bogduk N (2011): Diagnostic reliability in osteopathic medicine. *International Journal of Osteopathic Medicine*,1, 1 - 5

Marcus GM, Cohen J, Varosy PD, Vessey J, Rose E, Massie BM, Chatterjee K, Waters D. (2007): The utility of Gestures in Patients with Chest Discomfort. *The American Journal of Medicine*, 120, 83 – 89

McNeill D. (1992): *Hand and mind: What gestures reveal about thought.* Chicago: The University of Chicago Press

Sackett DL, Rosenberg WMC, Gray Muir JA, Haynes RB, Scott Richardson W. (1996): Evidence based medicine: what it is and what it isn't. *British Medical Journal*, 312, 71

Seffinger MA, Najm WI, Mishra SI, Adams A, Dickerson VM, Murphy LS, Reinsch S (2004): Reliability of spinal palpation for diagnosis of back and neck pain: a systematic review of the literature. *Spine*, 1, 29, 413 – 25

Siegel S, Castellan NJ Jr. (1988): *Nonparametric statistics for the behavioural sciences.* Boston: McGraw-Hill

Sommerfeld P (2006): Methodologie. Kleine und hoffentlich kurzweilige Einführung in den Sinn und Unsinn wissenschaftlichen Arbeitens auf der Grundlage quantitativer Methoden für OsteopathInnen. 3. Überarbeitete Version.

Stochkendahl MJ, Christensen HW, Hartvigsen J, Vach W, Haas M, Hestbaek L, Adams A. Bronfort G (2006): Manual Examination of the Spine: A Systematic Critical Literature Review of Reproducibility. *Journal of Manipulative and Physiological Therapeutics*, 29, 6, 475 – 85

Typaldos S. (1999): Orthopathische Medizin. Die Verbindung von Orthopädie und Osteopathie durch das Faszienmodell. Kötzing/Bayer. Wald: Verlag für Ganzheitliche Medizin Wühr

Typaldos S. (2002): FDM – Clinical and Theoretical Application of the Fascial Distortion Model within the Practice of Medicine and Surgery. Bangor: Published by the editor