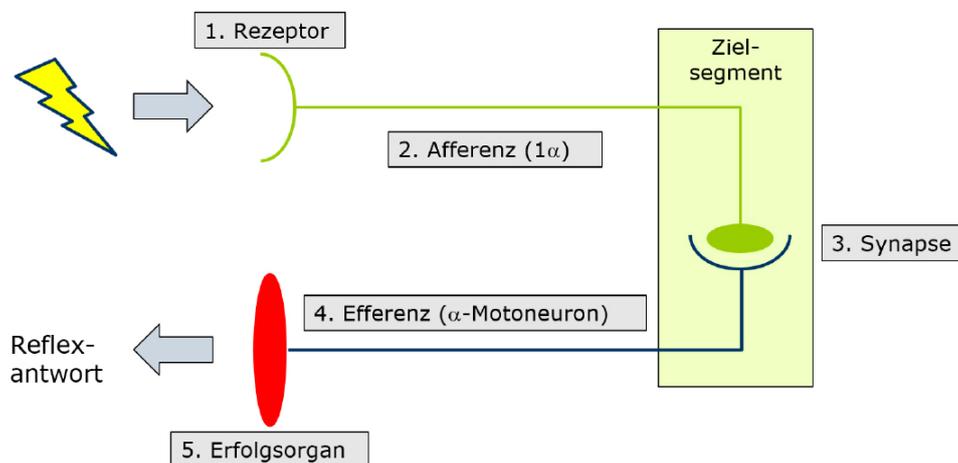


Reliabilitätsanalyse der subjektiven Beurteilung des Bizepssehnenreflexes beim gesunden Hund



Inaugural-Dissertation
zur Erlangung des Grades eines
Doktors der Veterinärmedizin
an der
Freien Universität Berlin

vorgelegt von
Felix Giebels
Tierarzt aus Krefeld-Hüls

Berlin 2018
Journal-Nr.: 4060

**Aus der Klinik für Kleine Haustiere
des Fachbereichs Veterinärmedizin
der Freien Universität Berlin**

**Reliabilitätsanalyse der subjektiven Beurteilung des Bizepssehnenreflexes beim
gesunden Hund**

**Inaugural-Dissertation
zur Erlangung des Grades eines
Doktors der Veterinärmedizin
an der
Freien Universität Berlin**

**vorgelegt von
Felix Giebels
Tierarzt
aus Krefeld-Hüls**

**Berlin 2018
Journal-Nr.: 4060**

Gedruckt mit Genehmigung des Fachbereichs Veterinärmedizin
der Freien Universität Berlin

Dekan: Univ.-Prof. Dr. Jürgen Zentek
Erster Gutachter: Univ.-Prof. Dr. Barbara Kohn
Zweiter Gutachter: Univ.-Prof. Dr. Franck Forterre
Dritter Gutachter: PD Dr. Roswitha Merle

Deskriptoren (nach CAB-Thesaurus):
dogs, tendons, muscles, reflexes, patella, video recordings, diagnostic techniques

Tag der Promotion: 08.06.2018

Bibliografische Information der *Deutschen Nationalbibliothek*
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen
Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über
<<https://dnb.de>> abrufbar.

ISBN: 978-3-86387-906-8
Zugl.: Berlin, Freie Univ., Diss., 2018
Dissertation, Freie Universität Berlin
D188

Dieses Werk ist urheberrechtlich geschützt.
Alle Rechte, auch die der Übersetzung, des Nachdruckes und der Vervielfältigung des Buches, oder
Teilen daraus, vorbehalten. Kein Teil des Werkes darf ohne schriftliche Genehmigung des Verlages in
irgendeiner Form reproduziert oder unter Verwendung elektronischer Systeme verarbeitet,
vervielfältigt oder verbreitet werden.

Die Wiedergabe von Gebrauchsnamen, Warenbezeichnungen, usw. in diesem Werk berechtigt auch
ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen-
und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt
werden dürfen.

This document is protected by copyright law.
No part of this document may be reproduced in any form by any means without prior written
authorization of the publisher.

Alle Rechte vorbehalten | all rights reserved
© Mensch und Buch Verlag 2018 Choriner Str. 85 - 10119 Berlin
verlag@menschundbuch.de – www.menschundbuch.de

Alles Wissen beruht auf der Übereinstimmung eines Objektiven mit einem Subjektiven.

(Friedrich Schelling)

Meiner Familie

Inhaltsverzeichnis

1. Einleitung	5
2. Literaturübersicht	6
2.1. Allgemeine neurophysiologische und neuroanatomische Grundlagen des Spinalreflexes	6
2.1.1. Der Reflexbogen	6
2.1.2. Motoneurone	9
2.1.2.1. Oberes Motoneuron	9
2.1.2.2. Unteres Motoneuron	11
2.2. Spezielle topographische Anatomie.....	12
2.2.1. Anatomie des Bizepssehnenreflexes	12
2.2.2. Anatomie des Patellarsehnenreflexes	13
2.3. Die Reflexevaluierung als medizinischer Parameter	14
2.3.1. Definitionen	14
2.3.1.1. Reflexantwort	14
2.3.1.2. Reflexpräsenz	14
2.3.1.3. Reflexaktivität	14
2.3.2. Die Entdeckung der Reflexantwort.....	15
2.3.3. Beeinflussung der Reflexantwort.....	15
2.3.3.1. Einfluss der neuroanatomischen Lokalisation der Läsion	15
2.3.3.1.1. Läsion des oberen Motoneurons.....	16
2.3.3.1.2. Läsion des unteren Motoneurons.....	16
2.3.3.2. Subjektive Einflussfaktoren	17
2.3.4. Methoden der Standardisierung der Reflexevaluierung.....	19
2.3.4.1. Standardisierung der neurologischen Untersuchung	19
2.3.4.2. Standardisierung der Beurteilung der Reflexantwort	21
2.3.5. Einsatzbereich der Reflexevaluierung.....	27
3. Material und Methoden	28
3.1. Versuchsaufbau	28
3.2. Probanden.....	28

3.3. Methoden.....	29
3.3.1. Unterteilung der Probanden.....	29
3.3.2. Auswahl und Unterteilung der Beurteiler.....	30
3.3.3. Voruntersuchung.....	32
3.3.4. Videodokumentation.....	33
3.3.4.1. Durchführung der Reflexevaluierung.....	33
3.3.4.2. Randomisierung.....	34
3.3.4.3. Digitale Videobearbeitung.....	35
3.3.5. Videoanalyse.....	38
3.3.6. Statistische Auswertung.....	38
3.3.6.1. Deskriptive Statistik.....	38
3.3.6.2. Reliabilitätsanalyse.....	38
3.3.6.2.1. Prozentuale Übereinstimmung.....	39
3.3.6.2.2. Kappa-Analyse.....	40
3.3.6.2.3. Intraklassen-Korrelationskoeffizient.....	49
3.3.6.2.4. Table of Agreement.....	50
3.3.6.3. Interpretation der Reliabilitätsanalyse.....	51
3.3.6.3.1. Prävalenz-Index (<i>PI</i>).....	52
3.3.6.3.2. Bias-Index (<i>BI</i>).....	54
3.3.6.3.3. Maximal erzielbarer Kappa (K_{max}).....	55
3.3.6.3.4. Interpretation der Ergebnisse unter dem Kriterium der <i>klinischen Akzeptanz</i>	57
3.3.6.4. Regressionsanalyse.....	57
4. Ergebnisse.....	64
4.1. Studie 1: Einfluss des Untersuchers auf das Interobserver Agreement bei der Evaluierung des Bizepssehnenreflexes.....	64
4.1.1. Probanden in Studie 1.....	64
4.1.2. Kategoriale Unterteilung der Probanden.....	67
4.1.3. Digitale Videobearbeitung.....	67
4.1.4. Beurteilerübereinstimmung.....	69
4.1.5. Regressionsanalyse.....	75
4.2. Studie 2: Vergleich des Interobserver Agreement zwischen dem Bizeps- und dem Patellarsehnenreflex.....	78

4.2.1. Studie 2a: Reliabilitätsanalyse des Bizepssehnenreflexes.....	78
4.2.1.1. Probanden in Studie 2a	78
4.2.1.2. Kategoriale Unterteilung der Probanden.....	78
4.2.1.3. Digitale Videobearbeitung	78
4.2.1.4. Beurteilerübereinstimmung	79
4.2.1.5. Regressionsanalyse	86
4.2.2. Studie 2b: Reliabilitätsanalyse des Patellarsehnenreflexes	89
4.2.2.1. Probanden in Studie 2b	89
4.2.2.2. Kategoriale Unterteilung der Probanden.....	92
4.2.2.3. Digitale Videobearbeitung	93
4.2.2.4. Beurteilerübereinstimmung	93
4.2.2.5. Regressionsanalyse	100
5. Diskussion	102
5.1. Hypothese	102
5.2. Auswahl der zu untersuchenden Reflexe	102
5.3. Skizzierung des Versuchsaufbaus und Versuchsdurchführung.....	103
5.4. Beurteilung der Reflexantwort anhand eines Videos	106
5.5. Auswahl und Limitationen der verwendeten Reliabilitätskoeffizienten.....	108
5.6. Beeinflussung des Interobserver Agreement.....	113
5.7. Ausblick	120
6. Zusammenfassung	122
7. Summary	124
8. Literaturverzeichnis	CXXVI
9. Anhang.....	CXLIX
9.1. Evaluierungsbögen.....	CXLIX
9.2. Evaluierungen der Beurteiler in Studie 1	CLIV
9.3. Evaluierungen der Beurteiler in Studie 2a	CLXXII
9.4. Evaluierungen der Beurteiler in Studie 2b	CLXXXII
10. Vorabveröffentlichungen.....	CXCV

11. Danksagung.....	CXCVI
12. Selbstständigkeitserklärung	CXCVII

1. Einleitung

Die Evaluierung der Reflexantwort segmentaler Reflexe ist ein fester Bestandteil der neurologischen Untersuchung in der Human- und Veterinärmedizin. Durch sie erhält der Untersucher wichtige Informationen über die Integrität des untersuchten Reflexbogens, über die neuroanatomische Lokalisation und den Schweregrad einer Läsion (Dick, 2003; Forterre et al., 2008; Jaggy und Spiess, 2007; Kornberg et al., 2012; Lebedowska und Fisk, 2003; Lemoyne et al., 2008; Litvan et al., 1996; Seim, 2009). Sie dient als klinischer Verlaufspareter während eines Krankheitsprozesses (Carlstedt et al., 2012; Cozens et al., 2000; Kiernan et al., 1999; Lemoyne et al., 2008; Muguet-Chanoit et al., 2012) und als Indikator für die Wirksamkeit verschiedener Pharmazeutika (Lemoyne et al., 2008; Lervik et al., 2012; Rohrbach et al., 2009).

Die Qualität eines Parameters zeichnet sich im Allgemeinen durch seine Messbarkeit und seine Reliabilität in der Beurteilung aus. Die Messbarkeit eines Reflexes stützt sich im klinischen Alltag meist auf subjektiven Beurteilungen, was dazu führt, dass seine Reliabilität in der Literatur unterschiedlich bewertet wird. Verschiedene Versuche wurden unternommen, die klinische Evaluation der Reflexantwort zu standardisieren (Borin et al., 2012; Dafkin et al., 2013; Dafkin et al., 2014). Unter klinischen Bedingungen sind, im Gegensatz zur Humanmedizin, die Möglichkeiten einer Standardisierung in der Veterinärmedizin allerdings deutlich begrenzt.

Die Untersuchung des Bizepssehnenreflexes beim Hund findet in der veterinärmedizinischen Literatur nur selten Erwähnung. Angeführte Gründe hierfür sind die vermeintliche Schwierigkeit der Durchführung (Braund und Sharp, 2002; Oliver et al., 1997; Schatzberg et al., 2012) sowie seine häufig angezweifelte Aussagekraft (Braund und Sharp, 2002; Oliver et al., 1997; Schatzberg et al., 2012). In der Humanmedizin findet die Untersuchung des Bizepssehnenreflexes hingegen breitere Beachtung (Dick, 2003). Sie dient hier unter anderem der Diagnose und Verlaufsbeurteilung zervikaler Läsionen (Kiernan et al., 1999), der Integritätsprüfung des fünften und sechsten Zervikalsegments sowie des Plexus brachialis (Miller et al., 1995b) oder als Indikator einer Hemiparese infolge eines Schlaganfalls (Condliffe et al., 2005). Demgegenüber stellt der Patellarsehnenreflex in der Human- wie auch in der Veterinärmedizin einen sehr zuverlässigen Parameter dar und ist somit Standard im Rahmen der neurologischen Untersuchung (De Lahunta und Glass, 2009a; Litvan et al., 1996; Stoffel, 2011a, Schatzberg et al., 2012).

Ziel dieser Arbeit ist eine vergleichende Reliabilitätsanalyse der Beurteilung der Reflexantwort des Bizepssehnen- und Patellarsehnenreflexes bei klinisch, neurologisch sowie orthopädisch unauffälligen Hunden. Es soll untersucht werden, wie sich die Zuverlässigkeit ihrer Beurteilung vor dem Hintergrund der Beurteilerexpertise und verschiedener in der Literatur diskutierter individueller Einflussfaktoren verändert und unterscheidet.

2. Literaturübersicht

2.1. Allgemeine neurophysiologische und neuroanatomische Grundlagen des Spinalreflexes

2.1.1. Der Reflexbogen

Der Begriff *Reflexbogen* bezeichnet einen „Schaltkreis“ von mindestens zwei Neuronen und stellt die kürzeste Verbindung zwischen Rezeptor und Effektor eines bestimmten neuronalen Erregungsablaufs dar (Bronisch, 1979; Pschyrembel und Dornblüth, 2002; Stoffel, 2011b).

Reflexbögen unterscheidet man zum einen anhand der Anzahl der verschalteten Synapsen und zum anderen anhand der anatomischen Lage des Detektororgans. So steht beim *monosynaptischen* Reflexbogen das sensorische Neuron direkt mit dem motorischen Neuron (siehe unten) in Verbindung, während beim *polysynaptischen* Reflexbogen Interneurone zwischengeschaltet sind (Klein und Cunningham, 2012; Pearson und Jordan, 2000; Stoffel, 2011b). Während beim *Eigenreflex* der Rezeptor selbst im Effektororgan liegt, befindet sich der Rezeptor beim *Fremdreflex* außerhalb des Effektororgans (Stoffel, 2011b). Die Eigenreflexbögen spinaler Reflexe sind meist monosynaptisch (Pearson und Jordan, 2000; Stoffel, 2011b). Fremdreflexe sind im Rückenmark mit motorischen Neuronen der ipsi- und auch der kontralateralen Seite verschaltet und somit polysynaptisch (Stoffel, 2011b). Da es sich sowohl beim Bizepssehnen- wie auch beim Patellarsehnenreflex definitionsgemäß um monosynaptische Eigenreflexe handelt, werden für die Charakterisierung des Reflexbogens im Weiteren exemplarisch die Komponenten des monosynaptischen Eigenreflexes erörtert.

Hinsichtlich der Richtung der Erregungsleitung lässt sich der Reflexbogen in eine afferente und eine efferente Bahn unterteilen, wobei sich ihre Modalitäten im Rahmen der beiden oben

genannten Reflexe auf generell somatische Afferenzen und generell somatische Efferenzen beschränken (Stoffel, 2011a).

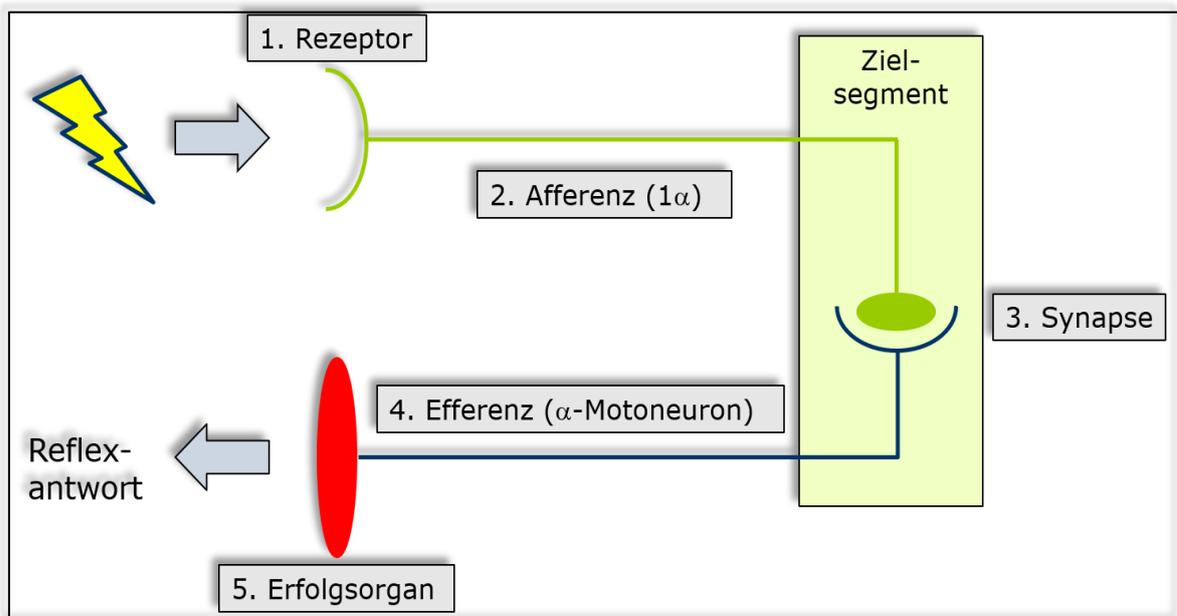
Der Reflexbogen besteht aus fünf Komponenten (Klein und Cunningham, 2012), von denen jede einzelne für die physiologische Reflexantwort funktionsfähig sein muss (s. Abbildung 1):

1. Der sensorische **Rezeptor** detektiert den Stimulus in der Peripherie und wandelt diesen in einen elektrischen Impuls in Form eines Aktionspotentials um (Klein und Cunningham, 2012). Beim Eigenreflex fungieren Muskelspindeln innerhalb des Erfolgsorgans (siehe unten) als Propriozeptoren (Stoffel, 2011b). Die Perkussion eines Muskelbauches oder seiner Sehne mit einem Reflexhammer im Rahmen der neurologischen Untersuchung führt zu einer passiven Überdehnung der Muskelfasern und somit der zwischen ihnen liegenden Muskelspindeln (Pearson und Jordan, 2000). Die Dehnung der, in der Äquatorialebene der Muskelspindel liegenden, intrafusalen Kernsackfasern erzeugt ein Aktionspotential an der Synapse in der annulospiraldendritischen Zone.
 2. Das Aktionspotential wird über die schnellleitenden 1a-Fasern des Dendrons des **sensorischen Neurons** (= Afferenz) zentripetal zum zentralen Nervensystem geleitet (Bear et al., 2009a; Burke et al., 1983; De Lahunta und Glass, 2009c; Pearson und Jordan, 2000). Das sensorische Neuron, eine pseudounipolare Nervenzelle, bildet das Spinalganglion der Dorsalwurzel (Radix dorsalis) des Rückenmarks (De Lahunta und Glass, 2009a). Über die Radix dorsalis des Rückenmarks treten die spinalen, segmentalen Afferenzen in das Rückenmark ein (Böhme et al., 2004d; De Lahunta und Glass, 2009a).
 3. Innerhalb des Rückenmarks ist das sensorische Neuron über jeweils eine **Synapse** mit mindestens einem Neuron verbunden (Klein und Cunningham, 2012). Der jeweilige Rückenmarksabschnitt, in dem diese Synapse liegt, wird als *Zielsegment* bezeichnet (Böhme et al., 2004d; Stoffel, 2011b). Bei einem motorischen Rückenmarksreflex ziehen afferente Axone ins Ventralhorn und aktivieren efferente Motoneurone. Ihre Axone verlassen das Rückenmark über die Radix ventralis des jeweiligen Rückenmarkssegments (Böhme et al., 2004d). Innerhalb des Eigenreflexbogens verläuft
-

das sensorische und motorische Neuron im selben Nerven des peripheren Nervensystems (Klein und Cunningham, 2012).

4. Das **α -Motoneuron** innerviert als motorische Efferenz die extrafusalen Muskelfasern des jeweiligen Muskels (Bear et al., 2009a; De Lahunta und Glass, 2009c).
5. Der durch das Motoneuron innervierte Muskel stellt das **Erfolgsorgan** des Reflexbogens dar und führt die Reflexantwort in Form einer kompensatorischen Kontraktion zur Aufrechterhaltung der Muskellänge (Myostatik) aus (Klein und Cunningham, 2012; Stoffel, 2011b).

Abbildung 1: Schematische Darstellung eines monosynaptischen Reflexbogens.



2.1.2. Motoneurone

2.1.2.1. Oberes Motoneuron

Das obere Motoneuron (OMN) ist dem unteren Motoneuron (UMN) hierarchisch übergeordnet (De Lahunta und Glass, 2009c). Es ist in zwei Anteile unterteilt, wobei sich eine komplett morphologisch und funktionell isolierte Betrachtung beider Anteile verbietet:

1. *Das pyramidale System*

Die Zellkörper des pyramidalen Systems liegen im motorischen Kortex (= Area motoria) des Großhirns (De Lahunta und Glass, 2009c; Garosi, 2013; Garosi und Lowrie, 2013; Stoffel, 2011b) unmittelbar kaudal des Sulcus cruciatus (Stoffel, 2011b). Ihre Axone ziehen als Pyramidenbahn in Form des Tractus corticospinalis und Tractus corticonuclearis durch die weiße Substanz des Großhirns und Hirnstamms (Stoffel, 2011b). Während der Tractus corticonuclearis die Verbindung zwischen Area motoria und den motorischen Kerngebieten diverser Hirnnerven darstellt, bildet der Tractus corticospinalis die Verbindung zwischen motorischem Kortex und den Motoneuronen des Rückenmarks (Stoffel, 2011b). Bei Mensch und Primat ist diese Verbindung monosynaptisch (von Engelhardt und Breves, 2005), bei den Haussäugetieren sind meist Interneurone in den Laminae VI und VII des Rückenmarks beziehungsweise weitere Kerngebiete, wie zum Beispiel der Nucleus ruber, zwischengeschaltet (Böhme et al., 2004b; Stoffel, 2011b). Alle Fasern der Pyramidenbahn kreuzen auf ihrem Weg die Mediane, wenngleich an unterschiedlichen Lokalisationen (De Lahunta und Glass, 2009c; Stoffel, 2011b). Somit liegt das motorische Kerngebiet jeweils kontralateral der innervierten Region. Auf Höhe der spinalen Intumneszenzen ist das OMN mit dem UMN der Gliedmaßen innerhalb des Dorsalhorns und der Substantia intermedia über eine Synapse miteinander verschaltet (Böhme et al., 2004e).

2. *Das extrapyramidale System*

Das extrapyramidale System ist ein multisynaptisches neuronales Konzept diverser miteinander verbundener und sich gegenseitig funktionell beeinflussender neuronaler Strukturen, die neuroanatomisch zwar außerhalb des pyramidalen Systems liegen, dieses jedoch funktionell

unterstützen (De Lahunta und Glass, 2009c; Stoffel, 2011b). Da jedoch ausschließlich die extrapyramidalen Kerngebiete in Mes- und Rhombencephalon über spinale Axone mit dem UMN der Gliedmaßen verbunden sind (De Lahunta und Glass, 2009c), werden nur diese hier Erwähnung finden. Der im Mesencephalon liegende Nucleus ruber bildet den Ursprung für den innerhalb des Funiculus lateralis verlaufenden deszendierenden Tractus rubrospinalis, der über Interneurone oder direkt mit dem UMN, den α - und γ -Motoneuronen, verbunden ist (Böhme et al., 2004b; De Lahunta und Glass, 2009c; Hongo et al., 1972). Der Nucleus ruber vermittelt im Zielsegment hauptsächlich inhibitorische Impulse auf die Streck- und exzitatorische Impulse auf die Beugemuskulatur (De Lahunta und Glass, 2009c; Hongo et al., 1972; Hongo et al., 1969). Innerhalb des Rhombencephalon entspringen aus der Formatio reticularis der Tractus reticulospinalis ventralis et lateralis (Bear et al., 2009b; Böhme et al., 2004b). Während der Tractus reticulospinalis ventralis innerhalb des Funiculus ventralis im Rückenmark verläuft und exzitatorisch auf die Motoneuronen der Extensorenmuskulatur wirkt, verläuft der Tractus reticulospinalis lateralis innerhalb des spinalen Funiculus lateralis und wirkt inhibitorisch auf die Motoneuronen der Extensorenmuskulatur (De Lahunta und Glass, 2009c).

Die willkürliche Motorik beziehungsweise Zielmotorik steht unter dem direkten Einfluss des motorischen Kortex und wird somit hauptsächlich durch das pyramidale System kontrolliert (Stoffel, 2011b). Die Entwicklung des pyramidalen Systems beziehungsweise die Anzahl der Synapsen, über die das pyramidale System mit dem UMN verbunden ist, korreliert direkt mit der Fähigkeit, bestimmte Bewegungsmuster zu erlernen (De Lahunta und Glass, 2009c). Da das pyramidale System bei den Haussäugetieren gegenüber dem Menschen eher gering ausgebildet ist und funktionell eine vergleichsweise untergeordnete Rolle spielt, ist das Erlernen neuer Bewegungsabläufe bei unseren Haussäugetieren auch nur begrenzt möglich (Stoffel, 2011b). Experimentell durchgeführte Resektionen der kortikalen Area motoria oder eine Durchtrennung der Pyramidenbahn führten, wenn überhaupt, nur zu einer geringen Beeinträchtigung der Motorik beim Tier (Böhme et al., 2004b). Bei den Haussäugetieren überwiegt die unwillkürliche Motorik, die zwar bewusst erlebt wird, jedoch aus erlernten, automatisierten Bewegungsmustern besteht. Die unwillkürliche Motorik wird beinahe ausschließlich durch das extrapyramidale System und nur zu einem geringen Anteil durch den motorischen Kortex gesteuert (Stoffel, 2011b).

2.1.2.2. Unteres Motoneuron

Die UMN-Axone verlassen das Rückenmark über den Ramus ventralis nervi spinalis. Das UMN bildet die direkte Verbindung zwischen OMN und Erfolgsorgan. Man unterscheidet zwei Typen spinaler Motoneurone (Bear et al., 2009a):

1. α -Motoneurone

α -Motoneurone bilden die direkte Verbindung zwischen OMN und der innervierten extrafusalen Muskelfaser (s. auch Abbildung 1) und stellen somit einen wesentlichen Bestandteil der motorischen Einheit dar (Liddell und Sherrington, 1925). Die Gesamtheit aller α -Motoneuronen, die einen einzelnen Muskel innervieren, bezeichnet man als Motoneuronenpool (Bear et al., 2009b). Die Entladungsrate der α -Motoneurone wird innerhalb des Ventralhorns durch die präsynaptische Aktivität sensorischer Afferenzen (1 α -Fasern der Muskelspindeln), supraspinaler, deszendierender Bahnen des zentralen Nervensystems (OMN) und Interneuronen des Rückenmarks moduliert (Bear et al., 2009b). Letztere können inhibitorische oder exzitatorische Stimuli an Motoneurone vermitteln und so kommt es beispielsweise im Rahmen des Dehnungsreflexes durch reziproke Hemmung sowohl zu einer Kontraktion einer bestimmten Muskelgruppe, jedoch auch zu einer gleichzeitigen Relaxation seiner Antagonisten (Bear et al., 2009b).

2. γ -Motoneurone

γ -Motoneurone bilden die motorischen Efferenzen zu den intrafusalen Muskelfasern. Sie stellen einen wesentlichen Anteil der sogenannten γ -Spindelschleife: γ -Motoneurone innervieren die Polregionen der Muskelspindeln und gewährleisten so die Aufrechterhaltung der propriozeptiven Funktion der Muskelspindeln (Bear et al., 2009b). Durch die induzierte Kontraktion der aktivierten Muskelspindel kehrt diese wieder in den Ausgangszustand zurück. Der Ausgangszustand wird anschließend wieder über eine afferente 1a-Faser an das Zielsegment geleitet. Die supraspinalen Motoneurone (OMN) sind innerhalb des Ventralhorns über Synapsen ebenfalls mit den γ -Motoneuronen verbunden und wirken inhibitorisch auf deren Aktivität (De Lahunta und Glass, 2009c).

2.2. Spezielle topographische Anatomie

2.2.1. Anatomie des Bizepssehnenreflexes

Der Plexus brachialis ist ein Nervengeflecht des sechsten, siebten und achten zervikalen (Cnv₆ - Cnv₈) sowie des ersten und zweiten thorakalen Nervus spinalis ventralis (Tnv₁ - Tnv₂) (Böhme et al., 2004a). Allam et al. (1952) fanden in ihrer Sektionsstudie bei 58 Armgeflechten, dass es vereinzelt auch zu einer Beteiligung des Cnv₅ kommt.

Der Nervus musculocutaneus entspringt dem Plexus brachialis (Allam et al., 1952; Böhme et al., 2004a). Über die spinalen Ursprungssegmente des Nervus musculocutaneus herrscht in der veterinäranatomischen Literatur Uneinigkeit. Allam et al. (1952) beschrieben den Nervus musculocutaneus als Fortsetzung von Cnv₇. Böhme et al. (2004a) postulierten ergänzend, dass der Nervus musculocutaneus beim Hund meist auch Fasern von Cnv₈ führt. Sharp et al. (1993) wiesen mittels Elektrostimulation nach, dass er Fasern der Segmentalnerven Cnv₆ – Tnv₁ führt, wobei der größte Anteil von den Segmentalnerven Cnv₆ und Cnv₇ stammt. Proximal des Schultergelenkes verläuft er schräg über die Medialseite des Musculus subscapularis und medial der Ansatzstelle des Musculus teres major (Allam et al., 1952). Der Ramus muscularis proximalis zweigt auf Höhe des Schultergelenks nach lateral vom Nervus musculocutaneus ab und innerviert den Musculus biceps brachii und Musculus coracobrachialis (Böhme et al., 2004a).

Der spindelförmige Musculus biceps brachii bildet das Erfolgsorgan des Bizepssehnenreflexes. Proximal setzt seine Ursprungssehne am Tuberculum supraglenoidale scapulae an. Sein Muskelbauch liegt dem Humerus proximal medial, nach distal zunehmend kranial an. In der Ellbogenbeuge teilt er sich in zwei Schenkel auf, von denen der kräftigere breitbasig am proximalen Rand des Processus coronoideus medialis ulnae, der schwächere an der Tuberositas radii ansetzt (Nickel et al., 2004b; Villamonte-Chevalier et al., 2012). Der Musculus biceps brachii fungiert somit als Flexor des Ellbogen- und Strecker des Schultergelenks mit geringer Hubhöhe, aber hoher Hubkraft (Nickel et al., 2004b).

Die Reflexantwort des Bizepssehnenreflexes ist eine Kontraktur des Musculus biceps brachii mit oder ohne Beugung des Ellbogens (Schatzberg et al., 2012).

2.2.2. Anatomie des Patellarsehnenreflexes

Beim Hund bilden die Ventraläste der 3. - 6. Lumbalnerven (Lnv₃₋₆) das Lendengeflecht (Plexus lumbalis) (Budras et al., 2004; Monticelli et al., 2016). Aus dem Plexus lumbalis geht der Nervus femoralis hervor, der seine Fasern hauptsächlich aus den Lnv₄₋₆ bezieht (Böhme et al., 2004c; Budras et al., 2004; Wilson, 1978), er kann aber auch Fasern von Lnv₃ (Miller et al., 1979) und Lnv₇ (Mihelic et al., 2007) enthalten. Der Nervus femoralis verläuft zunächst ventral der Wirbelsäule innerhalb des Musculus iliopsoas und gibt kurz vor Austritt aus der Bauchhöhle den Nervus saphenus ab (Miller et al., 1979). Nach Abgang des Nervus saphenus verlässt der Nervus femoralis kraniallateral der Arteria iliaca externa die Bauchhöhle durch die Lacuna musculorum und tritt zwischen dem Musculus vastus medialis und Musculus rectus femoris in die Quadrizepsgruppe ein, die er innerviert (Böhme et al., 2004c; Mahler und Adogwa, 2008; Miller et al., 1979).

Der Musculus quadriceps femoris ist das Erfolgsorgan des Patellarsehnenreflexes und lässt sich in vier Muskelbäuche unterteilen, die distomedial konvergieren und gemeinsam das gerade Kniescheibenband (Ligamentum patellae) bilden (Nickel et al., 2004a; Salomon und Geyer, 2007):

1. Der *Musculus rectus femoris* setzt mit seiner Ursprungssehne oberhalb des Pfannenrandes am Os ilium an. Seine Endsehne bildet den Hauptteil des Ligamentum patellae.
 2. Der *Musculus vastus medialis* hat seinen Ursprung an der kranio-medialen Fläche des proximalen Femurschafts und strahlt direkt kranial in die Endsehne des Musculus rectus femoris ein.
 3. Der *Musculus vastus lateralis* entspringt kraniallateral am proximalen Femurschaft und zieht in den Sehnen Spiegel des Musculus vastus medialis.
 4. Der *Musculus vastus intermedius* ist der schwächste der vier Muskelbäuche und nimmt seinen Ursprung gemeinsam mit dem Musculus vastus lateralis. Sein Muskelbauch liegt dem Femur direkt an und strahlt distal in den Musculus vastus medialis ein.
-

Der Musculus quadriceps femoris fungiert als Strecker des Kniegelenks und Beuger des Hüftgelenks (Salomon und Geyer, 2007). Die physiologische Reflexantwort des Patellarsehnenreflexes ist eine Kontraktion des Musculus quadriceps femoris, die zu einer sprunghaften Streckung des Kniegelenks führt (De Lahunta und Glass, 2009a).

2.3. Die Reflexevaluierung als medizinischer Parameter

2.3.1. Definitionen

2.3.1.1. Reflexantwort

Der Begriff *Reflexantwort* wird vom Kliniker meist mit der in der Literatur zu findenden Definition für den Überbegriff *Reflex* verwendet: ein unwillkürlicher und regelhaft ablaufender Vorgang als physiologische Reaktion eines Erfolgsorgans auf einen adäquaten Reiz (Pschyrembel und Dornblüth, 2002). Sie stellt den messbaren Parameter der neurologisch-klinischen Reflexprüfung dar.

2.3.1.2. Reflexpräsenz

Die Reflexpräsenz wird in der vorliegenden Arbeit als Parameter der Detektion einer Reflexantwort definiert (vgl. Levine et al., 2002). Die Kategorisierung ist dichotom (Reflexantwort detektierbar/ vorhanden; Reflexantwort nicht detektierbar/ nicht vorhanden).

2.3.1.3. Reflexaktivität

Die Reflexaktivität bezeichnet das Ausmaß beziehungsweise die Stärke der Reflexantwort (engl.: *reflex briskness*) (Dafkin et al., 2013; Stam und van Crevel, 1990). Sie kann quantitativ durch Elektromyographie (Dietz et al., 1991; Stam und Vanleeuwen, 1984; Stam und van Crevel, 1990) oder, wie im Rahmen der neurologischen Untersuchung in der Veterinärmedizin üblich, semiquantitativ mit Hilfe einer Ordinalskala (De Lahunta und Glass, 2009b; Schatzberg et al., 2012; Sharp und Wheeler, 2005; Thomas und Dewey, 2008) beurteilt werden.

2.3.2. Die Entdeckung der Reflexantwort

Mit der Entdeckung der Reflexantwort war die Grundlage für die Evaluierung der Reflexe im klinischen Alltag gelegt. Ihre historische Betrachtung veranschaulicht die Problematik und den Fortschritt ihrer zuverlässigen Messung.

In der griechischen Antike unterschied man zunächst zwischen willkürlichen und unwillkürlichen Bewegungen (Clarke und O'Malley, 1996). Unwillkürliche Bewegungen waren beispielsweise der Herzschlag, das Blinzeln der Augen und der rhythmische Bewegungsablauf des Ganges. Galen vermutete, dass die Seele einen Einfluss auf die unwillkürlichen Bewegungen ausüben müsse (Boes, 2014).

Der französische Naturwissenschaftler und Philosoph René Descartes (1596-1650) war der erste, der eine primitive Konzeption eines Reflexes Mitte des 17. Jahrhunderts publizierte (Boes, 2014; Clarke und O'Malley, 1996; Dyck et al., 2005). Im Jahre 1649 verwendete Descartes den Begriff „*refleschis*“ zur Beschreibung einer Abwehrreaktion (Clarke und O'Malley, 1996). Der britische Neurophysiologe Charles Scott Sherrington (1857-1952) lieferte durch seine bahnbrechenden Entdeckungen das Fundament zu dem aktuellen Modell des Reflexbogens (Clarke und O'Malley, 1996). Die Neurologen Wilhelm Erb (Erb, 1875), Carl Westphal (Westphal, 1875) und Jean-Martin Charcot (Goetz et al., 1995) sorgten mit ihren Publikationen schließlich dafür, dass die Evaluierung der Reflexantwort fester Bestandteil der neurologischen Untersuchung wurde (Clarke und O'Malley, 1996; Haines et al., 2007).

2.3.3. Beeinflussung der Reflexantwort

2.3.3.1. Einfluss der neuroanatomischen Lokalisation der Läsion

Die Feststellung der neuroanatomischen Lokalisation mittels Reflexevaluierung ist ein zentraler Punkt im Rahmen der neurologischen Untersuchung. Die anatomischen Grundlagen zum UMN und OMN wurden in Kapitel 2.1.2. Motoneurone erläutert.

2.3.3.1.1. Läsion des oberen Motoneurons

Bei einer Läsion des OMN sind die Reflexbögen kranial und kaudal der geschädigten Rückenmarksegmente intakt. Über Synapsen innerhalb des Ventralhorns beeinflusst das OMN mittels inhibitorischer Stimuli die Aktivität des UMN (Hongo et al., 1972; Hongo et al., 1969; Nyberg-Hansen und Brodal, 1963) und somit auch die Qualität der Reflexantwort (Cozens et al., 2000; Pearson und Jordan, 2000; Simons und Dimitrijevic, 1972). Indem das OMN die Aktivität der γ -Motoneurone zu den Polregionen der Muskelspindeln inhibiert (Bear et al., 2009b; De Lahunta und Glass, 2009c), nimmt es direkten Einfluss auf die Reflexantwort myostatischer Reflexe (De Lahunta und Glass, 2009c; Garosi und Lowrie, 2013). Durch die γ -Rückkopplungsschleife reguliert das jeweilige Rückenmarkssegment den Muskeltonus der innervierten Gliedmaße (Bear et al., 2009b). Eine Dysinhibition der γ -Rückkopplungsschleife resultiert in einem gesteigerten Tonus der Extensorenmuskulatur sowie Hyperreflexie (De Lahunta und Glass, 2009c; Garosi und Lowrie, 2013).

Die Symptome einer OMN-Läsion resultieren aus einer reduzierten reziproken Inhibition durch Muskelspindeln der Antagonisten und einer reduzierten präsynaptischen Inhibition der Muskelspindeln der Extensoren (Kheder und Nair, 2012). Weitere Symptome einer OMN-Läsion sind pathologische Reflexe wie der gekreuzte Streckreflex am liegenden Tier sowie Massenreflexe (De Lahunta und Glass, 2009c).

2.3.3.1.2. Läsion des unteren Motoneurons

Bei einer Läsion des UMN ist der Reflexbogen des untersuchten Reflexes gestört. Die Läsion kann somit jede der fünf Komponenten (s. 2.1.1. Der Reflexbogen) betreffen und resultiert zwangsläufig in einer Dysfunktion der motorischen Komponente des Reflexbogens (De Lahunta und Glass, 2009a). Demzufolge ist die Reflexantwort im Falle einer UMN-Läsion hypo- bis areflexiv (De Lahunta und Glass, 2009a; Garosi und Lowrie, 2013). Durch die direkte Störung der γ -Rückkopplungsschleife und dem damit einhergehenden Funktionsverlust der Muskeltonusmodulation sowie der Denervation des Muskels kommt es zu einer Muskelhypotonie (Bear et al., 2009a; De Lahunta und Glass, 2009a).

2.3.3.2. Subjektive Einflussfaktoren

Es wurde eine Vielzahl von Einflussfaktoren durch das Untersuchungsobjekt auf die Reflexaktivität beschrieben. Die meisten Studien beziehen sich auf den Menschen. Für einige Faktoren existieren jedoch auch Untersuchungen aus der Veterinär- und Humanmedizin. Es gilt des Weiteren zu berücksichtigen, dass die meisten diskutierten Studien die Reflexaktivität mittels Elektrodiagnostik quantifizierten. Vor dem Hintergrund der Vielzahl an subjektiven Einflussfaktoren empfahlen Stam und van Crevel (1990) eine Interpretation der Reflexaktivität im klinischen Alltag nur im Seitenvergleich und in Verbindung mit der übrigen Symptomatik.

Sowohl in der Human- wie auch in der Veterinärmedizin ist bekannt, dass das Alter des Patienten einen wesentlichen Einfluss auf die Reflexaktivität nimmt. Aminoff (2012) beschrieb eine verlängerte Dauer der Aktionspotentiale motorischer Einheiten (*Motor Unit Action Potentials* = *MUAP*) mit zunehmendem Alter. Frijns et al. (1996) wiesen eine positive Korrelation zwischen Alter und Latenzzeit nach. Levine et al. (2002) fanden heraus, dass der Patellarsehnenreflex beim älteren Hund nicht immer ausgelöst beziehungsweise beurteilt werden kann und Chandrasekhar et al. (2013) registrierten, dass die Kniegelenkswinkelung bei der Reflexantwort des Patellarsehnenreflexes mit zunehmendem Alter beim Menschen abnimmt.

Die altersabhängige Veränderung der Reflexaktivität wird mit der durch Wachstum einhergehenden großflächigeren Verteilung der neuromuskulären Endplatten (Aminoff, 2012), der generellen altersbedingten Muskelatrophie (*Sarkopenie*) (Baumgartner et al., 1995; Deschenes, 2004) und den altersabhängigen histologischen Veränderungen der Muskulatur wie veränderte Muskelfaseranteile (Balagopal et al., 1997; Deschenes, 2004), dem Verlust von α -Motoneuronen (Deschenes, 2004; Larsson, 1978) oder der verminderten Syntheserate bestimmter Muskelfaserproteine (Balagopal et al., 1997) erklärt. Griffiths und Duncan (1975) wiesen darüber hinaus altersabhängige Veränderungen an den lumbalen Spinalnerven beim Hund wie zum Beispiel Ballonierung der Myelinscheiden, Demyelinisierung oder verlängerte Internodien nach. Ziff und Stark (2017) relativierten die Hypothese, dass altersbedingte Veränderungen zu einer falsch-negativen Reflexantwort führen. Vielmehr konnten sie beweisen, dass es beim älteren Patienten wesentlich darauf ankommt verschiedene Reflexevaluierungsmethoden anzuwenden, da die standardmäßigen Untersuchungsmethoden bei älteren Patienten aufgrund inadäquater Lagerung oder Paratonie („*frontal rigidity*“) zu falsch-negativen Ergebnissen führen können.

Die intramuskuläre Temperatur ist ein weiterer Faktor, der die Reflexaktivität beeinflusst (Aminoff, 2012; Bolton et al., 1981; Buchthal et al., 1954; Denys, 1991; Rutkove, 2001). Denys (1991) schrieb ihr den größten Einfluss auf artifizielle Veränderungen in der Elektrodiagnostik zu. Ursächlich führte er eine mit der Temperaturabnahme einhergehende verlangsamte Nervenleitgeschwindigkeit oder eine verlängerte Kontraktions- und Relaxationszeit der Muskelfasern an. Eine Abnahme der Nervenleitgeschwindigkeit in Zusammenhang mit einer Temperaturabnahme konnte auch bei der Katze nachgewiesen werden (Douglas und Malcolm, 1955). Rutkove (2001) diskutierte, dass es infolge einer Temperaturabnahme zu einer reduzierten Aktivität der $\text{Na}^+\text{-K}^+\text{-ATPase}$ und somit zu einer Erhöhung des Membranruhepotentials kommt. Hubbard et al. (1971) stellten die Hypothese auf, dass eine Temperaturabnahme zu einer reduzierten Acetylcholin-Ausschüttung an der neuromuskulären Endplatte führt und es somit zu einer schwächeren (Denys, 1991) beziehungsweise langsameren (Torbegsen und Stalberg, 1986) Muskelkontraktion kommt. Ein Temperaturanstieg führt hingegen zu einer Verkürzung der Latenzzeit und einer Abnahme der Amplitude (Bolton et al., 1981). Es gilt zu berücksichtigen, dass Stress oder Anspannung hierbei sowohl zu einer Temperaturabnahme der Gliedmaßen (Denys, 1991) jedoch auch durch Muskelanspannung zu einer Temperaturerhöhung führen kann.

Der Blutfluss begünstigt den Abtransport der durch die Muskelkontraktion anfallenden Metaboliten und beschleunigt die Wiederherstellung des Grundzustandes der Muskulatur (De Luca, 1997). Emery (1929) fand bereits zu Beginn des letzten Jahrhunderts heraus, dass der Muskeltonus und somit das Ausmaß des Patellarsehnenreflexes, verglichen mit der Diastole, während des systolischen Herzschlagphase größer ist.

Die Menge an subkutanem Fett kann ebenfalls zu artifiziellen Veränderungen bei der elektromyographischen Ableitung führen (Dafkin et al., 2012).

Aus der Humanmedizin ist bekannt, dass die Reflexantwort zu einem gewissen Grad durch mentale Aktivität und durch sensorische Stimuli (Davis und Beaton, 1968) beeinflusst werden kann (Bathien, 1971; Emery, 1931; Stam und Tan, 1987; Tuttle, 1924). So können die spinalen Motoneurone durch deszendierende Impulse der retikulospinalen Bahnen beeinflusst werden. Diese Impulse führen so zu einer verkürzten Latenzzeit der Reflexantwort (Péron et al., 2004). Der Beurteiler könnte so die Reflexaktivität fälschlicherweise als gesteigert interpretieren (Dick, 2003). Stress und Anspannung im Rahmen der Reflexevaluierung führen zudem zu einem gesteigerten Muskeltonus (Dafkin et al., 2012; Delwaide und Toulouse, 1981; Péron et

al., 2004) und sollten als wichtige Quellen artifizieller Veränderungen im Rahmen der neurologischen Untersuchung in der Veterinärmedizin berücksichtigt werden.

Die meisten der genannten patientenbezogenen Faktoren sind nur bis zu einem gewissen Grad am wachen Patienten zu beeinflussen, sollten aber bei der Interpretation der Ergebnisse beziehungsweise bei der Versuchsdurchführung Berücksichtigung finden (De Luca, 1997; Frijns et al., 1997).

2.3.4. Methoden der Standardisierung der Reflexevaluierung

Durch verschiedene Standardisierungsmethoden wird versucht, das Interobserver Agreement der Reflexevaluierung zu optimieren. Drei Faktoren beeinflussen die Übereinstimmung zwischen zwei Untersuchern: der Beurteiler, der Proband und die Untersuchungsdurchführung (Dept. of Clinical Epidemiology and Biostatistics, McMaster University, 1980). Zwangsläufig müssen diese Faktoren für eine vergleichbare Wertigkeit des Untersuchungsergebnisses standardisiert werden (Lehmann und McGill, 1999). Inzwischen wurden verschiedene Methoden etabliert, die im Rahmen der Reflexevaluierung eine Beurteilung objektivieren beziehungsweise standardisieren sollen. Die subjektiven Einflussfaktoren, die durch den Patienten verursacht werden, wurden bereits beschrieben (s. 2.3.3.2. Subjektive Einflussfaktoren).

2.3.4.1. Standardisierung der neurologischen Untersuchung

Zwar bildet eine mehr oder weniger detaillierte Beschreibung der neurologischen Untersuchung die Basis eines jeden neurologischen Werkes beziehungsweise Kapitels (De Lahunta und Glass, 2009b; Garosi und Lowrie, 2013; Schatzberg et al., 2012; Thomas und Dewey, 2008), dennoch wird offensichtlich, dass die neurologische Untersuchung sich zumindest klinikspezifisch individuell unterscheidet. Dabei ist gerade eine standardisierte Durchführung Grundlage für vergleichbare und zuverlässige Messergebnisse (Vreeling et al., 1993). Im Folgenden werden verschiedene Faktoren der neurologischen Untersuchung diskutiert, deren Standardisierung Gegenstand der Forschung ist.

Es ist bekannt, dass die Reflexantwort von der Impulsstärke (Andersen et al., 2001; Davis und Beale, 1970; De Lahunta und Glass, 2009b; Miller et al., 1995a; Pagliaro und Zamparo, 1999; Spadavecchia et al., 2003), –lokalisation (Andersen et al., 1999; Andersen et al., 2001; Clarke und Harris, 2004; Dafkin et al., 2012; O'Keeffe et al., 1994) und –dauer (Sawaya et al., 2008) beeinflusst wird. Aus der Notwendigkeit heraus ein Instrument zu entwickeln, mit dem sowohl oberflächliche als auch tiefe Reflexe provoziert werden können, dessen Schläge angenehm für den Patienten sind, das eine konstante Schlagkraft garantiert und sich für die neurologische Untersuchung am Neugeborenen wie auch am älteren Patienten eignet, wurden eine Vielzahl von Reflexhammermodellen entwickelt (Lanska, 1989; Lanska und Lanska, 1990; Lanska, 1995a; Lanska, 1995b; Lanska, 1999). Für die ersten neurologischen Untersuchungen wurde der von Leopold Auenbrugger entwickelte Perkussionshammer verwendet. Alternativ wurden auch Schläge mit der Handkante oder den Fingern empfohlen (Lanska, 1989). John Madison Taylor präsentierte 1888 den von ihm entwickelten ersten sogenannten Reflexhammer (Lanska, 1989; Lanska und Lanska, 1990), der auch heute noch weit verbreitet ist. Aus dem historischen Essay von Lanska und Lanska (1990) geht hervor, dass es keineswegs den universellen Reflexhammer gab und gibt, jedoch zur standardisierten Untersuchung jeweils ein einheitliches Modell in den jeweiligen Kliniken verwendet wurde. De Lahunta und Glass (2009) empfahlen für kleine Hunde und Katzen einen kleinen und für größere Hunde einen größeren Reflexhammer zu verwenden. Auch in der Reliabilitätsanalyse des Patellarsehnenreflexes beim Hund von Litvan et al. (1996) wurde darauf geachtet, dass die Untersucher jeweils dasselbe Reflexhammermodell nutzten.

In einigen experimentellen Studien wurden zur Standardisierung der Impulsstärke und -lokalisation verschiedene Apparaturen, wie zum Beispiel das „*Patellometer*“ (Pereira, 1925) eingeführt, das quantifizierbare und konstante Impulse setzen sollte (Cozens et al., 2000; Dafkin et al., 2013; Davis und Beale, 1970; Pereira, 1925). In anderen Arbeiten wurde durch die Verwendung von Elektrodiagnostik eine kontrollierte Impulsstärke und –lokalisation gewährleistet (Clarke, 1965; Dafkin et al., 2013; Pagliaro und Zamparo, 1999).

Die Reflexaktivität des Menschen verändert sich mit Positionsveränderungen des Körpers (Nuzzo et al., 2016; Peterson et al., 2014). In der Humanmedizin fand man heraus, dass es zu positionsabhängigen Veränderungen der reziproken Inhibition durch 1α -Fasern auf die Motoneurone kommt und sich somit die Reflexaktivität in Abhängigkeit der Körperposition verändert (Barbeau et al., 2000; Perez und Field-Fote, 2003; Rossi et al., 1988). In der

veterinärmedizinischen Literatur wird eine Untersuchung in Seitenlage des Hundes empfohlen, wobei sich die zu untersuchende Gliedmaße bodenfern befindet (Garosi und Lowrie, 2013; Schatzberg, 2010; Schatzberg et al., 2012). De Lahunta und Glass (2009) beschrieben das Phänomen, dass sich der Patellarsehnenreflex beim Hund bisweilen lediglich im bodennahen Bein auslösen ließe und rieten somit dazu eine Reflexevaluierung für eine Gliedmaße jeweils in rechter und linker Seitenlage durchzuführen.

Ausreichende Übung in der Durchführung der Untersuchung ist notwendig, um im Rahmen medizinischer Untersuchungen zuverlässige Ergebnisse erzielen zu können (Kottner et al., 2009b; Péréon et al., 2004; Raftery und Holland, 1967; Raijmakers et al., 1991; Schmid et al., 2009). O’Keeffe et al. (1994) vermuteten, dass eine unterschiedliche Technik der Reflexevaluierung zwischen den Untersuchern zu Evaluierungsdiskrepanzen führe. Andere Studien belegen, dass die Reflexevaluierung auch bei Neurologen nicht standardisiert erfolgt, jedoch dennoch zuverlässige Ergebnisse erzielt werden können (Giebels et al., 2014; Litvan et al., 1996; Vreeling et al., 1993). Gleichwohl sollte die Untersuchungsdurchführung einem Protokoll folgen um Ergebnisse verschiedener Studien vergleichbar zu machen (Koran, 1975; Vreeling et al., 1993).

2.3.4.2. Standardisierung der Beurteilung der Reflexantwort

Je mehr Beurteiler eine Untersuchung evaluieren, desto geringer ist die Wahrscheinlichkeit einer vollständig übereinstimmenden Beurteilung (Blendis et al., 1970; Koran, 1975; Raftery und Holland, 1967). Ferner ist zu berücksichtigen, dass es in Studien, in denen der Untersucher gleichzeitig Beurteiler ist, zu einer hohen Interobserver Variabilität kommt (Hansen et al., 1994; Manschot et al., 1998; Stam und van Crevel, 1990; Vogel, 1992).

Die Kenntnis über die Anamnese des Patienten senkt die Anzahl falsch-positiver Ergebnisse (Doubilet und Herman, 1981), führt zu einer solideren Diagnosestellung (Aideyan et al., 1995), verbessert die Interpretation der Untersuchungsergebnisse (Vogel, 1992) und erhöht das Interobserver Agreement zwischen den Untersuchern (Litvan et al., 1996; Vogel, 1992). Vogel et al. (1992) fanden allerdings keinen Effekt der Anamnese auf die dichotome Evaluierung der Reflexpräsenz. Dabei ist die Kenntnis über die Anamnese nicht zwangsläufig vorteilhaft für das Untersuchungsergebnis: sie kann sich positiv auf die Vigilanz des Untersuchers auswirken

(Fowkes, 1986), jedoch auch durch „Priming“ die Interpretation der Untersuchungsergebnisse verfälschen (Vogel, 1992).

Silas Weir Mitchell veröffentlichte im Jahr 1886 gemeinsam mit seinem Kollegen Morris J. Lewis erstmals eine Ordinalskala zur Beurteilung der Reflexaktivität (Mitchell und Lewis, 1886). Diese 6-Punkte-Skala umfasste bereits Abstufungen zur Beurteilung einer sowohl reduzierten als auch gesteigerten Reflexantwort. Im Laufe der Jahre wurden verschiedene Modelle entwickelt, wobei die Anzahl der Kategorien zwischen vier und acht lag (Dyck et al., 2005). Wichtige Kriterien einer Beurteilungsskala sind nach heutigen Gesichtspunkten *nicht zu viele Kategorien*, eine *eindeutige Benennung der Kategorien* und die Möglichkeit der *Unterscheidung zwischen „normal“ und „abnormal“* (Manshot et al., 1998).

In der Veterinärmedizin wird gegenwärtig meist eine 5-Punkte-Skala zur Evaluierung der Reflexaktivität empfohlen, wobei folgende Definitionen gelten: 0 = nicht vorhanden, 1 = reduziert, 2 = normal, 3 = gesteigert, 4 = tonisch (De Lahunta und Glass, 2009b; Schatzberg et al., 2012; Sharp und Wheeler, 2005; Thomas und Dewey, 2008).

Eine Evaluierung der Reflexantwort mittels Videoanalyse wurde in modernen Studien bereits durchgeführt (Borin et al., 2012; Dafkin et al., 2013; Dafkin et al., 2016; Essex et al., 2003; Van Gijn und Bonke, 1977) und soll eine standardisierte Evaluierung durch den Beurteiler gewährleisten (Dafkin et al., 2013; Dafkin et al., 2016). Dafkin et al. (2013) zeigten, dass die subjektive Videoevaluierung der Reflexaktivität mit objektiven Messdaten korreliert. Es gilt allerdings auch hier zu berücksichtigen, dass eine Reflexevaluierung am Monitor nicht mit einer direkten klinischen Untersuchung gleichzusetzen ist und es durch die Standardisierung eventuell sogar zu einem artifiziellen, falsch-hohen Interobserver Agreement kommen kann (Dafkin et al., 2016).

Verschiedene Studien belegen, dass die Erfahrung in der Beurteilung von klinischen Parametern das Interobserver Agreement erhöht (Kottner et al., 2009b; Meade et al., 1968; Raftery und Holland, 1967). Somit berücksichtigten Studien, die das Interobserver Agreement verschiedener Phänomene untersuchten (s. auch Tabelle 1), das Erfahrungsniveau der Beurteiler (Dafkin et al., 2013; Kottner et al., 2009b; Meade et al., 1968; Raftery und Holland, 1967; Vreeling et al., 1993). Dafkin et al. (2016) postulierten, dass die Expertise der Beurteiler nicht zwangsläufig in einem höheren Interobserver Agreement, sondern in einer substantielleren Interpretation der Reflexantwort resultiert.

Die größte Standardisierung der Beurteilung der Reflexantwort wird durch ihre Quantifizierung mittels Elektrodiagnostik erzielt. Hierbei werden vor allem die Parameter *Reflexamplitude* und *Latenzzeit* zur Beurteilung herangezogen (Frijns et al., 1997; Voerman et al., 2005). Aufgrund der schwierigen Anwendung am wachen Patienten findet die Elektrodiagnostik zur Reflexevaluierung im Rahmen der neurologischen Untersuchung in der Veterinärmedizin keine Routineanwendung. Die Elektrodiagnostik wird in dieser Arbeit nur marginal erwähnt.

Tabelle 1 bietet eine Übersicht ausgewählter Studien, die die Reliabilität verschiedener neurologischer Parameter untersuchten.

Tabelle 1. Übersicht ausgewählter Studien, die das Ziel der Quantifizierung des Interobserver Agreements bei der Beurteilung neurologischer Parameter hatten. Kein Anspruch auf Vollständigkeit. K_C = Cohen's Kappa; K_w = gewichtetes Kappa; ICC = Intraklassen-Korrelationskoeffizient; r% = Prozentuale Übereinstimmung

Autor	Anzahl Untersucher	Anzahl Kategorien	Beurteiler = Untersucher?	beurteilte Parameter	Interobserver Agreement			
					K_C	K_w	ICC	r%
Tomasello et al. (1982)	2	2	ja	diverse				21,40 - 63,60
Born et al. (1987)	6	2-5	ja	diverse		0,55 - 0,94		
Brott et al. (1989)	4	2	ja	diverse	0,42 - 1,00			
Stam und van Crevel (1990)	3	9	ja	diverse			0,70 - 0,89	26,00
Maher et al. (1992)	4	2	ja	Babinski-Reflex				50,00
Vogel (1992)	2	5	ja	diverse		0,41 - 0,84		35,00 - 79,00
	2	2	ja		0,14 - 0,96			85,00 - 93,00
Vreeling et al. (1993)	2	3	ja	diverse		0,44 - 1,00		

Autor	Anzahl Untersucher	Anzahl Kategorien	Beurteiler = Untersucher?	beurteilte Parameter	Interobserver Agreement			
					K _C	K _w	ICC	r%
Hansen et al. (1994)	2	2	ja	diverse	0,19 - 0,64			
Litvan et al. (1996)	4	5	ja	diverse		0,43 - 0,80		
Manschot et al. (1998)	2-3	5-9	ja	diverse	-0,06 - 0,60			
Vroomen et al. (2000)	2	2	ja	Achillessehnenreflex Patellarsehnenreflex	0,42 - 0,53			86,00 - 100,00
Miller und Johnston (2005)	5	2	ja	Babinski-Reflex	0,28 - 0,73			
Singerman und Lee (2008)	2-6	2	ja	Babinski-Reflex	0,35 - 0,55			
Cook et al. (2009)	2	3	ja	diverse	0,55 - 0,76			78,00 - 89,00
Muguet-Chanoit et al. (2010)	2	2	ja	Cutaneus Trunci Reflex	0,67			
Lee et al. (2011)	2	3	ja	Babinski-Reflex	0,43			65,60
	2	3	nein	Babinski-Reflex	0,62			79,40
Annaswamy et al. (2012)	2	2	ja	Hoffmann-Zeichen	0,65 - 0,72			

Autor	Anzahl Untersucher	Anzahl Kategorien	Beurteiler = Untersucher?	beurteilte Parameter	Interobserver Agreement			
					K _C	K _w	ICC	r%
Borin et al. (2012)	2	6	ja	Blinkreflex	0,79			85,90
Carswell et al. (2012)	2	2	1 Untersucher 1 Videobeurteiler	diverse	0,18			
Isaza Jaramillo et al. (2014)	2	2	ja	Babinski-Reflex	0,73			
Thaller und Hughes (2014)	2	4	ja	diverse	0,57			
Dafkin et al. (2016)	12	2	nein	Babinski-Reflex	0,67 - 0,72			

2.3.5. Einsatzbereich der Reflexevaluierung

Die Reflexevaluierung dient der neuroanatomischen Lokalisation einer Läsion (De Lahunta und Glass, 2009b; Forterre et al., 2008; Garosi und Lowrie, 2013; Muguet-Chanoit et al., 2012; Thomas und Dewey, 2008). Durch die Evaluierung der spinalen Reflexe der Vorder- und Hintergliedmaßen überprüft der Kliniker die Integrität der spinalen Segmente (kraniozervikal: C₁-C₅; zervikothorakal: C₆-T₂; thorakolumbar: T₃-L₃ und lumbosakral: L₄-S₃) und kann so zwischen einer Läsion innerhalb des oberen oder unteren Motoneurons (s. 2.3.3.1. Einfluss der neuroanatomischen Lokalisation der Läsion) unterscheiden (Garosi und Lowrie, 2013).

Ferner dient die wiederholte Evaluierung der Reflexaktivität als wichtiges prädiktives Diagnostikum zur Beurteilung der Rekonvaleszenz verschiedener neurologischer Erkrankungen in der Human- und Veterinärmedizin (Aminoff, 2012; Carlstedt et al., 2012; Cozens et al., 2000; Dietz et al., 2009; Kiernan et al., 1999; Muguet-Chanoit et al., 2012). Eine wiedereinsetzende Reflexaktivität nach vorheriger Areflexie untermauert die Diagnose des „spinalen Schocks“ und wird als Besserung der klinisch-neurologischen Symptome interpretiert (Full et al., 2016; Okada et al., 2010; Smith und Jeffery, 2005). In der Humanmedizin wird die Korrelation zwischen Reflexaktivität und Schweregrad einer Spastizität diskutiert (Jang et al., 2013; Lin und Sabbahi, 1999; Nielsen und Sinkjaer, 1996; Pisano et al., 2000).

Diverse Studien nutzten außerdem die Beurteilung der Reflexantwort als Indikator für die Wirksamkeit therapeutischer Interventionen (Lervik et al., 2012) beziehungsweise von Anästhetika (Aghighi et al., 2012; Bergadano et al., 2006; Lervik et al., 2012; Martin et al., 1964; Pascoe, 2016; Rohrbach et al., 2009).

Experimentelle Untersuchungen am Tiermodell dienen der Erforschung der Nozizeption für die Human- und Veterinärmedizin. Hierzu wurde der Flexorreflex beim Kaninchen (Clarke et al., 2001), beim Hund (Bergadano et al., 2006; Martin et al., 1964), bei der Katze (Schomburg, 1990), beim Schaf (Rohrbach et al., 2014; Rohrbach et al., 2015) und beim Pferd (Spadavecchia et al., 2002; Spadavecchia et al., 2003) als Indikator verwendet.

3. Material und Methoden

3.1. Versuchsaufbau

Alle Untersuchungen wurden im selben Raum der Klinik für Kleine Haustiere der FU Berlin (KTK Berlin) und mit demselben Reflexhammer durchgeführt. Vor Studienbeginn wurde der Autor dieser Arbeit (= Untersucher 2; FG) von einem zertifizierten Veterinärneurologen des *European College of Veterinary Neurology* (= Untersucher 1; SL) in der Untersuchungstechnik der beiden Reflexe unterwiesen. Die Untersuchungen wurden gefilmt, digitalisiert und zu einem Film zusammengefügt. Für jeden der beiden Reflexe wurde jeweils ein Film angefertigt. Anhand der Filme evaluierten anschließend die Beurteiler die Reflexantwort. Das Interobserver Agreement (IA) zwischen den Beurteilern wurde mittels Reliabilitätsanalyse statistisch berechnet und ausgewertet.

In der ersten Untersuchung (Studie 1) wurde untersucht, ob sich das IA in Abhängigkeit des Untersuchers, des Geschlechts, des Gewichts, des Alters, der Felllänge, der Körperseite des Hundes und der Expertise des Beurteilers unterscheidet. Hierzu untersuchten Untersucher 1 und 2 den Bizepssehnenreflex beim Hund jeweils unabhängig voneinander.

In einer zweiten Untersuchung (Studie 2) wurde das IA zwischen dem Bizepssehnenreflex (Studie 2a) und dem Patellarsehnenreflex (Studie 2b) miteinander verglichen. In dieser Studie wurden ausschließlich Untersuchungen, die von Untersucher 1 durchgeführt wurden waren, ausgewertet. In Studie 2 wurde ebenfalls statistisch untersucht, ob die Faktoren *Alter*, *Geschlecht*, *Gewicht*, *Felllänge*, *Körperseite* und *Beurteilerexpertise* einen Einfluss auf das IA der Beurteilung Reflexantwort der beiden Reflexe nehmen.

3.2. Probanden

Die Hundebesitzer wurden vor Teilnahme an der Studie über den Ablauf der Testdurchführung informiert, erklärten sich mit der Teilnahme ihrer Hunde an der Studie einverstanden und waren während der Durchführung anwesend. Einschlusskriterien für die Hunde waren eine unauffällige allgemeine, neurologische sowie orthopädische Untersuchung am Tag der Versuchsdurchführung (s. 3.3.3. Voruntersuchung). Hunde, die vorberichtlich eine neurologische Erkrankung hatten, wurden von den Studien exkludiert. Des Weiteren wurden

Untersuchungen, die aufgrund starker Unruhe des Tieres, mangelhafter Videoaufzeichnung oder technischer Probleme nicht zu beurteilen waren, ebenfalls von der Evaluierung durch die Beurteiler ausgeschlossen.

Das Alter der Hunde wurde jeweils auf das Quartal auf- beziehungsweise abgerundet, das Gewicht wurde am Tag der Untersuchungen mittels einer Hundewaage (BIZERBA-Werke Wilhelm Kraut KG, Balingen (Württ.)) in der KTK Berlin ermittelt.

3.3. Methoden

3.3.1. Unterteilung der Probanden

In Anlehnung an die Literatur wurden die Gliedmaßen für beide Studien gemäß der untersuchten Einflussfaktoren *Geschlecht*, *Alter*, *Gewicht*, *Körperseite*, *Felllänge* und *Untersucher* in jeweils zwei Kategorien unterteilt (s. Tabelle 2). Das Fell wurde als „lang“ kategorisiert, wenn die Felllänge mindestens 8 cm betrug.

Tabelle 2. Kategoriale Unterteilung der untersuchten Gliedmaßen gemäß den untersuchten Risikofaktoren in Studie 1.

Einflussfaktor	Kategorie 1	Kategorie 2
Intrinsische Faktoren		
Geschlecht	weiblich	männlich
Alter	< 6 Jahre	≥ 6 Jahre
Gewicht	< 20 kg	≥ 20 kg
Felllänge	kurz	lang
Extrinsische Faktoren		
Untersucher	Untersucher 1	Untersucher 2
Körperhälfte	rechts	links

3.3.2. Auswahl und Unterteilung der Beurteiler

Die Beurteiler wurden gemäß ihrem Ausbildungsgrad in drei Gruppen unterteilt. Gruppe 1 bestand aus dem Autor dieser Arbeit (FG) und drei zertifizierten Veterinärneurologen des *European College of Veterinary Neurology*. Gruppe 2 setzte sich zusammen aus praktizierenden Tierärzten der KTK Berlin ohne besondere Affinität zur Veterinärneurologie. Die Beurteiler der Gruppe 2 hatten einen vergleichbaren Ausbildungsgrad und zwei bis maximal drei Jahre zuvor ihre Approbation erlangt. Gruppe 3 bestand aus Studenten der Veterinärmedizin an der Freien Universität Berlin. Die drei Studenten befanden sich im letzten Studienjahr und nahmen an der Studie während ihrer klinischen Rotation an der KTK Berlin teil.

Tabelle 3. Beurteiler in Studie 1. Namenskürzel in Klammern.

	Veterinärneurologen	Praktizierende Tierärzte	Studenten der Veterinärmedizin
1.	Neurologe 1 (SL)	Tierarzt 1 (KB)	Student 1 (MW)
2.	Neurologe 2 (NS)	Tierarzt 2 (SB)	Student 2 (CD)
3.	Neurologe 3 (HV)	Tierarzt 3 (AD)	Student 3 (SW)
4.	Neurologe 4 (FG)	Tierarzt 4 (AC)	Student 4 (HM)

Gemäß dieser Unterteilung setzte sich Gruppe 1 aus Beurteilern mit der erwartbar höchsten, Gruppe 2 mittleren und Gruppe 3 geringsten Beurteilerexpertise zusammen. Alle Beurteiler evaluierten die Reflexantwort separat voneinander und ohne Kenntnis der Evaluierungen der jeweils anderen Beurteiler.

Es gilt zu berücksichtigen, dass in Studie 1 die beiden Untersucher gleichzeitig Beurteiler (*Neurologe 1* und *4*) waren.

Für Studie 2, in der der Autor dieser Arbeit alleiniger Untersucher war, wurde dieser von der Beurteilung ausgeschlossen (*Neurologe 4*). Damit die Beurteilergruppen vergleichbar groß blieben, wurde ebenfalls jeweils ein Beurteiler der anderen beiden Gruppen von der Beurteilung ausgeschlossen. Hierzu wurde die Kombination von Beurteilern ausgewählt, die in Studie 1

gemeinsam den höchsten *ICC*-Wert (s. Tabelle 4 und 5) bei der Reflexaktivitätsbeurteilung erzielte.

Tabelle 4. Vergleich des *ICC* aus Studie 1 für alle möglichen Beurteilerkombinationen à drei Beurteiler aus Gruppe 2. Die Beurteilerkombination mit dem höchsten *ICC* (grau hinterlegt) wurde für Studie 2 genutzt.

	Beurteilerkombination	ICC
1	Tierarzt 1 + 2 + 3	0,79
2	Tierarzt 3 + 4 + 1	0,75
3	Tierarzt 4 + 1 + 2	0,75
4	Tierarzt 2 + 3 + 4	0,73

Tabelle 5. Vergleich des *ICC* aus Studie 1 für alle möglichen Beurteilerkombinationen à drei Beurteiler aus Gruppe 3. Die Beurteilerkombination mit dem höchsten *ICC* (grau hinterlegt) wurde für Studie 2 genutzt.

	Beurteilerkombination	ICC
1	Student 1 + 2 + 3	0,68
2	Student 3 + 4 + 1	0,61
3	Student 2 + 3 + 4	0,55
4	Student 4 + 1 + 2	0,50

3.3.3. Voruntersuchung

Die neurologische Untersuchung im Vorfeld der Studienteilnahme umfasste:

1. Erfragung neurologischer Vorerkrankungen
2. Beurteilung von Bewusstsein, Haltung und Gang
3. Untersuchung der Propriozeption der Vorder- und Hinterextremitäten
 - a. Pfotenkorrekturreaktion
 - b. Hüpfreaktion
4. Untersuchung der spinalen Reflexe
 - a. Vorderextremität
 - i. Flexorreflex
 - ii. Bizepssehnenreflex
 - iii. Musculus extensor carpi radialis Reflex
 - b. Hinterextremität
 - i. Flexorreflex
 - ii. Patellarsehnenreflex
 - iii. Musculus tibialis cranialis Reflex
5. Untersuchung der Kopfnerven
 - a. Untersuchung der Drohanwort
 - b. Untersuchung des Pupillarlichtreflexes
 - c. Untersuchung der Gesichtssensibilität
 - d. Untersuchung des physiologischen Nystagmus
 - e. Untersuchung der positionellen Augenstellung
 - f. Untersuchung des Schluckreflexes

Jeder Hund wurde ebenfalls auf Dolenz der Halswirbelsäule bei Extension, Flexion ventral und lateral zu beiden Seiten sowie auf Schmerzhaftigkeit oberhalb der Thorax- und Lendenwirbelsäule untersucht. Die neurologischen Untersuchungen wurden ausnahmslos vom Autor dieser Arbeit durchgeführt.

Hunde, die vorberichtlich Erkrankungen des Bewegungsapparates hatten, von denen ein Einfluss auf die Reflexevaluierung erwartet wurde, oder die eine schmerzlose Positionierung des Tieres gefährdeten (beispielsweise eine hochgradige Arthrose), wurden von der

Reflexevaluierung der betroffenen Gliedmaße ausgeschlossen. Eine beispielsweise länger zurückliegende, ausgeheilte Fraktur hingegen war kein Ausschlusskriterium.

Die orthopädische Untersuchung umfasste:

1. Erfragung orthopädischer Vorerkrankungen
2. Beurteilung des Gangbildes
3. Vergleichende Palpation der Vorder- und Hinterextremitäten am stehenden und liegenden Tier
4. Flexion und Extension der Gelenke der Vorder- und Hinterextremitäten am stehenden und liegenden Tier.

3.3.4. Videodokumentation

3.3.4.1. Durchführung der Reflexevaluierung

Die Hunde wurden zur Testung in Seitenlage verbracht, die zu untersuchende Gliedmaße war hierbei bodenfern (Thomas und Dewey, 2008). Die Probanden wurden bei Bedarf vor Beginn der Testung beruhigt, so dass die Gliedmaßen zur Untersuchung entspannt waren.

Die Lichtverhältnisse passten sich innerhalb des Untersuchungsraumes automatisch an die Lichtstärke des einfallenden Lichtes an. Die Raumtemperatur wurde automatisch konstant gehalten. Die Kamera (HDR-FX7E, Sony®, Japan) wurde auf einer Höhe von 110 cm an einem Stativ befestigt. Der Kamerawinkel betrug 30°. Die Probanden wurden so positioniert, dass die zu untersuchende Gliedmaße in einer Flucht mit dem Kameraobjektiv lag (s. Abbildung 2).

Das Bild wurde manuell so heraus- oder hereingezoomt, dass jeweils sowohl der Hammerschlag, als auch die Reflexantwort wahrgenommen werden konnten. Beim Bizepssehnenreflex wurde hierbei besonders darauf geachtet, dass die Kontraktion des Musculus biceps brachii beurteilbar war. Beim Patellarsehnenreflex wurde die Einstellung so gewählt, dass die Streckung des Unterschenkels detektierbar war.

Abbildung 2. Durchführung der Untersuchung des Bizepssehnenreflexes (Hund Nr. 3, Studie 1).



Die Kamera verfügte über einen automatischen Weißabgleich und einen Autofokus, so dass auch Hunde mit dunklem Fell beurteilt werden konnten und die Bilder stets, also auch nach eventuellen Bewegungen des Tieres während der Testung, scharf blieben. Vor jeder Videoaufnahme wurden die Einstellungen hinsichtlich ihrer Beurteilbarkeit getestet.

3.3.4.2. Randomisierung

Den teilnehmenden Hunden wurde für jede der zu untersuchenden Gliedmaße eine randomisierte Aufzeichnungsnummer zugeteilt. Hierzu zog der Untersucher eine Nummer von 1 bis einschließlich 100 aus einer hierfür angefertigten Losbox. Die zugeloste Nummer wurde mit weißem Klebeband zur Wiedererkennung bei der späteren digitalen Videobearbeitung (s.

3.3.4.3. Digitale Videobearbeitung) gut sichtbar auf der zu untersuchenden Gliedmaße befestigt.

Die patientenbezogenen Daten wurden mit den dazugehörigen Aufzeichnungsnummern in einer Excel-Tabelle dokumentiert.

3.3.4.3. Digitale Videobearbeitung

Die digitale Videobearbeitung wurde mit dem Programm *Windows Movie Maker* (Version 2012, *Microsoft Corporation*®) durchgeführt. Hierbei wurde das Rohmaterial so geschnitten, dass eine Serie von zehn Schlägen mit dem Reflexhammer mit konstanter Frequenz beurteilt werden konnte. Insgesamt wurden zwei Filme (Studie 1 und Studie 2b) angefertigt und im mp4-Format gespeichert.

Am Anfang der Videos wurde eine Sequenz von sieben (Film 1) (s. Abbildung 4) beziehungsweise fünf (Film 2) Sekunden Länge eingespielt, in denen noch einmal die Codierung für die Reflexpräsenz- und –aktivitätsbeurteilung auf Englisch erläutert wurde. Zwischen den einzelnen Sequenzen wurden Animationen mit der jeweils folgenden Aufzeichnungsnummer eingefügt (s. Abbildung 4). Diese Animationen hatten eine Länge von jeweils 7 Sekunden. Abbildung 5 zeigt einen Screenshot einer Untersuchungssequenz aus Film 1. Beide Filme endeten mit einer Animation, in der sich bei den Beurteilern für ihre Teilnahme bedankt wurde (s. Abbildung 6).

Abbildung 4. Screenshot von Film 1: Schema zur Beurteilung der Reflexantwort des Bizepssehnenreflexes zu Beginn des Filmes.

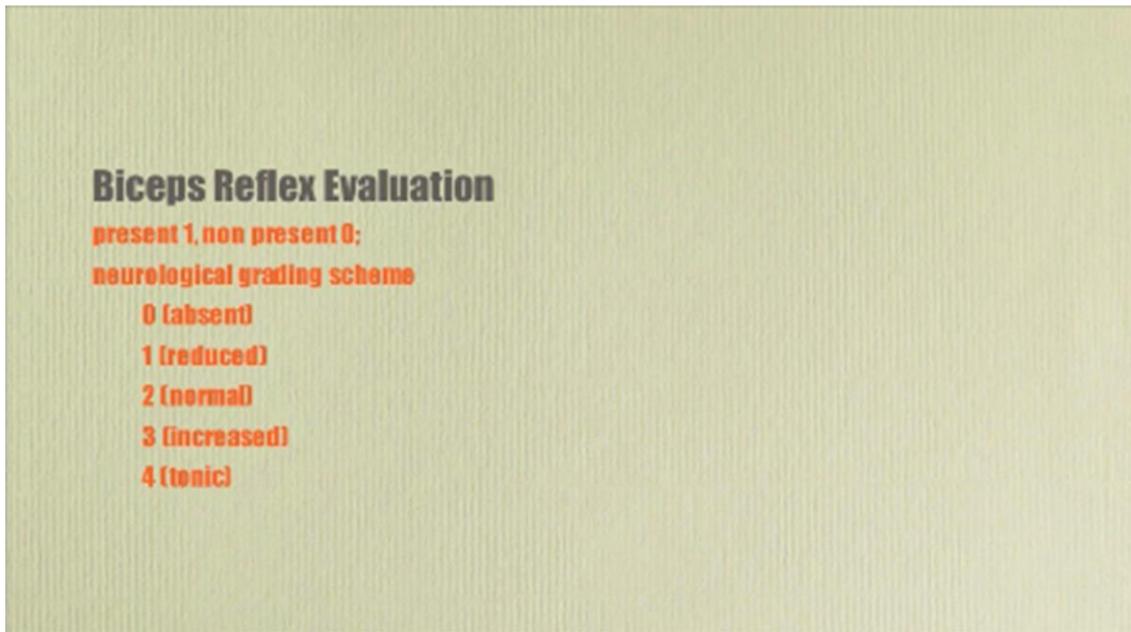


Abbildung 4. Screenshot von Film 1: die einzelnen Sequenzen wurden durch kurze Animationen, in denen die jeweils folgende Aufzeichnungsnummer eingeblendet wurde, voneinander getrennt.



Abbildung 5. Screenshot von Film 1: Untersuchung des Bizepssehnenreflexes (Aufzeichnungsnummer 48).



Abbildung 6. Screenshot von Film 1: in beiden Filmen wurde den Beurteilern in einer Abschlussequenz für ihre Teilnahme gedankt.



3.3.5. Videoanalyse

Beide Filme wurden den Beurteilern entweder direkt auf einem USB-Stick ausgehändigt oder mittels eines Zugangslinks zu einem erstellten Dropbox-Ordner (*Dropbox Inc.*®), in dem die Filme hinterlegt waren, zur Beurteilung zur Verfügung gestellt. Den Beurteilern wurde für jede der beiden Studien ein separater Evaluierungsbogen ausgehändigt. Auf diesem waren die Aufzeichnungsnummern in der Reihenfolge, wie sie im Film abliefen, in der linken Spalte (Spalte 1) eingetragen. Der Beurteiler sollte in den Spalten rechts daneben ankreuzen, ob die Reflexantwort *sichtbar/ vorhanden* oder *nicht sichtbar/ nicht vorhanden* war (Reflexpräsenzbeurteilung) (Spalten 2 - 3) und wie er die Stärke der Reflexantwort gemäß ordinaler Beurteilungsskala beurteilt (Reflexaktivitätsbeurteilung) (Spalten 4 - 8) (s. 9.1. Evaluierungsbögen). Die Beurteilungsskala umfasste fünf Kategorien (s. auch Abbildung 4): *0 = nicht vorhanden, 1 = reduziert, 2 = normal, 3 = gesteigert, 4 = tonisch* (Thomas und Dewey, 2008).

Die Evaluierungsbögen wurden dem Autor per Email zugesandt oder persönlich übergeben. Im Anhang sind die einzelnen Evaluierungsbögen der Beurteiler angefügt (s. Anhang, Kapitel 9.2 – 9.4).

3.3.6. Statistische Auswertung

3.3.6.1. Deskriptive Statistik

Für die unabhängigen Variablen *Alter* und *Gewicht* wurde der Median und der Minimal- und Maximalwert ermittelt. Für die Variablen *Felllänge*, *Geschlecht*, *Alter*, *Körperseite* und *Untersucher* wurde die Anzahl der Hunde beziehungsweise der Gliedmaßen für die jeweiligen Unterkategorien angegeben.

3.3.6.2. Reliabilitätsanalyse

Unter Reliabilität versteht man die Zuverlässigkeit einer Messung. Sie wird gemessen anhand der Übereinstimmungen der codierten Evaluierungen der Beurteiler zu einem bestimmten Zeitpunkt. Da jeder der etablierten Reliabilitätskoeffizienten seine Vor- und Nachteile hat, wird

die Erhebung verschiedener Reliabilitätskoeffizienten empfohlen. Zudem gilt es das Skalenniveau und die Anzahl der Beurteiler einer Untersuchung bei der Wahl des Reliabilitätskoeffizienten zu definieren (Gwet, 2014).

Im Folgenden werden die in dieser Arbeit ermittelten Reliabilitätskoeffizienten vorgestellt.

3.3.6.2.1. Prozentuale Übereinstimmung

Die *Prozentuale Übereinstimmung* ($r\%$) ist einer der bekanntesten Reliabilitätskoeffizienten und relativ einfach zu berechnen. Sie bildet den prozentualen Anteil der Anzahl n an Übereinstimmungen an der Gesamtzahl der Evaluierungen:

$$r\% = \frac{n_{\text{Übereinstimmungen}}}{n_{\text{Evaluierungen}}} \times 100$$

Formel I

Ein Vorteil liegt in der relativ einfachen Berechnung dieses Koeffizienten. Er wird allerdings wesentlich durch die Anzahl der Kategorien und Beurteiler beeinflusst, womit seine größte Aussagekraft bei dichotomer Beurteilung und zwei Beurteilern liegt (Gwet, 2014). Ferner berücksichtigt er nicht die Höhe der Differenz zwischen zwei Beurteilungen, weshalb er streng genommen nur zur Berechnung nominalskaliert Variablen herangezogen werden darf. Ein Konfidenzintervall kann nicht berechnet werden und eine Korrektur für zufällige Übereinstimmungen wird zur Berechnung nicht berücksichtigt.

Die *Prozentuale Übereinstimmung* wurde für die Reflexpräsenz- und –aktivitätsevaluierung berechnet, wobei sie aufgrund des ordinalen Skalenniveaus für letztere lediglich als Vergleich zu $r\%$ der Reflexpräsenzevaluierung diente. Der $r\%$ -Koeffizient wurde für jedes der möglichen Beurteilerpaare einer Gruppe berechnet.

Aus den ermittelten $r\%$ -Koeffizienten wurde das arithmetische Mittel ($\bar{X}r\%$) für jede Gruppe errechnet (vgl. Holsti, 1969):

$$\bar{X}r\% = \frac{\sum r\%}{n_{\text{Beurteilerkombinationen}}}$$

Formel II

3.3.6.2.2. Kappa-Analyse

Der Kappa-Koeffizient bietet eine zufallskorrigierte Berechnung der Reliabilität und ist daher einer der meistverwendeten Koeffizienten bei der Reliabilitätsanalyse kategorialer Variablen (Grouven et al., 2007).

Man unterscheidet drei verschiedene Modifikationen, die jeweils in Abhängigkeit der Anzahl der Beurteiler und verwendeten Kategorien eingesetzt werden:

- *Cohen's Kappa* (K_C)
- *gewichtetes Kappa* (K_w)
- *Fleiss' Kappa* (K_F)

Cohen's Kappa

Cohen's Kappa wurde erstmals 1960 von dem amerikanischen Psychologen Jacob Cohen beschrieben (Cohen, 1960) und bildet die mathematische Grundlage für die oben genannten Modifikationen des Kappa-Koeffizienten.

Die Basis seiner Berechnung bildet eine Übereinstimmungsmatrix (s. Abbildung 5), in der die beobachteten Häufigkeiten aller Beurteilungen (h_{1-k}) mit k Kategorien (Kat_{1-k}) eines Experiments durch zwei Beurteiler ($B1$ und $B2$) tabellarisch aufgelistet werden.

Abbildung 5. Übereinstimmungsmatrix als Grundlage zur Berechnung der Kappa-Koeffizienten.

		Beurteiler 1			Zeilensumme (h_i):
		Kat_1	...	Kat_k	
Beurteiler 2	Kat_1	h_{11}	...	h_{1k}	$h_{1.} = \sum_i^k h_{1i}$

	Kat_k	h_{k1}	...	h_{kk}	$h_{z.} = \sum_i^k h_{zi}$
Spaltensumme (h_i):		$h_{.1} = \sum_i^k h_{i1}$...	$h_{.k} = \sum_i^k h_{ik}$	$\Sigma\Sigma = N$

Die beobachteten übereinstimmenden Beurteilungen finden sich auf der Hauptdiagonalen (s. Abbildung 5, grau hinterlegt), Nichtübereinstimmungen in den übrigen Zellen.

Der relative Anteil der beobachteten übereinstimmenden Beurteilungen (p_o) ist der Quotient aus der Summe der beobachteten Übereinstimmungen dividiert durch die Gesamtzahl der Untersuchungen:

$$p_o = \frac{\sum_{i=1}^k h_{ii}}{N}$$

Formel III

- k = Anzahl der Merkmalskategorien
- i = Merkmalskategorie i (1,...,k)
- h_{ii} = Häufigkeiten auf der Hauptdiagonalen
- N = Gesamtzahl der Beurteilungen

Den Anteil der erwarteten übereinstimmenden Beurteilungen (p_e) errechnet man, indem man die Summe der jeweiligen Produkte der Randsummen (Zeilensumme x Spaltensumme) einer Kategorie durch das Quadrat der Anzahl aller Beurteilungen dividiert:

$$p_e = \frac{\sum_{i=1}^k h_{i.} \cdot h_{.i}}{N^2}$$

Formel IV

- k = Anzahl der Merkmalskategorien
- i = Merkmalskategorie i (1,...,k)
- $h_{i.}$ = Häufigkeiten in einer Zeile
- $h_{.i}$ = Häufigkeiten in einer Spalte
- N = Gesamtzahl der Beurteilungen

Mit Hilfe von p_o und p_e lässt sich nun K_C ausrechnen:

$$K_C = \frac{p_o - p_e}{1 - p_e}$$

Formel V

Es gilt bei der Verteilung der Beurteilungen innerhalb der Übereinstimmungsmatrix zwei Kriterien zu berücksichtigen:

1. Symmetrie der Verteilung

- a. Symmetrie: eine symmetrische Verteilung liegt vor, wenn die Proportionen über den Zeilensummen und den Spaltensummen gleich verteilt sind. Perfekte Symmetrie liegt dann vor, wenn die jeweils korrespondierenden Zeilen- und Spaltensummen einer Kategorie gleich sind.
- b. Asymmetrie: eine asymmetrische Verteilung liegt vor, wenn die Proportionen über den Zeilensummen entgegen gesetzt derer der Spaltensummen sind und vice versa.

2. Ausgewogenheit der Verteilung

- a. Balance: der Anteil der Beurteilungen ist über den Kategorien gleich verteilt. Bei einer dichotomen Beurteilungsskala, also mit jeweils 50% über jeder der beiden Kategorien.
- b. Inbalance: der Anteil der Beurteilungen ist über den Kategorien ungleichmäßig verteilt.

Die K_C -Koeffizienten und das jeweilige Konfidenzintervall (KI95%) wurden für jede mögliche Beurteilerkombination mittels *MedCalc* (Version 12.5.0.0, *MedCalc*®) berechnet. Für jede Gruppe wurde das arithmetische Mittel der K_C -Werte ($\bar{X}K_C$) ermittelt (Brott et al., 1989).

Die Verteilung der K_C -Koeffizienten innerhalb der Gruppen wurde auf signifikante Unterschiede mittels Kruskal-Wallis-Test zwischen den Gruppen geprüft. Hierbei war die Nullhypothese durch die Annahme definiert, dass die K_C -Koeffizienten über den Gruppen gleich verteilt sind. Lag eine signifikant unterschiedliche Verteilung zwischen den Gruppen vor, wurde die Nullhypothese abgelehnt. Das Signifikanzniveau lag bei 5%.

Die Interpretation des Kappa-Koeffizienten erfolgte gemäß dem von Landis und Koch (1977) vorgeschlagenen Modell (s. Tabelle 6).

Tabelle 6. Interpretation des Kappa-Koeffizienten (nach Landis und Koch (1977)).

Kappa-Koeffizient	Übereinstimmungsmaß
< 0,00	schlecht
0,00 – 0,20	kaum
0,21 – 0,40	gering
0,41 – 0,60	mittelmäßig
0,61 – 0,80	ordentlich
0,81 – 1,00	(fast) perfekt

Gewichtetes Kappa

Das *gewichtete Kappa* ist eine Modifikation des *Intraklassen-Korrelationskoeffizienten* und von K_C und berücksichtigt die Höhe der Diskrepanz zwischen zwei Beurteilungen (Fleiss und Cohen, 1973).

Für die ordinalskalierte Reflexaktivitätsanalyse wurde zur Einbeziehung der Höhe der Diskordanz zwischen zwei Beurteilern eine lineare Gewichtung (w_{ij}) verwendet. Wie oben bereits beschrieben, liegen übereinstimmende Beurteilungen auf der Hauptdiagonalen der Übereinstimmungsmatrix (s. Abbildung 5). Alle anderen Zellen der Übereinstimmungsmatrix beinhalten somit diskordante Beurteilungen. Es gilt allerdings zu berücksichtigen, dass diese unterschiedlich interpretiert werden müssen: je näher bei diskordanten Beurteilungen die beiden von den Beurteilern gewählten Kategorien beieinanderliegen, desto näher liegen sie an der Hauptdiagonalen. Im Umkehrschluss ist es so, dass je größer die Differenz zwischen den beiden gewählten Kategorien eines Beurteilerpaares ist, desto weiter entfernt von der Hauptdiagonalen liegen ihre Beurteilungen somit in der Übereinstimmungsmatrix.

Im folgenden Beispiel können die beiden Beurteiler aus fünf Kategorien (0-4) wählen:

		Beurteiler 1				
		0 (abwesend)	1 (reduziert)	2 (normal)	3 (gesteigert)	4 (tonisch)
Beurteiler 2	0	0-0 = 0
	1	1-0 = 1	1-1 = 0
	2	2-0 = 2	...	2-2 = 0
	3	3-0 = 3	3-3 = 0	...
	4	4-0 = 4	4-4 = 0

Mittels K_w werden nun die Zellen außerhalb der Hauptdiagonalen mit dem Faktor w_{ij} multipliziert und somit gewichtet.

Der Gewichtung der Zellen liegt die mathematische Formel zugrunde:

$$w_{ij} = \frac{i}{k-1}$$

Formel VI

w_{ij} = Gewichtungsfaktor

i = Differenz zwischen den Beurteilern

k = Anzahl der Kategorien (= 5)

Je weiter die Zellen von der Hauptdiagonalen entfernt liegen, umso größer der verwendete Gewichtungsfaktor. Somit wurden die Zellen in der vorliegenden Arbeit folgendermaßen gewichtet:

Tabelle 7. Übereinstimmungsmatrix mit den zugrunde gelegten Gewichtungsfaktoren der jeweiligen Zellen.

	0	1	2	3	4
0	0	0,25	0,5	0,75	1
1	0,25	0	0,25	0,5	0,75
2	0,5	0,25	0	0,25	0,5
3	0,75	0,5	0,25	0	0,25
4	1	0,75	0,5	0,25	0

In die Formel für K_w fließt w_{ij} wie folgt ein (Bortz und Lienert, 1998):

$$K_w = 1 - \frac{\sum_{i=1}^k \sum_{j=1}^k w_{ij} h_{ii}}{\sum_{i=1}^k \sum_{j=1}^k w_{ij} h_{i.} h_{.i}}$$

Formel VII

- k = Anzahl der Merkmalskategorien
- i = Merkmalskategorie i (1,...,k)
- j = Merkmalskategorie j (1,...,k)
- h_{ii} = Häufigkeiten auf der Hauptdiagonalen
- $h_{i.}$ = Häufigkeiten in einer Zeile
- $h_{.i}$ = Häufigkeiten in einer Spalte

Sind beispielsweise alle Felder außerhalb der Hauptdiagonalen (h_{ii}) leer, so werden lediglich die Zellen der Hauptdiagonalen mit dem Gewicht „0“ multipliziert und somit ist die Zählersumme = 0 und $K_w = 1$ (Bortz und Lienert, 1998). Die Interpretation von K_w erfolgt auf

der gleichen Grundlage wie die von K_C . Somit kann $K_w = 1$ als *perfekte* Übereinstimmung interpretiert werden (s. Tabelle 6).

Die K_w -Koeffizienten und das jeweilige KI95% wurden für jedes Untersucherpaar mittels des Programms *MedCalc* (Version 12.5.0.0, *MedCalc*®) berechnet. Für jede Gruppe wurde das arithmetische Mittel der K_w -Werte ($\bar{X}K_w$) errechnet.

Die K_w -Koeffizienten der drei Gruppen wurden mittels Kruskal-Wallis-Test auf signifikante Unterschiede geprüft. Hierbei war die Nullhypothese durch die Annahme definiert, dass die K_w -Koeffizienten über den Gruppen gleich verteilt sind. Lag eine signifikant unterschiedliche Verteilung zwischen den Gruppen vor, wurde die Nullhypothese abgelehnt. Das Signifikanzniveau lag bei 5%.

Fleiss' Kappa

Fleiss' Kappa ist ein Reliabilitätsmaß für die Übereinstimmung von mehr als zwei Beurteilern unter der Verwendung kategorialer Daten (Fleiss, 1971). Eine Übereinstimmungsmatrix für K_F listet die Anzahl der Beurteilungen jeder Kategorie k_j auf. Vorausgesetzt jeder Beurteiler gibt eine Beurteilung ab, so ist die Zeilensumme für ein Beurteilungsobjekt i immer gleich der Anzahl der Beurteiler m .

Zunächst definiert man die *Güte der Beurteilerübereinstimmung* P_i (Bortz und Lienert, 1998) bei einem Objekt i :

$$P_i = \frac{\sum_{j=1}^k n_{ij} \times (n_{ij} - 1)}{m \times (m - 1)}$$

Formel VIII

- k = Anzahl der Kategorien
- m = Anzahl der Beurteiler
- j = Merkmalskategorie j (1,...,k)
- i = Beurteilungsobjektvariable
- n_{ij} = Anzahl der Beurteiler, die bei Objekt i übereinstimmend die Kategorie j gewählt haben

Der Durchschnitt aller P_i -Werte (\bar{P}) gibt an, mit welcher Wahrscheinlichkeit die beiden Beurteiler eines beliebigen Beurteilerpaares identisch beurteilt haben (Bortz und Lienert, 1998):

$$\bar{P} = \frac{\sum_{i=1}^N P_i}{N}$$

Formel IX

N = Gesamtzahl der Beurteilungen

i = Merkmalskategorie i

Der relative Anteil an Beurteilungen für eine Kategorie j (P_j) ist (Bortz und Lienert, 1998):

$$P_j = \frac{\sum_{i=1}^N n_{ij}}{N \times m}$$

Formel X

Mithilfe P_j errechnet man die Wahrscheinlichkeit \bar{P}_e , bei der die beiden Beurteiler eines beliebigen Beurteilerpaares unabhängig voneinander dieselbe Kategorie j wählen:

$$\bar{P}_e = \sum_{j=1}^k P_j^2$$

Formel XI

Die Formel für K_F lautet nun:

$$K_F = \frac{\bar{P} - \bar{P}_e}{1 - \bar{P}_e}$$

Formel XII

Die K_F -Koeffizienten, das jeweilige KI95% und der dazugehörige Standardfehler (SF) wurden für jede Beurteilergruppe mittels des von Prof. Deen Freelon online zur Verfügung gestellten

Kalkulationsprogramms¹ berechnet. *Fleiss' Kappa* einer Gruppe galt als signifikant unterschiedlich, wenn sein KI95% nicht den jeweiligen K_F einer anderen Gruppe einschloss (Thaller und Hughes, 2014).

3.3.6.2.3. Intraklassen-Korrelationskoeffizient

Der *Intraklassen-Korrelationskoeffizient (ICC)* drückt die Varianz der Beurteilung zwischen den Messobjekten proportional zur Varianz aller Beurteilungen aus (Petrie und Watson, 2013). Durch die Berücksichtigung der Varianz ähnelt er dem *gewichteten Kappa*, ist aber für mehr als zwei Beurteiler geeignet und stellt somit vor allem für die nicht-dichotome Beurteilung der Reflexaktivität einen sinnvollen alternativen Reliabilitätskoeffizienten zu K_F dar. Dem Algorithmus von Shrout und Fleiss (1979) folgend, wurde der *ICC-Typ ICC(3,k)* zur statistischen Analyse verwendet:

$$ICC(3,k) = (BMS - EMS)/BMS$$

Formel XIII

BMS = („between targets mean square“) = Varianz zwischen den Messobjekten

EMS = („error of mean square“) = Restvarianz

Für die Interpretation des *ICC* existieren verschiedene Modelle (vgl. Altman, 1991; Chinn, 1991; LeBreton und Senter, 2008; Lewis et al., 2012; Portney und Watkins, 2015; Vincent und Weir, 1999).

¹ <http://www.dfreelon.org/> (abgerufen am 21.02.2018)

In dieser Arbeit wurde sich allerdings für ein relativ strenges Interpretationsmodell des *ICC* entschieden (Vincent und Weir, 1999):

ICC	Übereinstimmungsmaß
< 0,70	schlecht
0,71 – 0,80	zweifelhaft
0,81 – 0,90	mittelmäßig
0,91 – 1,00	hoch

Die Berechnung des *ICC* wurde mittels *SPSS*® (Version 21.0.0.0, *IBM*®) durchgeführt.

3.3.6.2.4. Table of Agreement

Die *Table of Agreement* orientiert sich an der Arbeit von Stam und van Crevel (1990) (Stam und van Crevel, 1990) und bildet im Grunde eine Erweiterung des *r*%-Koeffizienten (s. Formel I).

Die Ergebnisse der Reflexaktivitätsevaluierung wurden für jede Gruppe hinsichtlich ihrer Übereinstimmungsstufe ausgezählt und in einem kumulierten Balkendiagramm graphisch dargestellt. Die Übereinstimmungsstufe wurde definiert als die Anzahl der übereinstimmenden Evaluierungen einer Testung und die Differenz zwischen der jeweils höchsten und niedrigsten Evaluierung zwischen den Beurteilern innerhalb einer Testung.

Beispiel einer fiktiven Testung:

	Beurteiler 1	Beurteiler 2	Beurteiler 3
Testung x	1	1	3

In diesem Beispiel ist die:

- Anzahl der übereinstimmenden Evaluierungen = 2 (Beurteiler 1 und 2)
- maximale Differenz zwischen den Evaluierungen = (3-1) = 2.

Die Übereinstimmungsstufen wurden wie folgt für die beiden Studien von 1-4 kategorisiert:

Übereinstimmungsstufe	Bezeichnung	Anzahl übereinstimmender Evaluierungen	Differenz zwischen den Evaluierungen
1	Vollständige Übereinstimmung	Studie 1: 4 Studie 2: 3	0
2	Partielle (Nicht-) Übereinstimmung (1 Punkt)	Studie 1: 3 Studie 2: 2	1
3	Partielle (Nicht-) Übereinstimmung (≥ 2 Punkte)	Studie 1: 3 Studie 2: 2	2-4
4	Vollständige Nichtübereinstimmung	Studie 1: 0 oder 2 Studie 2: 0	2-4

Je höher die Kategorie der Übereinstimmungsstufe, desto größer ist also der Dissens zwischen den Beurteilern einer Gruppe beziehungsweise desto ungenauer ist die Evaluierung einer Gruppe für eine bestimmte Testung.

3.3.6.3. Interpretation der Reliabilitätsanalyse

Die Höhe von Kappa ist von verschiedenen Faktoren abhängig und so kann seine Interpretation in Kombination mit der *Prozentualen Übereinstimmung* schwierig und zum Teil widersprüchlich sein (Burn, 2011; Feinstein und Cicchetti, 1990).

Feinstein und Cicchetti (1990) proklamierten zwei zentrale Paradoxa, die es bei der Interpretation von Kappa zu berücksichtigen gilt:

1. Trotz eines hohen IA, können relativ niedrige Kappa-Werte resultieren.
2. Eine asymmetrische, inbalancierte Verteilung der Beurteilungen kann in einem höheren Kappa-Wert resultieren, als bei einer symmetrischen und balancierten Verteilung.

Wie aus Formel V hervorgeht, ist die Höhe von Kappa wesentlich von dem erwarteten Anteil an übereinstimmenden Beurteilungen (p_e) abhängig: je niedriger p_e , umso höher ist Kappa und vice versa (Flight und Julious, 2015). Aus Formel IV wird darüber hinaus deutlich, dass p_e abhängig von der Verteilung der Randsummen ist. Somit kann es zum ersten Paradoxon kommen, wenn die Beurteilungen symmetrisch inbalanciert verteilt sind.

Im Folgenden werden in der Literatur empfohlene Interpretationsparameter vorgestellt, die in dieser Arbeit verwendet wurden.

3.3.6.3.1. Prävalenz-Index (*PI*)

Kappa wird wesentlich von der Prävalenz eines untersuchten Merkmals innerhalb der beobachteten Gruppe beeinflusst. Die Prävalenz (lat.: *praevalare* = Übergewicht, Vorrang haben) ist definiert als die Häufigkeit eines bestimmten Merkmals innerhalb einer untersuchten Gruppe und stellt meist im Zusammenhang mit der Detektion bestimmter Erkrankungen ein epidemiologisches Maß zur Charakterisierung des Krankheitsgeschehens dar (Pschyrembel und Dornblüth, 2002). Kappa sinkt, wenn die Prävalenz eines untersuchten Merkmals gegen 1 tendiert und es somit zu einer inbalancierten Verteilung der Beurteilungen über den Kategorien kommt (Hansen et al., 1994; Sim und Wright, 2005). Weist eine untersuchte Gruppe beispielsweise eine starke Homogenität eines Merkmals auf, steigt die Wahrscheinlichkeit einer rein zufälligen Übereinstimmung und somit sinkt Formel V folgend, Kappa. Paradoxerweise können die Beurteiler dennoch eine relativ hohe *Prozentuale Übereinstimmung* aufweisen. Für jeden denkbaren $r\%$ -Wert gilt, dass wenn der *PI* steigt, Kappa sinkt. Die Höhe des *PI* gibt einen Hinweis über die Homogenität der untersuchten Gruppe (Burn und Weir, 2011).

Die in der Literatur vorhandenen Berechnungsmodelle für *PI* setzen eine dichotome Verteilung des untersuchten Merkmals voraus (vgl. Burn und Weir, 2011; Sim und Wright, 2005). Es wurden somit die Evaluierungen sowohl für die Reflexpräsenz als auch –aktivität in zwei Kategorien („*normal*“ vs. „*abnormal*“) unterteilt (s. Tabelle 8).

Tabelle 8. Dichotomisierung der Evaluierungskategorien zur Berechnung des Prävalenz-Indexes.

	Reflexpräsenz	Reflexaktivität
Normal	vorhanden (= 1)	normal (= 2)
Abnormal	nicht vorhanden (= 0)	abwesend (= 0) reduziert (= 1) gesteigert (= 3) tonisch (= 4)

Auf der Grundlage einer Übereinstimmungsmatrix für zwei Beurteiler ($B1$ und $B2$) und der dichotomen Verteilung der Reflexpräsenz und -aktivität (s. Tabelle 9) errechnet sich PI Formel XIV folgend.

Tabelle 9: Übereinstimmungsmatrix für zwei Beurteiler ($B1$ und $B2$) zur Berechnung des Prävalenz-Indexes (PI).

		B1		Zeilensumme
		normal	abnormal	
B2	normal	A	B	A+B
	abnormal	C	D	C+D
Spaltensumme		A+C	B+D	$N = (A+B) + (C+D)$

$$PI = \frac{|A - D|}{N}$$

Formel XIV

Für die Berechnung des PI für zwei bis vier Beurteiler wurde das von Burn und Weir (2011) vorgeschlagene Berechnungsmodell zugrunde gelegt (s. Formel XV). Hierzu wurden die Beurteilungen sowohl der Reflexpräsenz, als auch der -aktivität gemäß Tabelle 10 klassifiziert.

Tabelle 10: Klassifikation der Sequenzen zur Berechnung des Prävalenz-Indexes für zwei bis vier Beurteiler.

Anzahl der übereinstimmenden Beurteiler	Anzahl der Sequenzen als <i>normal</i> klassifiziert	Anzahl der Sequenzen als <i>abnormal</i> klassifiziert	Gesamtzahl der Sequenzen, die als <i>normal</i> klassifiziert wurden (R_1)	Gesamtzahl der Sequenzen, die als <i>abnormal</i> klassifiziert wurden (R_2)
4	A	B	4 x A	4 x B
3	C	D	3 x C	3 x D
2	E	F	2 x E	2 x F
1	G	H	1 x G	1 x H
0	H	I	0 x H	0 x I

$$PI = \frac{R_1 - R_2}{N \times n}$$

Formel XV

N = Anzahl der Beurteiler

n = Gesamtzahl aller Beurteilungen

Ein PI von 0 bedeutet somit, dass die Beurteilungen über den Kategorien gleich verteilt sind (= Balance), während ein PI von 1 bedeutet, dass die Übereinstimmungen alle in eine Kategorie fallen (= Inbalance) (Burn und Weir, 2011).

3.3.6.3.2. Bias-Index (BI)

Der BI ist die Höhe der Differenz zwischen den jeweils diskordanten und konkordanten Beurteilungen von zwei Beurteilern dividiert durch die Gesamtzahl aller Beurteilungen (n) (Sim und Wright, 2005).

Legt man Tabelle 9 zugrunde, errechnet sich der BI wie folgt:

$$BI = \frac{|B - C|}{N}$$

Formel XVI

Der BI ist somit umso höher, je asymmetrischer die Beurteilungen jeweils in den Feldern B und C liegen (Sim und Wright, 2005).

Vergleicht man nun die Kappa-Koeffizienten von zwei Studien, in denen jeweils die exakt gleiche Verteilung der übereinstimmenden Beurteilungen (Felder A und D) vorliegt, so ist Kappa paradoxerweise in der Studie höher, in der die nichtübereinstimmenden Beurteilungen asymmetrischer in den Feldern B und C der Übereinstimmungsmatrix verteilt sind (2. Paradoxon, s. o.).

3.3.6.3.3. Maximal erzielbarer Kappa (K_{max})

Der maximal mögliche Anteil der beobachteten Übereinstimmungen (p_o) in einer bestimmten Kategorie ist durch den Minimalwert der tatsächlich in dieser Kategorie abgegebenen Beurteilungen in den korrespondierenden Randsummen begrenzt (Sim und Wright, 2005). In einer stark asymmetrischen, inbalancierten Kontingenztabelle beispielsweise kann p_o Formel III folgend keinen hohen Wert annehmen und Kappa muss demzufolge relativ niedrig sein.

Zusammen mit dem jeweils berechneten Kappa-Wert wird der Anteil von K_{max} definiert, der bereits durch K_C oder K_w berechnet wurde (Feinstein und Cicchetti, 1990).

Die Summe der Minimalwerte der korrespondierenden Zeilen- beziehungsweise Randsummen bildet den maximalen p_o -Wert p_{oM} (Feinstein und Cicchetti, 1990; Flight und Julious, 2015):

$$p_{oM} = \min\left(\frac{h_{1.}}{N} \mid \frac{h_{.1}}{N}\right) + \dots + \min\left(\frac{h_{k.}}{N} \mid \frac{h_{.k}}{N}\right)$$

Formel XVII

$h_{1...k.}$ = Zeilensumme

$h_{.1...k}$ = Spaltensumme

N = Gesamtzahl aller Beurteilungen

K_{max} errechnet sich nun wie folgt:

$$K_{max} = \frac{p_{oM} - p_e}{1 - p_e}$$

Formel XVIII

K_{max} ist gleich dem berechneten K_C - oder K_w -Wert wenn p_{oM} gegen 1 tendiert, beziehungsweise den Wert 1 annimmt oder wenn die Beurteiler bereits mit maximal möglicher Übereinstimmung beurteilt haben ($p_o = p_{oM}$). K_{max} ist nur geringgradig höher als der berechnete Kappa-Wert wenn die Randsummen relativ symmetrisch verteilt sind und deutlich höher wenn die Randsummen stark asymmetrisch und inbalanciert verteilt sind oder aber die Beurteiler trotz imperfekt symmetrischer Inbalance den höchst möglichen p_o erzielen (Feinstein und Cicchetti, 1990).

Die Differenz zwischen K_{max} - und Kappa-Koeffizient definiert den Anteil des Einflusses, der auf das IA durch eine inbalancierte Randsummenverteilung wirkt (Sim und Wright, 2005). Eine inbalancierte Randsummenverteilung entsteht zum Beispiel durch individuelle Faktoren der Beurteiler wie zum Beispiel eine eingeschränkte Sensitivität ein bestimmtes Merkmal korrekt zu interpretieren.

3.3.6.3.4. Interpretation der Ergebnisse unter dem Kriterium der *klinischen Akzeptanz*

Bei der Einordnung und dem Vergleich der eigenen Reliabilitätsanalyse ist es hilfreich einen Grenzwert zu definieren, unter welchem das IA als *nicht klinisch akzeptabel* definiert ist (Burn und Weir, 2011; Sim und Wright, 2005). Sim und Wright (2005) und auch Burn et al., (2009, 2011) empfahlen hierzu einen Cut-off von 0,40 für den kalkulierten Kappa-Wert und von 75% für die kalkulierte *Prozentuale Übereinstimmung*.

Demzufolge wird die *klinische Akzeptanz* der in dieser Arbeit präsentierten Ergebnisse der Reliabilitätsanalyse unter Berücksichtigung der Grenzwerte für *r%* bei 75% und für Kappa bei 0,40 interpretiert:

Kategorie	r%	K _C / K _w	Interpretation
I.	≥ 75	≥ 0,40	klinisch akzeptabel
II.	< 75	0,00 – 1,00	klinisch nicht akzeptabel
III.	≥ 75	< 0,40	nicht eindeutig

Ein vermeintlich niedriges IA fällt nach Burn und Weir (2011) somit in zwei Kategorien:

1. statistische Nichtübereinstimmung zwischen den Beurteilern (Kategorie II)
2. statistische Ambiguität aufgrund einer inbalancierten Probandenpopulation (Kategorie III).

Vor allem bei den Ergebnissen, die in Kategorie III fallen, ist es notwendig die dazugehörigen Interpretationsparameter in die Interpretation einzubeziehen (s. o.).

3.3.6.4. Regressionsanalyse

Mittels logistischer Regressionsanalyse wurde untersucht inwieweit das IA von verschiedenen intrinsischen und extrinsischen Faktoren beeinflusst wird. Untersucht wurden die Risikofaktoren *Untersucher, Geschlecht, Gewicht, Alter, Körperseite, Felllänge* und *Beurteilerexpertise*.

Modellbildung

Das IA wurde als abhängige Variable (Regressand) definiert. Alle Nichtübereinstimmungen wurden mit 0 und alle Übereinstimmungen mit 1 codiert. Konkret bedeutet dies, dass die Anzahl der Nichtübereinstimmungen bei der Reflexpräsenzevaluierung der Anzahl der ausgezählten Nichtübereinstimmungen entspricht, bei der Reflexaktivitätsevaluierung hingegen entspricht die hier zur Analyse herangezogene Anzahl der Nichtübereinstimmungen der Summe der Übereinstimmungsstufen *partielle (Nicht-)Übereinstimmung (1 Punkt)*, *partielle (Nicht-)Übereinstimmung (≥ 2 Punkte)* und *vollständige Nichtübereinstimmung*.

Aufgrund dieser Dichotomie der abhängigen Variablen wurde ein binär logistisches Regressionsmodell verwendet. Die Risikofaktoren wurden als unabhängige Variablen (Regressoren) unterteilt und codiert (s. Tabelle 11).

Die Berechnung der Regressionsanalyse wurde mittels SPSS® (Version 21.0.0.0, IBM®) durchgeführt.

Tabelle 11. Übersicht der untersuchten Risikofaktoren. ¹= nur in Studie 1 untersucht

Risikofaktor	Bezeichnung	Kriterium	Codierung	Referenzkategorie (*)
Untersucher¹	Untersucher 1	SL	0	*
	Untersucher 2	FG	1	
Geschlecht	weiblich	weiblicher Hund	0	*
	männlich	männlicher Hund	1	
Gewicht	leicht	≤ 20kg	0	*
	schwer	> 20kg	1	
Alter	jung	< 6 Jahre	0	*
	alt	≥ 6 Jahre	1	
Felllänge	kurz	Felllänge (< 8cm)	0	*
	lang	Felllänge (≥ 8cm)	1	
Körperseite	rechts	rechte Gliedmaße	0	*
	links	linke Gliedmaße	1	
Beurteilerexpertise	Neurologe	Neurologe	1	*
	Praktizierender Tierarzt	Praktizierender Tierarzt	2	
	Student	Student	3	

Da die einzelnen Gliedmaßen jeweils von jedem Beurteiler evaluiert wurden, gelten die erhobenen Daten als Longitudinaldaten, weshalb ein gemischtes Regressionsmodell verwendet wurde.

Die Basisgleichung der binären logistischen Regression für ein Ereignis Z lautet:

$$Z = C + BX$$

Formel XIX

C = Konstante

B = Regressionskoeffizient

X = unabhängige Variable

Der Regressionskoeffizient im hier angewendeten Modell ist der standardisierte Regressionskoeffizient β ($R\beta$), der sich nach Standardisierung des Mittelwertes (= 0) und der Varianz (= 1) sowohl aus der abhängigen und unabhängigen Variablen errechnet.

Das hier angewendete Regressionsmodell folgt demnach:

$$\text{Nichtübereinstimmung} = C + R\beta \times X$$

Formel XX

Wahrscheinlichkeits- und Chancenanalyse

Die Wahrscheinlichkeit, dass das Ereignis Z in einem Experiment eintritt (P_Z), bildet der Quotient aus der Anzahl des tatsächlich beobachteten Ereignisses (n_Z) dividiert durch die Anzahl aller im Experiment durchgeführten Untersuchungen (n):

$$P_Z = \frac{n_Z}{n}$$

Formel XXI

Die Wahrscheinlichkeit einer Nichtübereinstimmung innerhalb einer Gruppe bei einer Evaluierung lässt sich somit wie folgt ausdrücken:

$$P_{\text{Nichtübereinstimmung}} = \frac{n_{\text{Nichtübereinstimmung}}}{n_{\text{Evaluierungen}}}$$

Formel XXII

Die Wahrscheinlichkeit, dass das Ereignis Z eintritt, errechnet sich aus der logistischen Funktion:

$$P_Z = \frac{1}{(1 + e^{-Z})}$$

Formel XXIII

e = Eulersche Zahl ($\approx 2,718$)

z = Logit (= lineares Regressionsmodell der unabhängigen Variablen)

Eine Chance (*odd*) ist definiert als das Verhältnis zweier Wahrscheinlichkeiten, beziehungsweise P_Z dividiert durch die Gegenwahrscheinlichkeit:

$$\text{odd}(Z) = \frac{P_Z}{1 - P_Z}$$

Formel XXIV

Das bedeutet:

$$\text{Chance}_{\text{Nichtübereinstimmung}} = \frac{\text{Wahrscheinlichkeit}_{\text{Nichtübereinstimmung}}}{\text{Wahrscheinlichkeit}_{\text{Übereinstimmung}}}$$

Formel XXV

Die Verwendung der logarithmierten Chance (*Logit*) im Regressionsmodell hat den Vorteil, dass sie zum einen unbegrenzt und zum anderen symmetrisch ist. Sie errechnet sich wie folgt:

$$\text{Logit} = \ln\left(\frac{P_Z}{1 - P_Z}\right)$$

Formel XXVI

Odds ratio

Die *Odds ratio* (*OR*) ist definiert als das Chancenverhältnis zwischen zwei Ereignissen *Z* und *Y*:

$$OR = \frac{\text{Chance}_Z}{\text{Chance}_Y}$$

Formel XXVII

Daraus ergibt sich:

$$OR = \frac{\text{Chance}_{\text{Nichtübereinstimmung}}}{\text{Chance}_{\text{Übereinstimmung}}}$$

Formel XXVIII

Ihr Wert gibt an wie stark der untersuchte Risikofaktor die Chance eine Nichtübereinstimmung durch die Beurteiler zu erzielen, erhöht oder verringert. Es gilt somit:

$OR < 1$:	Die Chance einer Nichtübereinstimmung ist kleiner als die Chance einer Übereinstimmung
$OR = 1$:	Die Chance einer Nichtübereinstimmung ist genauso hoch wie die Chance einer Übereinstimmung
$OR > 1$:	Die Chance einer Nichtübereinstimmung ist größer als die Chance einer Übereinstimmung

Im in dieser Arbeit angewendeten Regressionsmodell wurde die *OR* als entlogarithmierter Logit-Koeffizient ($Exp(B)$) angegeben. Dieser errechnet sich wie folgt:

$$Exp(B) = OR = e^{-Logit}$$

Formel XXIX

Er gibt also an, um wie viel sich die logarithmierte Chance erhöht oder verringert eine Nichtübereinstimmung zu erzielen, wenn man die unabhängige Variable um eine Einheit erhöht.

Mit Hilfe der *OR* lässt sich nun die Wahrscheinlichkeit für das Auftreten einer Nichtübereinstimmung berechnen:

$$Wahrscheinlichkeit_{Nichtübereinstimmung} = \frac{1}{1 + e^{-Logit}}$$

Formel XXX

Die errechneten Wahrscheinlichkeiten bilden somit Vorhersagewerte der logistischen Regression.

Selektive Vorwärtsselektion

Mittels Univarianzanalyse wurde zunächst untersucht welcher der Risikofaktoren (unabhängige Variablen) einen signifikanten Einfluss ($p < 0,05$) auf das Interobserver Agreement hat. Im Anschluss wurden die unabhängigen Variablen mit $p < 0,1$ gemeinsam zur weiteren multivariablen Analyse verwendet (manuelle Vorwärtsselektion). Der Grenzwert von $p < 0,1$ wurde hierbei gewählt, um die Kriterien für den möglichen Einschluss in das Modell liberal zu halten.

4. Ergebnisse

4.1. Studie 1: Einfluss des Untersuchers auf das Interobserver Agreement bei der Evaluierung des Bizepssehnenreflexes

4.1.1. Probanden in Studie 1

Insgesamt wurden 32 Hunde in Studie 1 aufgenommen. Die Hunde hatten ein medianes Alter von 5 (0,5 - 15) Jahren und ein medianes Gewicht von 17,5 (5,8 – 57) kg. Fünfzehn Hunde (46,88%) waren kastriert, 17 (53,12%) intakt.

Eine Übersicht über das Signalement der Probanden bietet Tabelle 12.

Tabelle 12. Untersuchte Hunde in den Studien 1 und 2a: grün = Vordergliedmaßen, von beiden Untersuchern untersucht; rot = Vordergliedmaßen nur von Untersucher 1 untersucht; blau = Vordergliedmaßen nur von Untersucher 2 untersucht; orange = von der Beurteilung ausgeschlossen (s. Text).

fortlaufende Nr.	Aufzeichnungsnummer		Geschlecht	Rasse	Alter (a)	Gewicht (kg)	Felllänge
	r	l					
1	85	21	w	Mischling	0,50	16,00	kurz
2	44	47	wk	Bayerischer Gebirgsschweißhund	2,00	19,00	kurz
3	73	37	wk	Marshall Beagle	4,00	10,00	kurz
4	--	33	wk	Mischling	3,00	8,00	lang
5	3	74	wk	Mischling	2,00	16,00	kurz
6	58	71	wk	Dalmatiner	6,00	26,50	kurz
7	41	45	wk	Schnauzer	2,00	18,00	lang
8	68	67	wk	Mischling	6,00	17,00	kurz
9	98	25	m	Golden Retriever	9,00	39,00	lang
10	1	34	m	Mischling	12,00	18,50	lang
11	82	22	m	Mischling	7,00	30,00	kurz
12	89	65	w	Mischling	2,00	27,50	kurz
13	91	54	w	Labrador	2,50	25,50	kurz
14	69	32	wk	Mischling	1,25	10,00	lang
15	42	13	w	Mischling	8,50	23,00	lang
16	87	18	w	Labradoodle	4,50	23,00	lang
17	94	70	wk	Malinois	7,50	24,70	kurz
18	4	86	w	Französische Bulldogge	4,50	10,60	kurz
19	36	96	m	Schweizer Sennenhund	7,00	57,00	kurz
20	90	48	wk	Mischling	6,00	11,80	kurz
21	39	77	m	Labrador	1,50	28,00	kurz

fortlaufende Nr.	Aufzeichnungsnummer		Geschlecht	Rasse	Alter (a)	Gewicht (kg)	Felllänge
	r	l					
22	31	27	mk	Australian Shepherd	1,00	12,00	kurz
23	72	60	mk	Australian Shepherd	2,50	15,00	lang
24	100	2	wk	Australian Cattle Dog	5,50	20,00	kurz
25	51	23	m	Boston Terrier	14,00	6,40	kurz
26	10	84	m	Berner Sennenhund	6,75	40,50	lang
27	99	66	wk	Dackel	12,25	14,75	kurz
28	11	97	m	Deutsch Kurzhaar	2,50	40,00	kurz
29	59	63	m	Rauhaardackel	12,00	7,35	kurz
30	81	93	m	Yorkshire Terrier	4,50	5,80	lang
31	95	83	wk	Australian Terrier	11,25	10,00	lang
32	24	49	w	Mischling	6,00	8,40	lang

4.1.2. Kategoriale Unterteilung der Probanden

Die kategoriale Unterteilung der Probanden ist in Tabelle 13 aufgelistet.

Bei einem Hund wurde zweieinhalb Jahre vor der Untersuchung die rechte Vorderextremität amputiert, so dass der Bizepssehnenreflex an 63 Vordergliedmaßen untersucht wurde. Bei 45 Vordergliedmaßen (71,43%) fand eine Untersuchung durch beide Untersucher statt, bei 18 (28,57%) nur durch jeweils einen der beiden Untersucher, was in einer Gesamtzahl von 108 Untersuchungen resultiert.

4.1.3. Digitale Videobearbeitung

Nach Sichtung des Filmmaterials wurden die Untersuchungen beider Untersucher der rechten Vorderextremität von Hund Nr. 21, sowie die Untersuchung der rechten Vorderextremität von Hund Nr. 6 durch Untersucher 2 aufgrund technischer Probleme oder starker Unruhe des Hundes von der Beurteilung ausgeschlossen. Aus denselben Gründen wurde auch die Untersuchung durch Untersucher 2 an den linken Vordergliedmaßen der Hunde Nr. 8 und Nr. 19 von dem Beurteilungsprozess ausgeschlossen.

Beurteilt wurden somit 103 Videosequenzen an 61 Gliedmaßen (s. Tabelle 13). Es handelte sich um 30 (49,18%) rechte und 31 (50,82%) linke Vorderextremitäten.

Der Film für Studie 1 hatte eine Gesamtlänge von 34:09 Minuten (928.948 KB). Die durchschnittliche Sequenzdauer betrug für Untersucher 1 12,66 und für Untersucher 2 13,71 Sekunden.

Tabelle 13. Kategoriale Unterteilung der in Studie 1 untersuchten Vordergliedmaßen.

Kategorie	Anzahl Gliedermaßen	%	Anzahl Sequenzen
Geschlecht			
männlich	23	37,70	38
weiblich	38	62,30	65
Total:	61	100,00	103
Alter			
jung	29	47,54	56
alt	32	52,46	47
Total:	61	100,00	103
Gewicht			
leicht	36	59,02	59
schwer	25	40,98	44
Total:	61	100,00	103
Felllänge			
kurz	38	62,30	64
lang	23	37,70	39
Total:	61	100,00	103
Körperseite			
rechts	30	49,18	50
links	31	50,82	53
Total:	61	100,00	103
Untersucher			
Untersucher 1	-	-	47
Untersucher 2			56
Total:			103

4.1.4. Beurteilerübereinstimmung

Interobserver Agreement der Reflexpräsenzanalyse

Jeder der für die Reflexpräsenz untersuchten Reliabilitätskoeffizienten in Studie 1 ist umso höher je höher das Expertenlevel der Beurteilergruppe ist (s. Tabelle 14). Dieser Unterschied ist zwischen allen drei Gruppen für $K_F Präs$ signifikant. Die Koeffizienten $K_F Präs$ und $\bar{X}K_C$ ergeben für Gruppe 1 eine *ordentliche* und für die Gruppen 2 und 3 eine *geringe* Übereinstimmung.

Die K_C -Werte sind über den Gruppen signifikant unterschiedlich verteilt ($p = 0,007$). Gruppe 1 zeigt gegenüber den anderen beiden Gruppen eine signifikant unterschiedliche Verteilung der K_C -Koeffizienten (s. Abbildung 6 und Tabelle 15). Die mittlere Differenz zwischen $K_{max Präs}$ und K_C beträgt für Neurologen und Studenten jeweils 0,00 und für praktizierende Tierärzte 0,25 (s. Tabelle 14). Der höchste K_C -Wert (1 = *perfekte* Übereinstimmung; KI95%: 1-1) wird von dem Beurteilerpaar T1-S1, der niedrigste von dem Beurteilerpaar N3-S3 (0,1 = *kaum* Übereinstimmung; KI95%: -0,214-0,607) erzielt. Der mediane K_C -Wert aller Beurteilerkombinationen beträgt 0,47 (s. Anhang, Tabelle 34).

Interobserver Agreement der Reflexaktivitätsanalyse

Eine Übersicht über die Ergebnisse der Reflexaktivitätsanalyse aus Studie 1 bietet Tabelle 16.

Für Gruppe 3 ergibt sich die signifikant niedrigste Übereinstimmung der beiden Reliabilitätskoeffizienten $K_F Akt$ und ICC . *Fleiss' Kappa* unterscheidet sich hinsichtlich seines Übereinstimmungsmaßes zwischen Neurologen und praktizierenden Tierärzten nicht signifikant. Beide weisen eine *geringe*, Studenten *kaum* Übereinstimmung auf. Der ICC wird für Studenten als *schlecht*, für praktizierende Tierärzte als *zweifelhaft* und für Neurologen als *mittelmäßig* interpretiert.

Das durchschnittliche K_w ergibt für Neurologen und praktizierende Tierärzte eine *mittelmäßige* und für Studenten eine *geringe* Übereinstimmung. Das *durchschnittliche* K_w ist zwischen der Gruppe der Neurologen und der Gruppe der Studenten signifikant unterschiedlich verteilt (s. Abbildung 7 und Tabelle 17). Die durchschnittliche Differenz zwischen $K_{max Akt}$ und K_w beträgt für Neurologen 0,12, für praktizierende Tierärzte 0,07 und für Studenten 0,10 (s. Tabelle 16).

Der höchste K_w -Wert (0,61 = *ordentlich*; KI95%: 0,463-0,757) wird für das Beurteilerpaar T1-S1 berechnet, der niedrigste für das Beurteilerpaar T2-S4 (0,08 = *kaum* Übereinstimmung; KI95%: -0,093-0,253). Der mediane K_w -Wert aller Beurteilerkombinationen beträgt 0,39 (s. Anhang, Tabelle 35).

Die Anzahl der Beurteilungen mit *vollständiger Übereinstimmung* ist umso höher, je höher das Level der Beurteilerexpertise ist. Umgekehrt verhält es sich mit den *partiellen* und *vollständigen (Nicht-)Übereinstimmungen*: die Anzahl ist hier umso niedriger, je höher das Level der Beurteilerexpertise ist (s. Abbildung 8).

Tabelle 14. Ergebnisübersicht der Reliabilitätsanalyse der Reflexpräsenzevaluierung aus Studie 1 mit zugehörigen Interpretationsparametern.
 * = signifikant (< 0,05).

	r%Präs	$\bar{X}r\%$ Präs	K _C	KA	PI	BI	K _{max} Präs	$\bar{X}K_C$	K _F Präs	SF	KI95%	
											unten	oben
Gruppe 1												
N1-N4	98,06	97,90	0,66	I	0,94	0,02	0,66	0,71	0,70*	0,040	0,621	0,779
N1-N3	99,03		0,80	I	0,95	0,01	0,80					
N1-N2	96,12		0,49	I	0,92	0,04	0,49					
N2-N4	98,06		0,79	I	0,90	0,02	0,79					
N2-N3	97,09		0,65	I	0,91	0,03	0,65					
N3-N4	99,03		0,85	I	0,93	0,01	0,85					
Gruppe 2												
T1-T4	94,18	94,50	0,23	III	0,92	0,04	0,49	0,40	0,40*	0,040	0,326	0,483
T1-T3	96,12		0,49	I	0,92	0,04	0,49					
T1-T2	96,12		0,49	I	0,92	0,04	0,49					
T2-T4	92,23		0,29	III	0,88	0,00	1,00					
T2-T3	96,12		0,65	I	0,88	0,00	1,00					
T3-T4	92,23		0,29	III	0,88	0,00	1,00					
Gruppe 3												
S1-S4	98,06	92,07	0,49	I	0,96	0,00	0,49	0,33	0,30*	0,040	0,226	0,383
S1-S3	91,26		0,28	III	0,87	0,09	0,28					
S1-S2	92,23		0,31	III	0,88	0,08	0,31					
S2-S4	92,23		0,31	III	0,88	0,08	0,31					
S2-S3	87,38		0,31	III	0,80	0,01	0,31					
S3-S4	91,26		0,28	III	0,87	0,09	0,28					

Abbildung 6. Boxplot der K_C -Werte für die jeweils sechs Beurteilerkombinationen der 3 Gruppen aus Studie 1. Die Strecke über den Boxplots markiert den signifikanten (*) Unterschied zwischen den jeweiligen Gruppen (s. auch Tabelle 15).

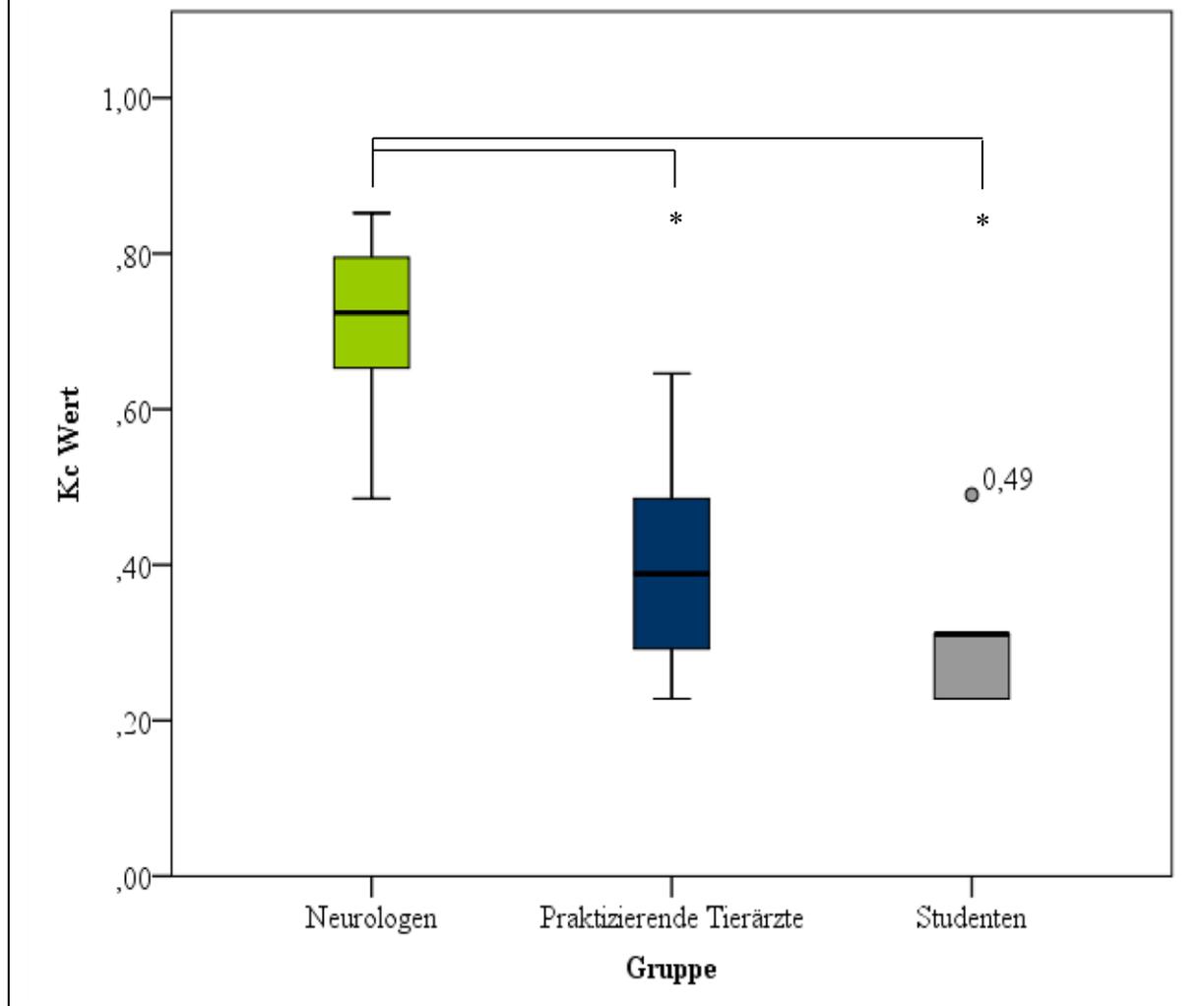


Tabelle 15. Einfaktorielle ANOVA nach Kruskal-Wallis der K_C -Mittelwerte aus Studie 1. Paarweiser Vergleich. *=signifikant ($p < 0,05$).

Gruppenvergleich	p-Wert
Neurologen vs. Praktizierende Tierärzte	0,032*
Neurologen vs. Studenten	0,014*
Praktizierende Tierärzte vs. Studenten	1,000

Tabelle 16. Ergebnisübersicht der Reliabilitätsanalyse der Reflexaktivitätsevaluierung aus Studie 1 mit zugehörigen Interpretationsparametern.
 * = signifikant (< 0,05).

	r% _{Akt}	$\bar{X}r\%_{Akt}$	K _w	KA	PI	BI	K _{max} _{Akt}	$\bar{X}K_w$	K _{F Akt}	SF	KI95%		ICC	KI95%	
											unten	oben		unten	oben
Gruppe 1															
N1-N4	75,73	80,91	0,26	III	0,69	0,02	0,72	0,44	0,33	0,031	0,270	0,392	0,84	0,782	0,855
N1-N3	82,52		0,43	I	0,70	0,01	0,89								
N1-N2	79,61		0,41	I	0,67	0,04	0,86								
N2-N4	82,52		0,58	I	0,69	0,02	0,72								
N2-N3	85,44		0,55	I	0,66	0,03	0,84								
N3-N4	79,61		0,41	I	0,68	0,01	1,00								
Gruppe 2															
T1-T4	63,11	61,17	0,39	II	0,26	0,06	0,84	0,36	0,33	0,029	0,271	0,383	0,80	0,734	0,859
T1-T3	69,90		0,50	II	0,12	0,09	0,70								
T1-T2	62,14		0,44	II	0,16	0,05	0,92								
T2-T4	53,40		0,30	II	0,21	0,11	0,81								
T2-T3	66,02		0,30	II	0,07	0,04	0,72								
T3-T4	52,43		0,28	II	0,17	0,15	0,75								
Gruppe 3															
S1-S4	64,08	54,70	0,25	II	0,46	0,16	0,49	0,25	0,17*	0,025	0,121	0,220	0,67*	0,553	0,763
S1-S3	51,46		0,32	II	0,12	0,18	0,66								
S1-S2	53,40		0,33	II	0,13	0,17	0,81								
S2-S4	47,57		0,14	II	0,28	0,33	0,44								
S2-S3	47,57		0,33	II	0,06	0,01	0,80								
S3-S4	64,08		0,13	II	0,27	0,34	0,31								

Abbildung 7. Boxplot der K_w -Werte für die jeweils sechs Beurteilerkombinationen der 3 Gruppen aus Studie 1. Die Strecke über den Boxplots markiert den signifikanten (*) Unterschied zwischen den jeweiligen Gruppen (s. auch Tabelle 17).

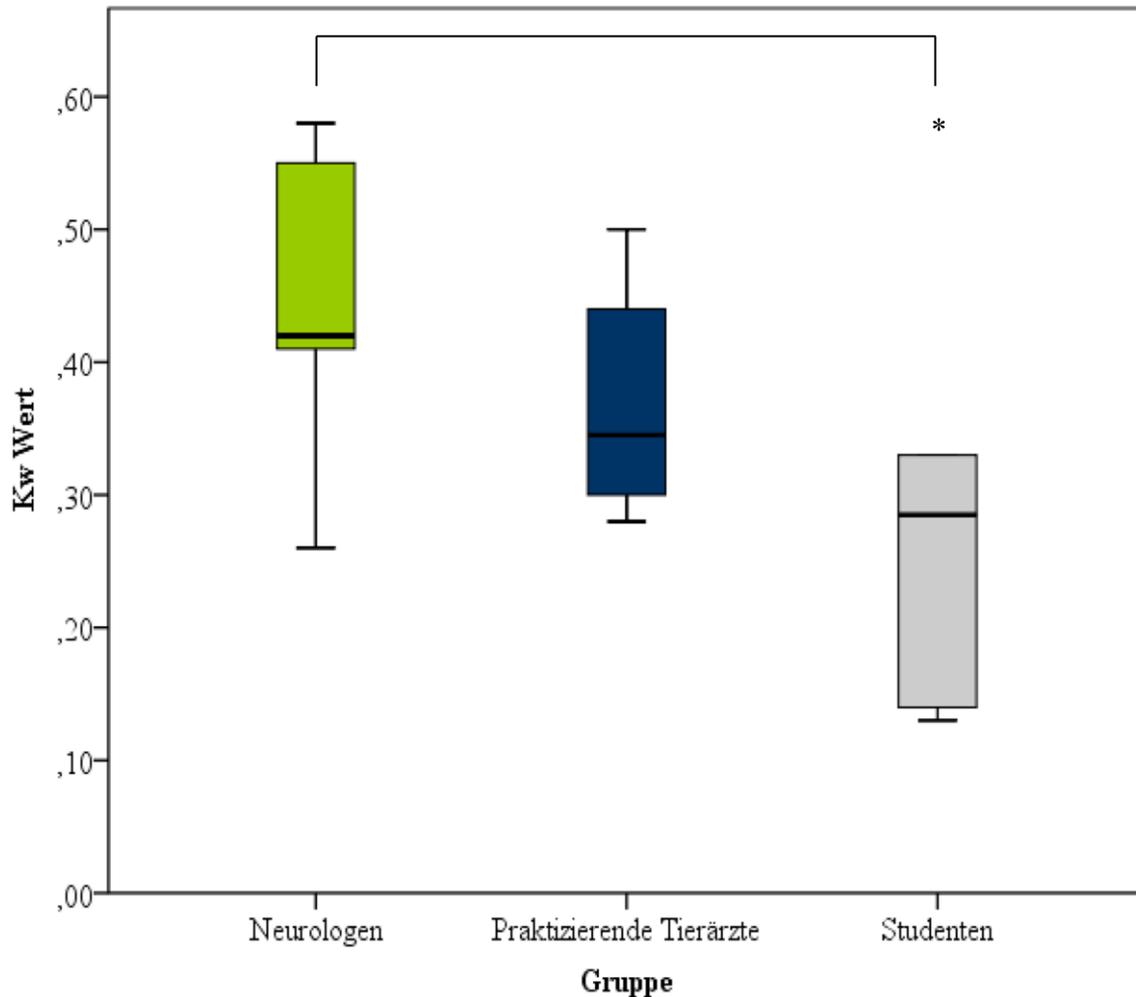
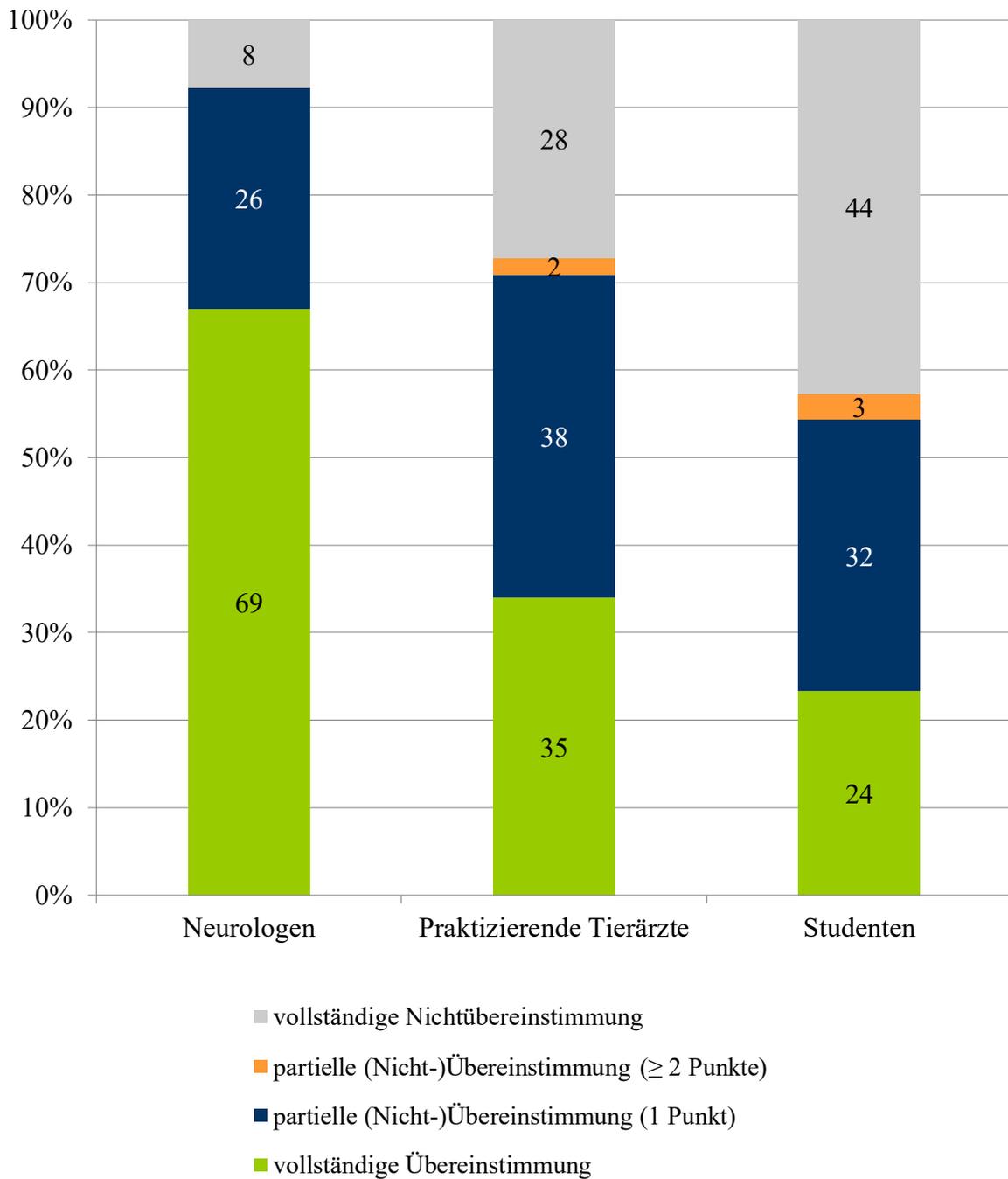


Tabelle 17. Einfaktorielle ANOVA nach Kruskal-Wallis der K_w -Mittelwerte aus Studie 1. Paarweiser Vergleich. *= signifikant ($p < 0,05$).

Gruppenvergleich	p-Wert
Neurologen vs. Praktizierende Tierärzte	0,989
Neurologen vs. Studenten	0,044*
Praktizierende Tierärzte vs. Studenten	0,431

Abbildung 8. *Table of Agreement* der Reflexaktivitätsanalyse in Studie 1. Die Anzahl der einzelnen Übereinstimmungsstufen ist angegeben.



4.1.5. Regressionsanalyse

Sowohl bei der Reflexpräsenz- als auch bei der –aktivitätsanalyse ist der einzige Risikofaktor, der einen signifikanten Einfluss auf das Interobserver Agreement der Reflexbeurteilung nimmt, die Beurteilerexpertise (s. Tabelle 18).

Bei der Beurteilung der Reflexpräsenz ist die Chance keine oder zumindest partiell keine Übereinstimmung zu erzielen bei Tierärzten 3,8-fach, bei Studenten 5,7-fach höher als bei einer Beurteilung durch Neurologen.

Bei der Beurteilung der Reflexaktivität wird der Unterschied zwischen den verschiedenen Beurteilungsgruppen deutlich: die Chance einer (partiellen) (Nicht-)Übereinstimmung ist bei Tierärzten 4,3-fach und bei Studenten 7,6-fach höher als bei Neurologen.

Der Risikofaktor *Untersucher* hat keinen signifikanten Einfluss auf die Chance einer Nichtübereinstimmung.

Tabelle 18. Ergebnisse der Regressionsanalyse aus Studie 1. *=signifikant

Risikofaktor	Reflexpräsenz			Reflexaktivität				
	Odds Ratio (OR)	KI95% für OR		p-Wert	Odds Ratio (OR)	KI95% für OR		p-Wert
		unten	oben			unten	oben	
Untersucher								
Untersucher 1	Referenzkategorie	.	.		Referenzkategorie	.	.	
Untersucher 2	1,022	0,380	2,748	0,965	1,515	0,902	2,545	0,116
Körperseite								
rechts	Referenzkategorie	.	.		Referenzkategorie	.	.	
links	1,429	0,524	3,895	0,484	0,903	0,527	1,547	0,710
Geschlecht								
männlich	Referenzkategorie	.	.		1,107	0,646	1,897	0,710
weiblich	1,168	-0,083	1,138	0,755	Referenzkategorie	.	.	
Felllänge								
kurz	Referenzkategorie				Referenzkategorie			
lang	1,432	0,528	3,884	0,480	1,64	0,958	2,810	0,710
Gewicht								
leicht	Referenzkategorie	.	.		Referenzkategorie	.	.	
schwer	0,925	0,337	2,539	0,879	1,491	-0,134	0,933	0,142
Alter								
jung	Referenzkategorie	.	.		Referenzkategorie	.	.	
alt	1,378	0,512	3,711	0,525	1,305	0,776	2,194	0,314
Beurteiler								
Neurologen	Referenzkategorie	.	.		Referenzkategorie	.	.	
Prakt. Tierärzte	3,818	1,104	13,208	0,034*	4,352	2,383	7,949	<0,001*
Studenten	5,662	1,689	18,979	0,005*	7,635	4,029	14,468	<0,001*

4.2. Studie 2: Vergleich des Interobserver Agreement zwischen dem Bizeps- und dem Patellarsehnenreflex

4.2.1. Studie 2a: Reliabilitätsanalyse des Bizepssehnenreflexes

4.2.1.1. Probanden in Studie 2a

Das mediane Alter der in dieser Studie inkludierten 30 Hunde lag bei 5,75 (1 – 14) Jahren und einem medianen Gewicht von 17,5 (5,8 – 57) kg. Fünfzehn (50%) Hunde waren kastriert, 15 (50%) intakt. Eine Übersicht über das Signalement der inkludierten Hunde bietet Tabelle 12.

4.2.1.2. Kategoriale Unterteilung der Probanden

Die kategoriale Unterteilung der in Studie 2a inkludierten Hunde ist Tabelle 19 zu entnehmen.

4.2.1.3. Digitale Videobearbeitung

Für Studie 2a wurde kein separater Film angefertigt, sondern die Sequenzen, die von Untersucher 2 durchgeführt wurden (n = 56, s. Tabelle 12), separat ausgewertet. Die durchschnittliche Dauer der 56 Sequenzen betrug 13,71 (8,57 - 34,63) Sekunden.

Tabelle 19. Kategoriale Unterteilung der in Studie 2a untersuchten Vordergliedmaßen.

Kategorie	Anzahl Gliedermaßen	%
Geschlecht		
männlich	23	41,07
weiblich	33	58,93
Total:	56	100,00
Alter		
jung	27	46,43
alt	29	53,57
Total:	56	100,00
Gewicht		
leicht	32	60,71
schwer	24	39,29
Total:	56	100,00
Felllänge		
kurz	33	71,88
lang	23	28,12
Total:	56	100,00
Körperseite		
rechts	27	48,21
links	29	51,79
Total:	56	100,00

4.2.1.4. Beurteilerübereinstimmung

Interobserver Agreement der Reflexpräsenzanalyse

Die beiden Reliabilitätskoeffizienten K_F *Präs* und $\bar{X}K_C$ sind umso höher, je höher das Expertenlevel der Beurteilergruppe ist (s. Tabelle 20). *Fleiss' Kappa* ist für Gruppe 3

signifikant am niedrigsten. Keine Signifikanz besteht zwischen Neurologen und praktizierenden Tierärzten. Sowohl $K_F Präs$, als auch $\bar{X}K_C$ liegen bei diesen beiden Gruppen im *ordentlichen* und bei Studenten im *mittelmäßigen* Bereich.

Die Verteilung der K_C -Werte ist über den Gruppen gleich, weist jedoch einen tendenziellen Unterschied ($p = 0,067$) auf (s. Abbildung 9). Die durchschnittliche Differenz zwischen $K_{max Präs}$ und K_C beträgt für Neurologen 0,06, für praktizierende Tierärzte 0,00 und für Studenten 0,16 (s. Tabelle 20). Der höchste K_C -Wert (1 = *perfekte* Übereinstimmung; KI95%: 1-1) wird von den Beurteilerpaaren N2-T2 und T1-S1, der niedrigste von dem Beurteilerpaar N3-S3 (0,20 = *kaum*; KI95%: -0,144-0,348) erzielt. Der mediane K_C -Wert aller Beurteilerkombinationen beträgt 0,56 (s. Anhang, Tabelle 39).

Interobserver Agreement der Reflexaktivitätsanalyse

Bei der Reflexaktivitätsanalyse in Studie 2a unterscheidet sich $K_F Akt$ signifikant zwischen den drei Gruppen und ist für die Neurologen signifikant am höchsten (0,49 = *mittelmäßig*) und für die Gruppe der Studenten signifikant am niedrigsten (0,24 = *gering*) (s. Tabelle 21). Der ICC ist für Neurologen ebenfalls signifikant am höchsten (0,87 = *mittelmäßig*) und liegt für die beiden anderen Gruppen jeweils im *zweifelhaften* Bereich (0,76).

Die Verteilung der K_w -Werte ist über den drei Gruppen gleich, weist jedoch einen tendenziellen Unterschied ($p = 0,054$) zwischen den Gruppen auf (s. Abbildung 10). Die durchschnittliche Differenz zwischen $K_{max Akt}$ und K_w beträgt für Neurologen 0,15, für praktizierende Tierärzte 0,11 und für Studenten 0,10 (s. Tabelle 21). Der höchste K_w -Wert wird von dem Beurteilerpaar N2-N3 (0,77 = *ordentlich*; KI95%: 0,549 – 0,998), der niedrigste von S1-S3 (0,27 = *gering*; KI95%: 0,265 – 0,730) erzielt. Der mediane K_w -Wert aller Beurteilerkombinationen beträgt 0,49 (s. Anhang, Tabelle 40).

Die Anzahl der Evaluierungen mit *vollständiger Übereinstimmung* ist umso höher, je höher das Level der Beurteilerexpertise ist (s. Abbildung 11). Umgekehrt ist die Anzahl an diskordanten Evaluierungen umso höher, je niedriger das Level der Beurteilerexpertise ist.

Tabelle 20. Ergebnisübersicht der Reliabilitätsanalyse der Reflexpräsenzevaluierung aus Studie 2a mit den jeweiligen Interpretationsparametern.
* = signifikant.

	r%Präs	$\bar{X}r\%$ Präs	K _C	KA	PI	BI	K _{max} Präs	$\bar{X}K_C$	K _F Präs	SF	KI95%		ICC	KI95%	
											unten	oben		unten	oben
Gruppe 1															
N1-N3	98,21	97,62	0,79	I	0,91	0,00	0,98	0,76	0,77	0,077	0,614	0,916	0,91*	0,859	0,944
N1-N2	96,43		0,65	I	0,89	0,04	0,96								
N2-N3	98,21		0,85	I	0,88	0,02	0,98								
Gruppe 2															
T1-T3	92,86	95,24	0,47	I	0,86	0,07	0,47	0,63	0,64	0,077	0,490	0,792	0,85*	0,768	0,908
T1-T2	96,43		0,65	I	0,89	0,04	0,65								
T2-T3	96,43		0,78	I	0,82	0,04	0,78								
Gruppe 3															
S1-S3	94,64	91,67	0,55	I	0,88	0,02	0,55	0,47	0,45*	0,077	0,303	0,606	0,73*	0,576	0,832
S1-S2	91,10		0,41	I	0,84	0,09	0,41								
S2-S3	89,29		0,44	I	0,79	0,04	0,81								

Abbildung 9. Boxplot der K_C -Werte für die jeweils drei Beurteilerkombinationen der 3 Gruppen aus Studie 2a.

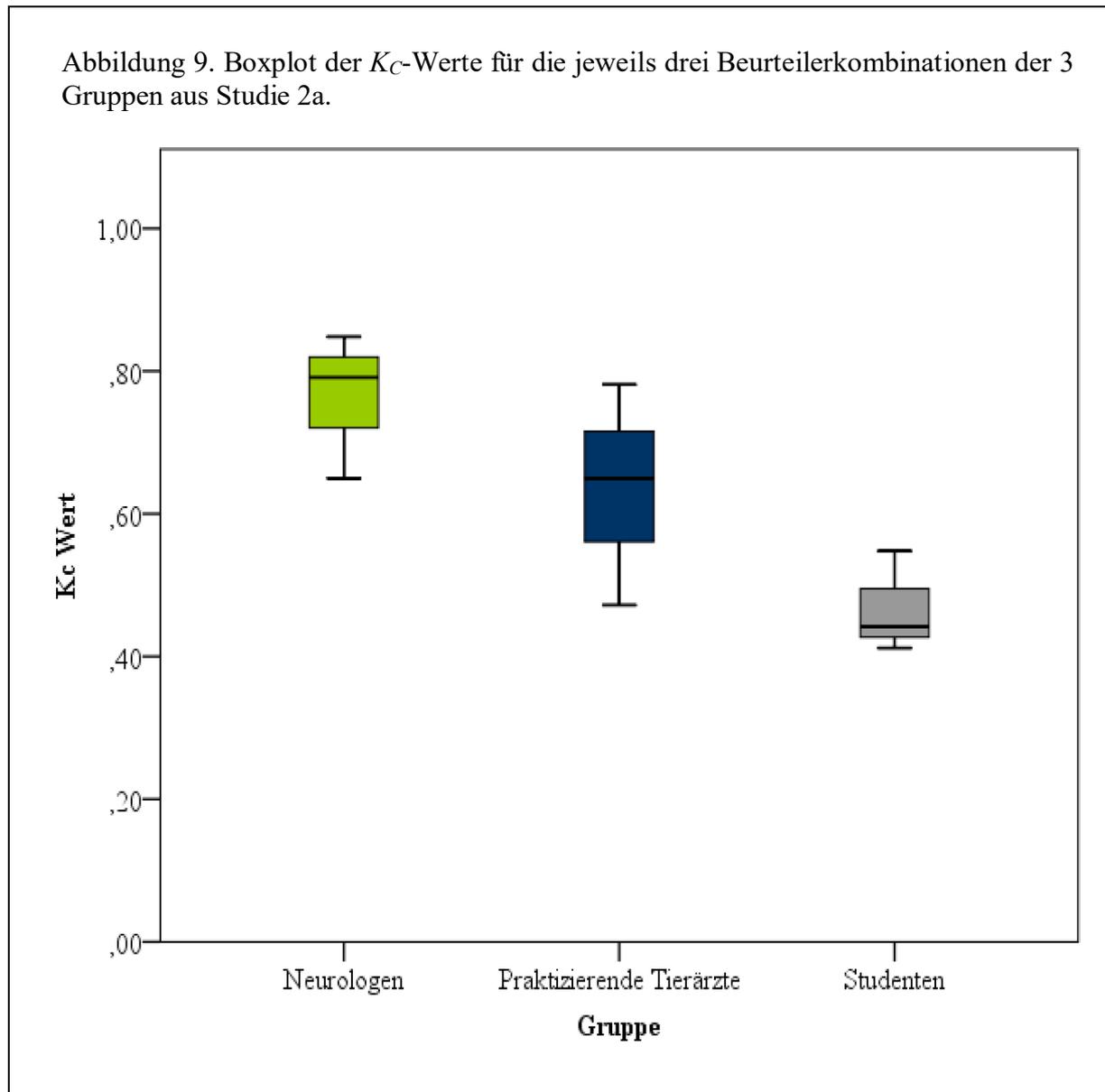


Tabelle 21. Ergebnisübersicht der Reliabilitätsanalyse der Reflexaktivitätsevaluierung aus Studie 2a mit den jeweiligen Interpretationsparametern.
 * = signifikant.

	r% _{Akt}	$\bar{X}r\%_{Akt}$	K _w	KA	PI	BI	K _{max} Akt	$\bar{X}K_w$	K _{F Akt}	SF	KI95%		ICC	KI95%	
											unten	oben		unten	oben
Gruppe 1															
N1-N3	80,36	81,55	0,55	I	0,31	0,04	0,85	0,57	0,49*	0,061	0,376	0,613	0,87*	0,795	0,918
N1-N2	75,00		0,43	I	0,31	0,04	0,90								
N2-N3	89,29		0,74	I	0,29	0,00	0,78								
Gruppe 2															
T1-T3	67,86	60,71	0,47	II	0,04	0,10	0,70	0,44	0,37*	0,053	0,268	0,474	0,76	0,630	0,853
T1-T2	55,36		0,41	II	0,04	0,11	0,91								
T2-T3	58,93		0,43	II	0,02	0,00	0,70								
Gruppe 3															
S1-S3	62,50	56,55	0,38	II	0,00	0,18	0,65	0,37	0,24*	0,048	0,149	0,337	0,76	0,622	0,850
S1-S2	55,36		0,33	II	0,00	0,18	0,81								
S2-S3	51,79		0,41	II	0,10	0,00	0,90								

Abbildung 10. Boxplot der K_w -Werte für die jeweils drei Beurteilerkombinationen der 3 Gruppen aus Studie 2a.

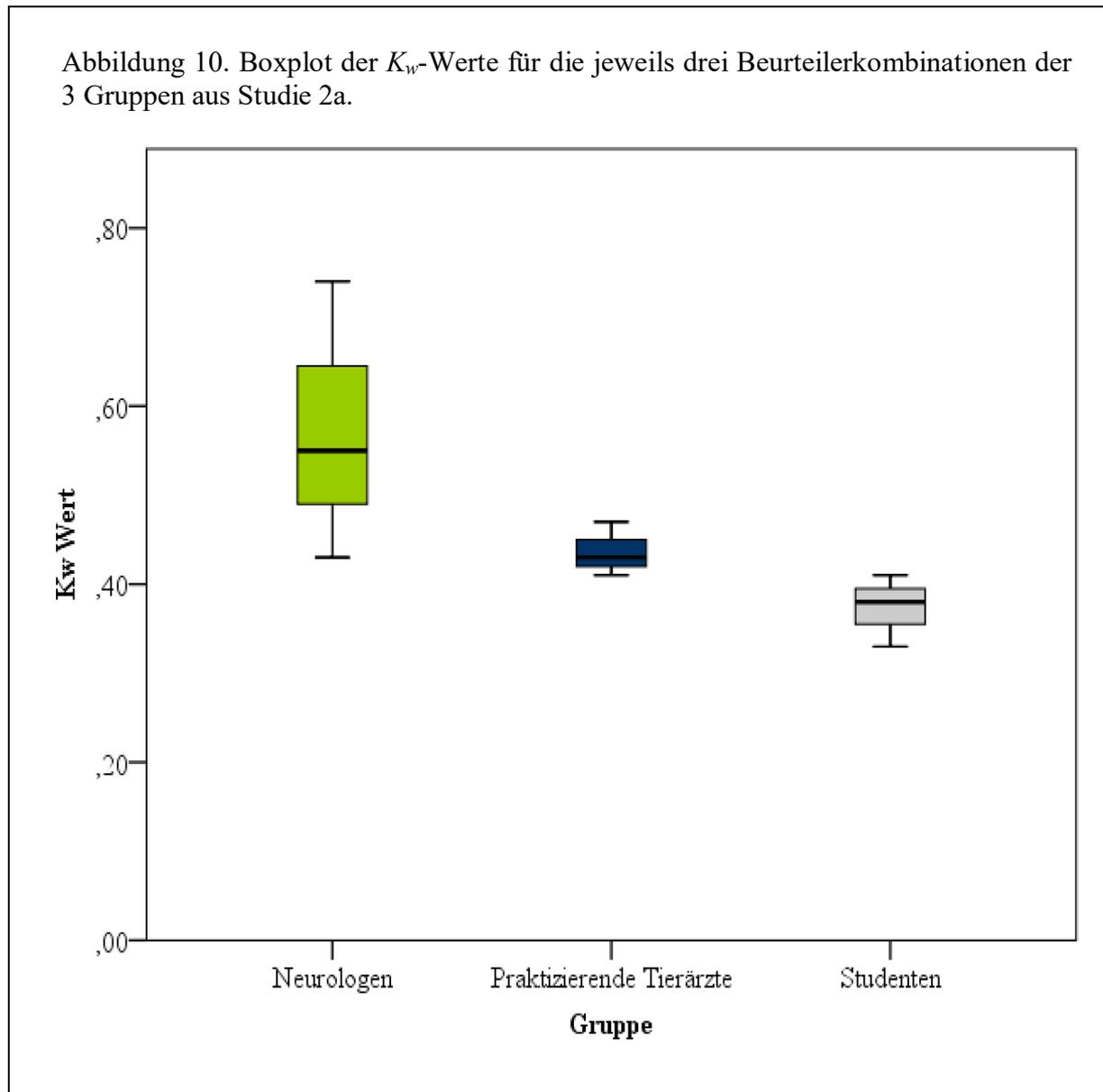
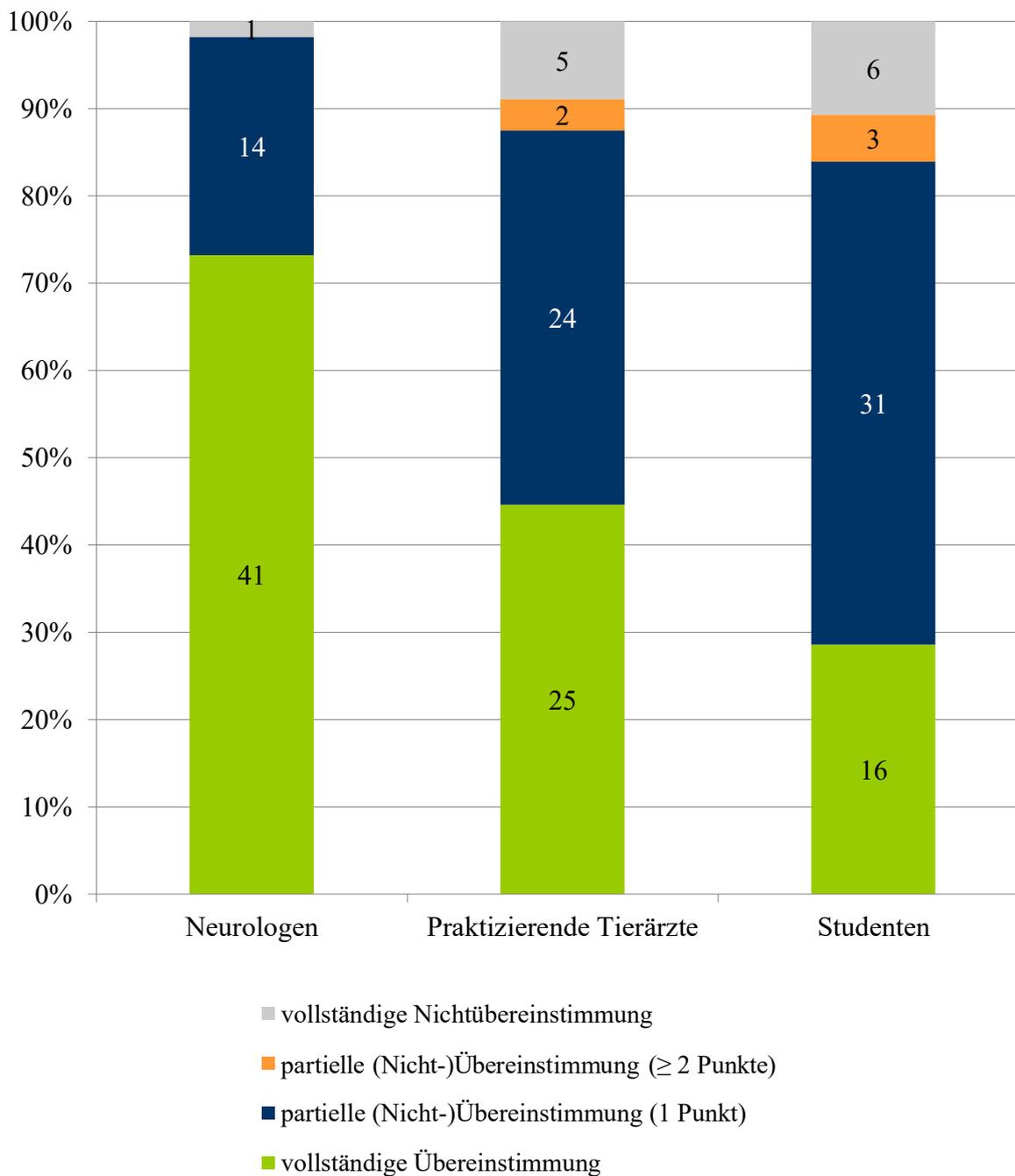


Abbildung 11. *Table of Agreement* der Reflexaktivitätsanalyse in Studie 2a. Die Anzahl der einzelnen Übereinstimmungsstufen ist angegeben.



4.2.1.5. Regressionsanalyse

Die Ergebnisse der Regressionsanalyse aus Studie 2a sind in Tabelle 22 und 23.

Das Ergebnis zeigt, dass keiner der für die Reflexpräsenz untersuchten Risikofaktoren in Studie 2a einen signifikanten Einfluss auf die Chance hat, nicht übereinstimmend zu beurteilen. Zwar ist die Chance einer Nichtübereinstimmung hier bei Studenten 4,3-mal höher als bei Neurologen, allerdings ist dieser Effekt nicht signifikant, zeigt jedoch eine statistische Tendenz ($p = 0,09$).

Bei der Regressionsanalyse der Reflexaktivität wird deutlich, dass sowohl die Felllänge, als auch die Beurteilerexpertise signifikant die Chance einer Nichtübereinstimmung erhöhen. So ist die Chance einer Nichtübereinstimmung bei Hunden mit langem Fell 2,6-mal höher verglichen mit kurzhaarigen Hunden ($p = 0,003$). Tierärzte haben eine 3,4-mal höhere ($p = 0,003$), Studenten eine fast 7-mal höhere Chance ($p < 0,001$) einer Nichtübereinstimmung gegenüber Neurologen.

Die schrittweise Regressionsanalyse ergibt, dass unter dem Einschluss der in Schritt 0 signifikanten Risikofaktoren *Felllänge* und *Beurteilerexpertise*, der signifikante Effekt der *Beurteilerexpertise* auch nach Korrektur für den Risikofaktor *Felllänge* bestehen bleibt (Schritt 1). Andersherum bleibt der Effekt der Felllänge signifikant, egal ob ein Neurologe, praktizierender Tierarzt oder ein Student die Reflexaktivität beurteilt (s. Tabelle 23).

Tabelle 22. Ergebnisse der Regressionsanalyse (Schritt 0) von Studie 2a. *=signifikant

Risikofaktor	Reflexpräsenz			Reflexaktivität				
	Odds Ratio (OR)	KI95% für OR		p-Wert	Odds Ratio (OR)	KI95% für OR		p-Wert
		unten	oben			unten	oben	
Körperseite								
rechts	Referenzkategorie	.	.		Referenzkategorie	.	.	
links	1,274	0,300	5,407	0,741	1,045	0,562	1,945	0,888
Geschlecht								
männlich	Referenzkategorie	.	.		Referenzkategorie	.	.	
weiblich	1,277	0,291	5,608	0,745	0,935	0,498	1,757	0,835
Felllänge								
kurz	Referenzkategorie	.	.		Referenzkategorie	.	.	
lang	1,382	0,327	5,838	0,659	2,646	1,393	5,023	0,003*
Gewicht								
leicht	Referenzkategorie	.	.		Referenzkategorie	.	.	
schwer	0,544	0,114	2,595	0,443	1,248	0,662	2,355	0,492
Alter								
jung	Referenzkategorie	.	.		Referenzkategorie	.	.	
alt	0,802	0,190	3,386	0,762	0,993	0,533	1,850	0,983
Beurteiler								
Neurologen	Referenzkategorie	.	.		Referenzkategorie	.	.	
Prakt. Tierärzte	1,575	0,234	10,591	0,639	3,434	1,540	7,659	0,003*
Studenten	4,337	0,795	23,654	0,090	6,975	3,015	16,132	<0,001*

Tabelle 23. Ergebnisse der Regressionsanalyse (Schritt 1) von Studie 2a. *=signifikant

Risikofaktor	Odds Ratio (OR)	Reflexaktivität		p-Wert
		unten	oben	
Felllänge				
kurz	Referenzkategorie	.	.	
lang	3,166	1,520	6,596	0,002*
Beurteiler				
Neurologen	Referenzkategorie	.	.	
Prakt. Tierärzte	3,736	1,620	8,623	0,002*
Studenten	7,922	3,288	19,085	<0,001*

4.2.2. Studie 2b: Reliabilitätsanalyse des Patellarsehnenreflexes

4.2.2.1. Probanden in Studie 2b

Der Patellarsehnenreflex wurde an 64 Hintergliedmaßen von 32 Hunden untersucht. Eine Übersicht über die in Studie 2b inkludierten Hunde bietet Tabelle 24. Die Hunde hatten ein medianes Alter von 6,38 (0,75 – 11) Jahren und ein medianes Gewicht von 25,5 (2 – 45) kg. Sechszehn (50%) Hunde waren kastriert, 16 (50%) waren intakt.

Alle untersuchten Gliedmaßen wurden der Beurteilung zugeführt.

Tabelle 24. Untersuchte Hunde in Studie 2b.

fortlaufende Nr.	Aufzeichnungsnummer		Geschlecht	Rasse	Alter (a)	Gewicht (kg)	Felllänge
	rechts	links					
1	4	2	wk	Border Collie	4,50	15,50	lang
2	51	50	w	Kurzhaarteckel	1,00	5,00	kurz
3	73	7	wk	Dalmatiner	7,25	26,20	kurz
4	48	67	wk	Mischling	4,00	16,00	kurz
5	92	42	mk	Mischling	11,00	34,00	kurz
6	75	32	w	Mischling	0,75	6,00	lang
7	29	11	w	Irish Setter	7,25	29,00	lang
8	9	35	m	Foxhound	8,00	18,00	kurz
9	19	25	wk	Deutscher Schäferhund	10,00	34,40	lang
10	84	17	w	Mischling	7,00	2,00	kurz
11	78	33	wk	Pudel	10,00	23,00	lang
12	31	96	m	Labrador	10,50	32,40	kurz
13	62	36	wk	Mischling	4,50	8,00	lang
14	63	24	wk	Deutscher Schäferhund	8,00	38,00	kurz
15	27	38	mk	Mischling	3,00	34,00	kurz
16	66	58	w	Deutscher Schäferhund	1,00	45,00	lang
17	54	3	mk	Labrador	5,50	34,00	kurz

fortlaufende Nr.	Aufzeichnungsnummer		Geschlecht	Rasse	Alter (a)	Gewicht (kg)	Felllänge
	rechts	links					
18	15	39	w	Mischling	2,50	13,00	kurz
19	70	100	mk	Mischling	6,70	29,00	lang
20	52	34	w	Labrador	2,00	27,00	kurz
21	57	87	w	Mischling	7,00	13,00	kurz
22	28	55	m	Dobermann	1,00	40,00	kurz
23	1	23	m	Mischling	2,50	16,00	kurz
24	68	74	w	Magyar Vizsla	3,00	19,00	kurz
25	41	93	w	Labrador	6,00	28,00	kurz
26	12	76	w	Fox-Terrier	9,00	9,00	kurz
27	14	72	mk	Mischling	11,00	4,50	lang
28	89	85	wk	Mischling	11,00	31,00	kurz
29	5	79	mk	Labrador	8,50	27,20	kurz
30	37	98	wk	Labrador	8,50	24,80	kurz
31	18	77	wk	Labrador	2,75	28,30	kurz
32	94	47	w	Labrador	1,00	24,20	kurz

4.2.2.2. Kategoriale Unterteilung der Probanden

Die kategoriale Unterteilung der untersuchten Gliedmaßen ist in Tabelle 25 aufgeführt.

Tabelle 25. Kategoriale Unterteilung der in Studie 2b untersuchten Hintergliedmaßen.

Kategorie	Anzahl Gliedermaßen	%
Geschlecht		
männlich	20	31,25
weiblich	44	68,75
Total:	64	100,00
Alter		
jung	32	50,00
alt	32	50,00
Total:	64	100,00
Gewicht		
leicht	26	40,63
schwer	38	59,37
Total:	64	100,00
Felllänge		
kurz	46	71,88
lang	18	28,12
Total:	64	100,00
Körperseite		
rechts	32	50,00
links	32	50,00
Total:	64	100,00

4.2.2.3. Digitale Videobearbeitung

Der Film für Studie 2b hatte eine Gesamtlänge von 11:42 Minuten (250.209 KB). Die durchschnittliche Länge der 64 Videosequenzen betrug nach digitaler Bearbeitung 7,76 (5,4-10,03) Sekunden.

4.2.2.4. Beurteilerübereinstimmung

Interobserver Agreement der Reflexpräsenzanalyse

Bei der Reflexpräsenzevaluierung in Studie 2b korreliert die Höhe der Reliabilitätskoeffizienten $K_{F\text{Präs}}$ und $\bar{X}K_C$ mit dem Level der Beurteilerexpertise. Hierbei ist $K_{F\text{Präs}}$ für die Neurologen signifikant am höchsten, zwischen praktizierenden Tierärzten und Studenten besteht kein signifikanter Unterschied. Die Übereinstimmung bei den Neurologen ist *mittelmäßig*, bei den beiden anderen Gruppen jeweils *gering* (s. Tabelle 26).

Die Verteilung von K_C ist über den drei Gruppen gleich ($p = 0,831$) (s. Abbildung 12). Die durchschnittliche Differenz zwischen $K_{\max\text{Präs}}$ und K_C beträgt für Neurologen 0,45, für praktizierende Tierärzte 0,00 und für Studenten 0,23. Drei Beurteilerpaare (N3-S1, N3-T2, T2-S1) erzielen eine *perfekte* Übereinstimmung. Den niedrigsten K_C -Wert (-0,02; KI95%: -0,0379-0,00612) erzielen vier Beurteilerpaare (N1-N3, N1-S1, N1-T1, N1-T2). Der mediane K_C -Wert aller Beurteilerkombinationen ist 0,24 (s. Anhang, Tabelle 44).

Interobserver Agreement der Reflexaktivitätsanalyse

Die Reliabilitätsanalyse der Reflexaktivitätsevaluierungen zeigt eine variierende Reihenfolge der Beurteilerübereinstimmung. So ist $\bar{X}r\% \text{Akt}$ und $K_{F\text{Akt}}$ jeweils umso höher, je höher das Expertenlevel der Beurteiler ist; $K_{F\text{Akt}}$ ist für Studenten signifikant am niedrigsten. Der ICC ist für praktizierende Tierärzte am höchsten und am niedrigsten für Studenten (s. Tabelle 27).

Die Übereinstimmung wird für $K_{F\text{Akt}}$ und den ICC für alle Gruppen als *mittelmäßig* interpretiert.

Die Verteilung von K_w ist über den Gruppen gleich ($p = 0,193$) (s. Abbildung 13). Die durchschnittliche Differenz zwischen $K_{\max\text{Akt}}$ und K_w beträgt für Neurologen 0,10 und für

praktizierende Tierärzte und Studenten jeweils 0,03 (s. Tabelle 27). Den höchsten K_w -Wert (0,80 = *ordentlich*) erzielt das Beurteilerpaar T2-T3, den niedrigsten das Beurteilerpaar T3-S2 (0,51 = *mittelmäßig*). Der mediane K_C -Wert aller Beurteilerkombinationen ist 0,62 (s. Anhang, Tabelle 45).

Die Anzahl der jeweiligen Übereinstimmungsstufen ist für die Gruppen *Neurologen* und *Praktizierende Tierärzte* identisch. In beiden Gruppen gibt es keine *teilweisen (Nicht-)Übereinstimmungen* (≥ 2 Punkte), bei 61% *vollständigen Übereinstimmungen* (s. Abbildung 14). Studenten haben die vergleichsweise geringste Anzahl an *vollständigen Übereinstimmungen* (42%) ($n = 27$). In 3% ($n = 2$) der Beurteilungen urteilen sie mit einer *partiellen (Nicht-)Übereinstimmung* (≥ 2 Punkte), in ca. 9% ($n = 6$) mit einer *vollständigen Nichtübereinstimmung* (s. Abbildung 14).

Tabelle 26. Ergebnisübersicht der Reliabilitätsanalyse der Reflexpräsenzevaluierung aus Studie 2b mit den jeweiligen Interpretationsparametern.
 * = signifikant.

	r%Präs	$\bar{X}r\%$ Präs	K _C	KA	PI	BI	K _{max} Präs	$\bar{X}K_C$	K _F Präs	SF	KI95%		ICC	KI95%	
											unten	oben		unten	oben
Gruppe 1															
N1-N3	96,88	97,92	-0,02	III	0,96	0,00	1,00	0,43	0,49*	0,072	0,348	0,631	0,74	0,612	0,836
N1-N2	98,44		0,66	I	0,95	0,02	0,66								
N2-N3	98,44		0,66	I	0,66	0,02	0,66								
Gruppe 2															
T1-T3	96,88	97,92	0,00	III	0,97	0,03	0,00	0,22	0,32	0,072	0,181	0,464	0,60	0,391	0,743
T1-T2	98,44		0,66	I	0,95	0,02	0,66								
T2-T3	98,44		0,00	III	0,98	0,02	0,00								
Gruppe 3															
S1-S3	93,75	89,58	0,32	III	0,91	0,06	0,32	0,25	0,23	0,072	0,088	0,371	0,50	0,246	0,682
S1-S2	89,06		0,20	III	0,86	0,12	0,20								
S2-S3	85,94		0,23	III	0,80	0,05	0,74								

Abbildung 12. Boxplot der K_C -Werte für die jeweils drei Beurteilerkombinationen der 3 Gruppen aus Studie 2b.

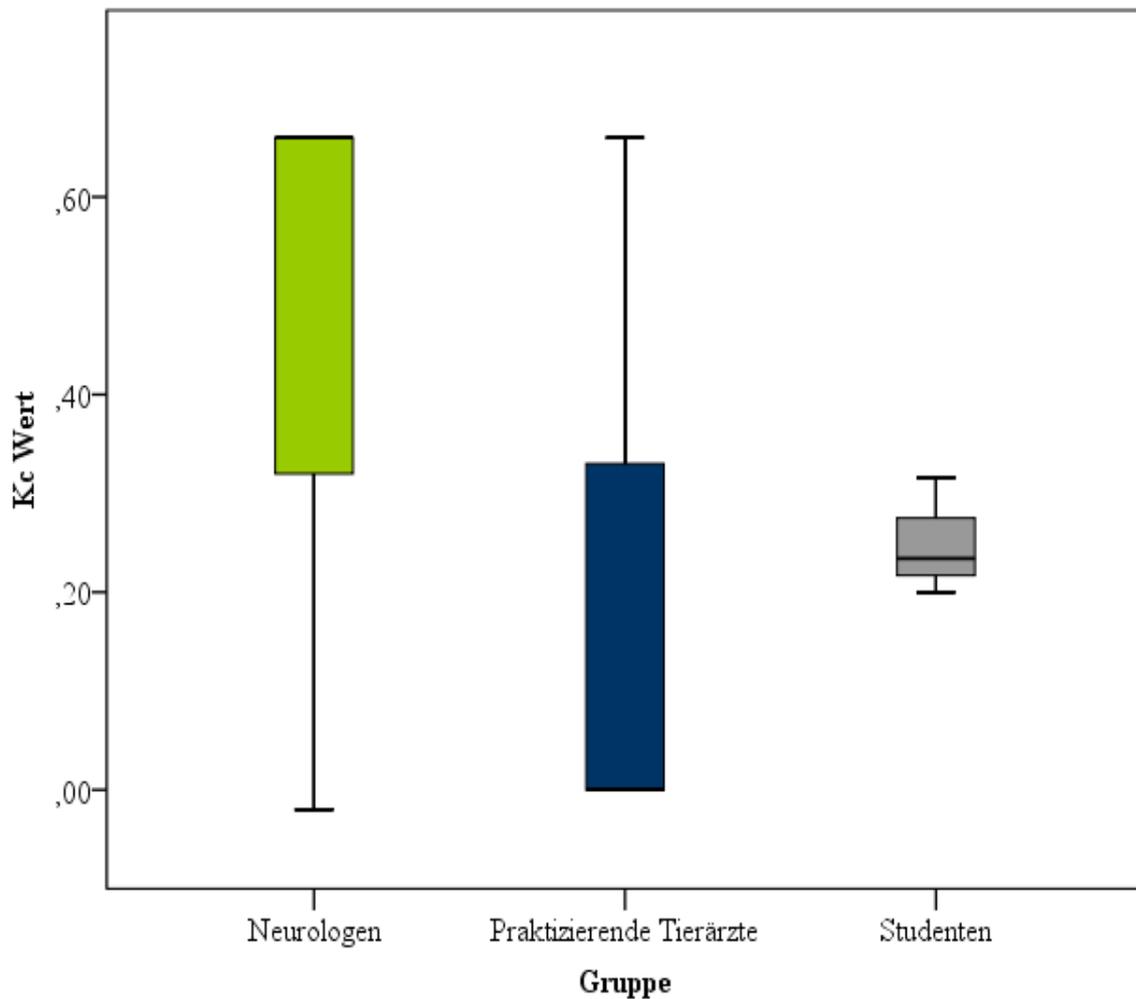


Tabelle 27. Ergebnisübersicht der Reliabilitätsanalyse der Reflexaktivitätsevaluierung aus Studie 2b mit den jeweiligen Interpretationsparametern.
 * = signifikant.

	r% _{Akt}	$\bar{X}r\%_{Akt}$	K _w	KA	PI	BI	K _{max} Akt	$\bar{X}K_w$	K _{F Akt}	SF	KI95%		ICC	KI95%	
											unten	oben		unten	oben
Gruppe 1															
N1-N3	68,75	73,96	0,46	II	0,14	0,09	0,85	0,56	0,49	0,057	0,376	0,601	0,85	0,768	0,902
N1-N2	84,38		0,72	I	0,22	0,05	0,87								
N2-N3	68,75		0,49	II	0,17	0,14	0,68								
Gruppe 2															
T1-T3	67,20	73,96	0,57	II	0,03	0,27	0,59	0,65	0,57	0,051	0,469	0,670	0,90	0,813	0,921
T1-T2	70,31		0,64	II	0,02	0,19	0,71								
T2-T3	84,38		0,75	I	0,15	0,08	0,86								
Gruppe 3															
S1-S3	62,50	58,33	0,53	II	0,12	0,09	0,89	0,54	0,43*	0,042	0,345	0,511	0,80	0,764	0,900
S1-S2	57,81		0,54	II	0,18	0,20	0,81								
S2-S3	54,69		0,56	II	0,24	0,11	0,86								

Abbildung 13. Boxplot der K_w -Werte für die jeweils drei Beurteilerkombinationen der 3 Gruppen aus Studie 2b.

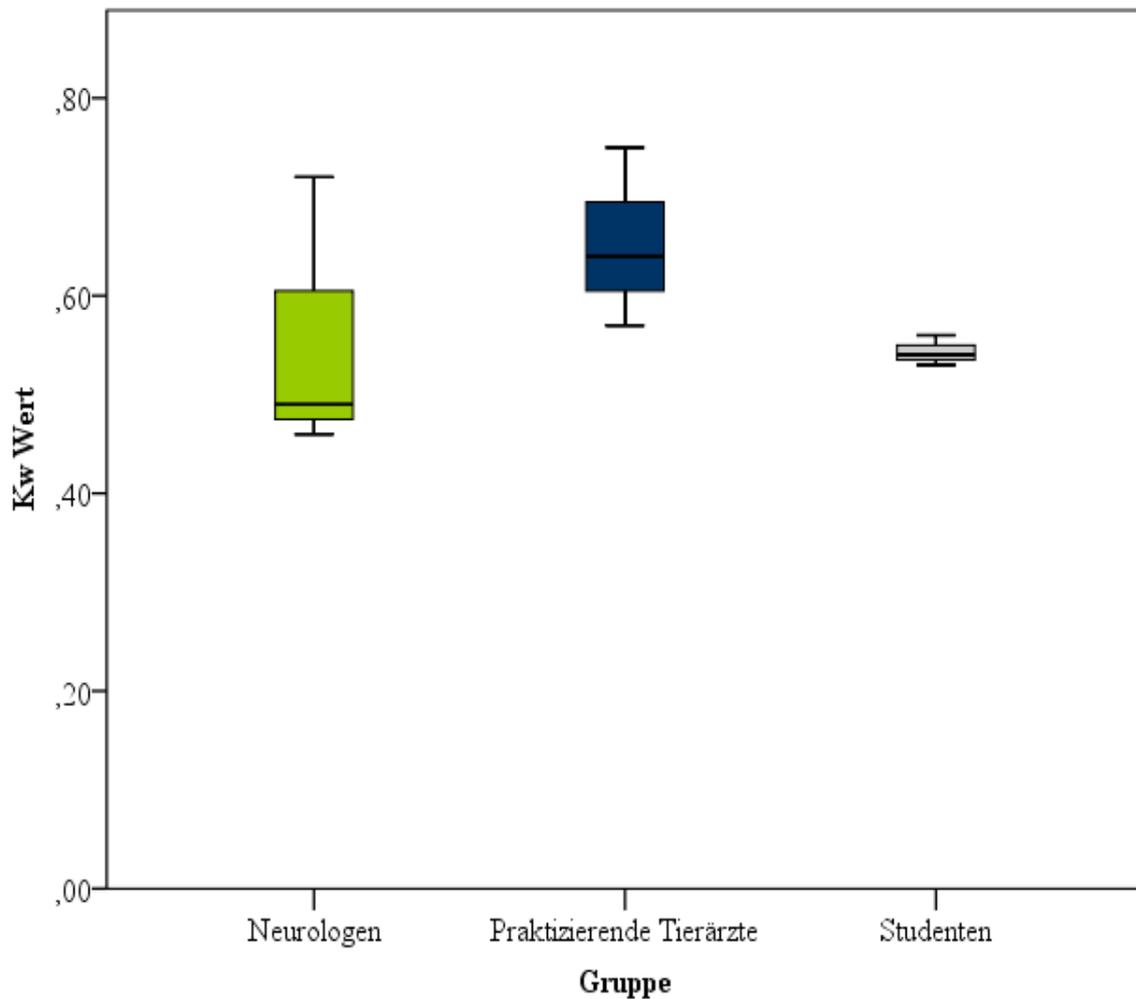
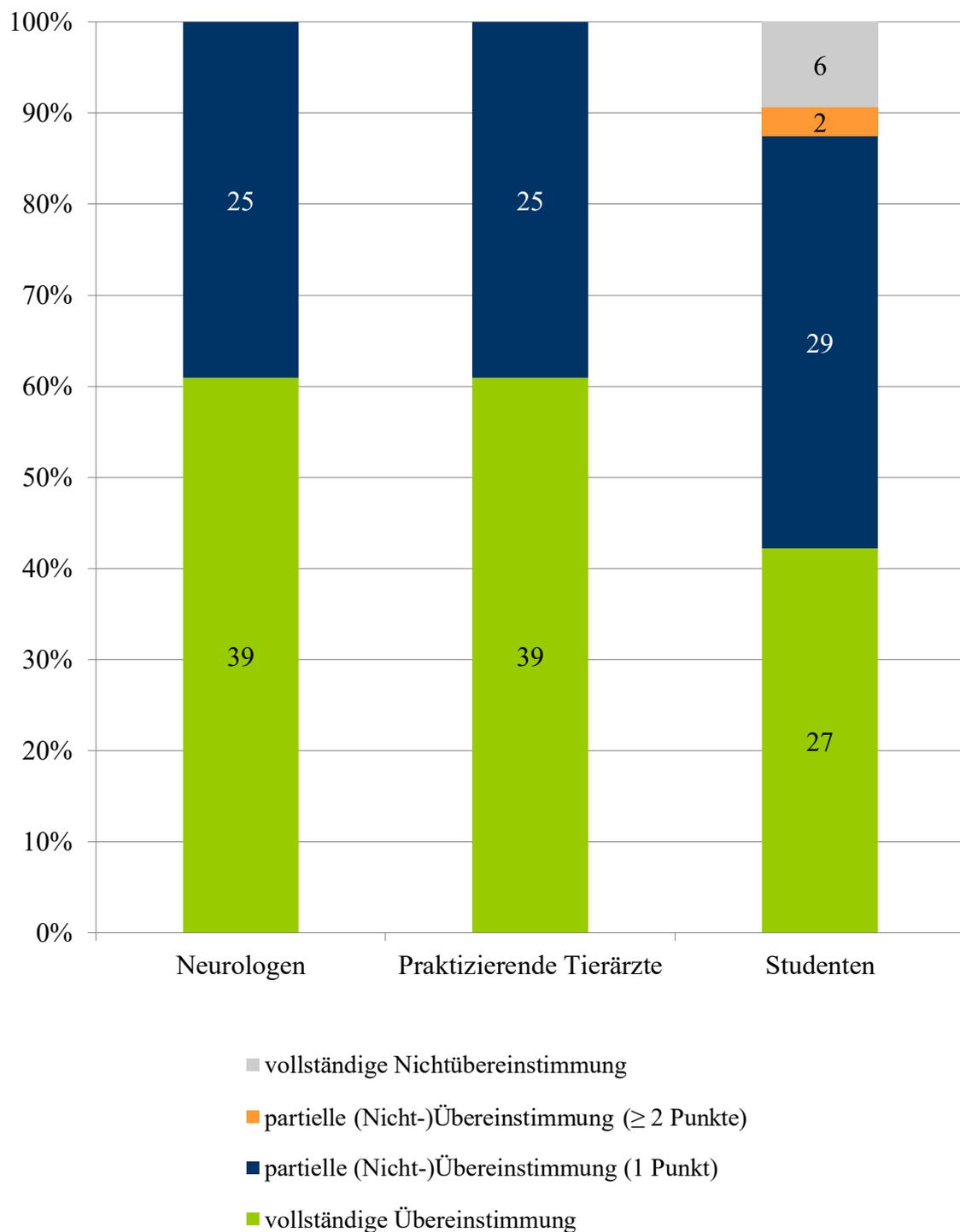


Abbildung 14. *Table of Agreement* der Reflexaktivitätsanalyse in Studie 2b. Die Anzahl der einzelnen Übereinstimmungsstufen ist angegeben.



4.2.2.5. Regressionsanalyse

Keiner der untersuchten Risikofaktoren hat einen signifikanten Einfluss auf die Chance einer Nichtübereinstimmung.

Bei der Reflexpräsenzevaluierung haben Studenten eine 3,7-fach höhere Chance einer Nichtübereinstimmung verglichen mit Neurologen, dieser Effekt gründet allerdings nur auf einer statistischen Tendenz ($p = 0,054$; s.

Bei der Reflexaktivität besitzen Studenten eine 4,3-mal höhere Chance einer Nichtübereinstimmung gegenüber Neurologen und Tierärzten, deren Chance einer Nichtübereinstimmung jeweils identisch ist. Der Unterschied zwischen Studenten und den beiden anderen Beurteilergruppen ist hochsignifikant ($p < 0,001$). Die Risikofaktoren *Körperseite* ($p = 0,063$) und *Felllänge* ($p = 0,093$) nehmen keinen signifikanten Einfluss auf die Chance einer Nichtübereinstimmung, weisen jedoch eine statistische Tendenz auf (s. Tabelle 28.)

Tabelle 28. Ergebnisse der Regressionsanalyse von Studie 2b. *= signifikant

Risikofaktor	Reflexpräsenz				Reflexaktivität			
	Odds Ratio (OR)	KI95% für OR		p-Wert	Odds Ratio (OR)	KI95% für OR		p-Wert
		unten	oben			unten	oben	
Körperseite								
rechts	Referenzkategorie	.	.		Referenzkategorie	.	.	
links	1,671	0,534	5,230	0,376	1,725	0,971	3,064	0,063
Geschlecht								
männlich	Referenzkategorie	.	.		Referenzkategorie	.	.	
weiblich	0,697	0,221	2,197	0,536	0,633	0,339	1,180	0,149
Felllänge								
kurz	Referenzkategorie	.	.		Referenzkategorie	.	.	
lang	0,887	0,248	3,165	0,852	0,577	0,304	1,098	0,093
Gewicht								
leicht	Referenzkategorie	.	.		Referenzkategorie	.	.	
schwer	1,664	0,497	5,578	0,406	1,129	0,627	2,035	0,684
Alter								
jung	Referenzkategorie	.	.		Referenzkategorie	.	.	
alt	1,292	0,420	3,972	0,653	0,747	0,419	1,329	0,319
Beurteiler								
Neurologen	Referenzkategorie	.	.		Referenzkategorie	.	.	
Prakt. Tierärzte	0,601	0,092	3,942	0,594	1,000	0,488	2,048	1,000
Studenten	3,761	0,976	14,497	0,054	4,351	2,045	9,260	<0,001*

5. Diskussion

5.1. Hypothese

In der vorgelegten Arbeit wird erstmals die Reliabilität der subjektiven Beurteilung des Bizeps- und Patellarsehnenreflexes beim Hund untersucht. Die zugrunde gelegte Methodik und die Interpretation vor dem Hintergrund verschiedener potentieller Einflussfaktoren liegen nach Kenntnis des Autors in diesem Ausmaß bislang in der Literatur nicht vor. Die zentrale Fragestellung dieser Arbeit ist, warum sich die Reflexbeurteilung hinsichtlich ihrer Aussagekraft zwischen den verschiedenen spinalen Reflexen unterscheidet bzw. ob und welche Risikofaktoren das Interobserver Agreement beeinflussen. Beim Studium der Literatur wird deutlich, dass die Beurteilung der Zuverlässigkeit eines bestimmten Spinalreflexes häufiger subjektiven Präferenzen unterliegt als wissenschaftlichen Untersuchungen. Einig scheint sich der Großteil der Autorenschaft über die Tatsache, dass der Flexor- und Patellarsehnenreflex die aussagekräftigsten Reflexe im Rahmen der neurologischen Untersuchung sind (De Lahunta und Glass, 2009b; Schatzberg, 2010; Schatzberg et al., 2012). Dennoch fehlt es bislang an einer evidenzbasierten Untersuchung dieser Hypothese.

5.2. Auswahl der zu untersuchenden Reflexe

Bezogen auf die diskutierte Reliabilität der spinalen Reflexe hätte man der Literatur folgend in dieser Arbeit ebenso die Evaluierung des Flexorreflexes der des Bizepssehnenreflexes gegenüberstellen können. Bei der Beurteilung des Flexorreflexes wird allerdings nicht nur die Beugung der Gliedmaße in ihren Gelenken, sondern auch die dabei aufgebrauchte Beugungskraft beurteilt, mit der die Gliedmaße zurückgezogen wird. Dies ist für den Betrachter eines Videos nicht nachvollziehbar. Nach Meinung des Autors muss die Beurteilung der Reflexantwort des Flexorreflexes mittels Videoanalyse stark störungsanfällig sein. Darüber hinaus handelt es sich beim Flexorreflex um einen polysynaptischen Fremdreflex.

Verschiedene Autoren nähren Zweifel an der vermeintlichen Zuverlässigkeit etablierter spinaler Reflexe. So zeigten Forterre et al. (2008), dass die Untersucher in ca. 30% der untersuchten Fälle die Läsion aufgrund eines reduzierten Flexorreflexes der Vorderextremität falsch in die Rückenmarkssegmente C₆-T₂ lokalisierten. Für die Reflexantwort des

Patellarsehnenreflexes konnte eine alters- (Chandrasekhar et al., 2013; Levine et al., 2002) und positionsbedingte (De Lahunta und Glass, 2009a) reduzierte Aktivität nachgewiesen werden. Einen Zweifel an der Sensitivität des Patellarsehnenreflexes stellten auch Abdelhakiem et al. (2015), die bei keinem der Hunde mit einem Bandscheibenvorfall innerhalb der Rückenmarkssegmente zwischen dem vierten Lenden- und ersten Sakralwirbel einen reduzierten Patellarsehnenreflex fanden und in 30% der Fälle jedoch aufgrund eines reduzierten Patellarsehnenreflexes fälschlicherweise die Diagnose einer Läsion innerhalb der Rückenmarkssegmente L₄-S₁ stellten. Murakami et al. (2014) beschrieben, dass eine Hyporeflexie des Flexorreflexes beim Hund in lediglich 37,5% und eine Hyporeflexie des Patellarsehnenreflexes sogar nur in 16,7% der untersuchten Fälle eine Läsion zwischen viertem Lenden- und erstem Sakralwirbel richtig lokalisiert. Für die häufig als leicht durchführ- und beurteilbar beschriebenen Muskelmassenreflexe des Musculus tibialis cranialis und Musculus extensor carpi radialis konnte von Tudury et al. (2017) nachgewiesen werden, dass diese auch noch nach lokaler Nervenblockade der verantwortlichen peripheren Nerven und sogar noch kurz nach Eintritt des Todes des Tieres (Tudury et al., 2013) auslösbar waren.

Mit dem Patellarsehnenreflex wurde dem Bizepssehnenreflex ein als zuverlässig zu evaluierender und ebenfalls monosynaptischer Muskelsehnenreflex gegenübergestellt.

5.3. Skizzierung des Versuchsaufbaus und Versuchsdurchführung

Es wurde ein Versuchsaufbau entworfen, der es erlaubt die Unterschiede in der Reflexbeurteilung unter klinischen Bedingungen zu untersuchen (Murakami et al., 2014). Die Reflexuntersuchung in dieser Arbeit bildet die klinischen Bedingungen gut ab und die Ergebnisse untermauern ein vieldiskutiertes Problem in der Medizin: die Reliabilität zweier subjektiv erhobener Beurteilungen. Im klinischen Alltag ist es häufig so, dass zur Einordnung bestimmter Merkmale die subjektive Beurteilung als klinischer Parameter herangezogen wird, auf dessen Basis Diagnose und Prognose gestellt und der klinische Verlauf bewertet werden. In diversen medizinischen Bereichen wie zum Beispiel der Radiologie (Cales et al., 1993; Doubilet und Herman, 1981; Reetz et al., 2013; Robinson et al., 1999; Tudor et al., 1997), der Krankenpflege (Kottner et al., 2009a; Kottner et al., 2009b) oder der Kardiologie (Davies, 1958) wurde daher die Quantifizierung der Beurteilerübereinstimmung zur Einordnung ihrer Reliabilität durchgeführt.

Bei der Skizzierung des Versuchsaufbaus wurde sich an der bereits existierenden Literatur orientiert. Levine et al. (2002) beispielsweise legten in ihrer Studie einem verblindeten Beurteiler ein Video zur Evaluierung der Patellarsehnenreflexantwort vor, die er als *präsent* oder *abwesend* kategorisieren sollte. Stam und van Crevel (1990) ließen drei verblindete Neurologen die Reflexaktivität beim Menschen anhand eines Videos mittels einer 9-Punkte-Ordinalskala beurteilen.

Die Selektion der Hunde orientierte sich ebenfalls an bereits vorhandenen vergleichbaren Studien. Muguet-Chanoit et al. (2011) schlossen in ihrer Arbeit zur Definition des sensorischen Feldes und der Wiederholbarkeit des Musculus cutaneus trunci-Reflexes nur solche Hunde ein, die eine normale neurologische Untersuchung, einen normalen Hydrationsstatus und vorberichtlich keine Myelopathien aufwiesen. Auch Levine et al. (2002) wählten als Einschlusskriterien eine unauffällige allgemeine, neurologische oder orthopädische Untersuchung. Rohrbach et al. (2014) führten ihre Untersuchungen an 25 Schafen durch, die im Vorfeld sogar bereits innerhalb einer orthopädischen Studie verwendet wurden, zum Zeitpunkt der Testungen allerdings klinisch gesund waren.

Die Integrität der untersuchten Reflexbögen wurde in der vorliegenden Studie klinisch bestätigt, allerdings nicht radiologisch verifiziert. Studien, die die Reliabilität neurologischer Symptome untersuchten, verzichteten meist auf die Analyse eines Interobserver Agreement. Vielmehr stellten sie zur Einordnung der Beurteilungsgüte oder zur Untersuchung einer Korrelation zwischen subjektiver Beurteilung und objektiv messbarer Daten den subjektiv erhobenen Evaluierungen der Reflexantwort eine als Gold-Standard definierte Untersuchung wie zum Beispiel bildgebende (Abdelhakiem et al., 2015; Forterre et al., 2008; Garosi et al., 2001) oder elektrodiagnostische Verfahren (Dafkin et al., 2013; Stam und van Crevel, 1990) gegenüber. So untersuchte Forterre et al. (2008) zum Beispiel eine Korrelation zwischen subjektiver Untersuchung des Flexorreflexes beim Hund und objektiver Befundung der magnetresonanztomographischen Untersuchung der Halswirbelsäule.

Da in der hier vorgelegten Arbeit der Beurteilerevaluierung keine objektive Messgröße gegenübergestellt und die Evaluierung der Integrität des jeweiligen Reflexbogens ausschließlich auf der im Vorfeld durchgeführten neurologischen Untersuchung basiert, kann keine Aussage zu der Evaluierungsgüte innerhalb der Gruppe getroffen werden. Dafkin et al. (2013) verwendeten den Begriff „*accuracy*“ als Definition des Anteils übereinstimmender

Beurteilungen an der Gesamtzahl aller Beurteilungen. Dieser im Deutschen mit *Genauigkeit* oder *Exaktheit* zu übersetzende Terminus ist allerdings nach Meinung des Autors dieser Arbeit falsch, da die Genauigkeit eines medizinischen Indikators auch dadurch definiert ist, Abweichungen vom Physiologischen zu detektieren. Dafkin et al. (2013) untersuchten allerdings nur gesunde Probanden. Es ist wichtig zu unterstreichen, dass sich die Resultate in der vorgelegten Arbeit somit allein auf eine Aussage über die Beurteilerübereinstimmung beschränken.

Es war also nicht das Ziel der vorgelegten Arbeit, die Sensitivität, beziehungsweise Spezifität des Bizeps- und Patellarsehnenreflexes zu ermitteln, sondern das Interobserver Agreement dieser Reflexe zu erörtern beziehungsweise Unterschiede in der Beurteilerübereinstimmung zwischen diesen beiden Reflexen abzubilden und somit die Problematik der Reliabilität subjektiver Untersuchungsergebnisse im Rahmen der neurologischen Untersuchung zu veranschaulichen. Die Prävalenz-Indices aller drei Studien dieser Arbeit belegen, dass Neurologen die Videosequenzen der Reflexuntersuchung klinisch-neurologisch unauffälliger Hunde am häufigsten als *normal* kategorisieren. Dies lässt den Rückschluss zu, dass die Einschlusskriterien durch den Autor dieser Arbeit im Vorfeld korrekt evaluiert wurden.

Ursprünglich sah der Versuchsaufbau dieser Arbeit vor, in Anlehnung an die Publikation von Dafkin et al. (2013) und Stam und van Crevel (1990), die mittels Oberflächenelektroden ermittelte maximale Reflexamplitude und -dauer den ordinalskalierten Beurteilungen der Beurteiler gegenüberzustellen. Dem Autor wurde dafür ein eigens entwickelter piezoelektrischer Reflexhammer (Fa. Sigma Medizin Technik) zur Verfügung gestellt, der Zeitpunkt und Stärke des Impulses detektiert. Dieser Versuchsaufbau erwies sich in Vorversuchen allerdings als nicht durchführbar. Es war unmöglich, am nichtsedierten Hund reproduzierbare Untersuchungsergebnisse zu erzielen. Schon die für eine fehlerfreie Messung notwendige ruhige Grundlinie des Elektromyogramms zu erhalten, war nicht möglich, da sich diese bereits durch minimale Bewegungen oder Zittern des Hundes veränderte. Eine Sedation hätte nicht zu im klinischen Alltag reproduzierbaren Ergebnissen geführt, da die neurologische Untersuchung routinemäßig am unsedierten Tier durchgeführt wird und anzunehmen ist, dass die Reflexaktivität hierdurch herabgesetzt wird. Darüber hinaus ist die Nutzung von Oberflächenelektroden stark störungsanfällig (Hodges und Gandevia, 2000; Stam und van Crevel, 1989). Rohrbach et al. (2009, 2014) untersuchten den Flexorreflex am unsedierten Schaf und Pferd mit Elektrostimulation und Oberflächenelektroden und erzielten verwendbare

Ergebnisse. Aufgrund der Größe der Muskelbäuche (Musculus deltoideus beim Schaf, Musculus digitalis communis und Musculus tibialis cranialis beim Pferd) ist eine zentrale Platzierung der Oberflächendetektorelektroden auf der Zielmuskulatur eventuell besser möglich als beim Hund. Eine artifizielle Detektion von Kontraktionen angrenzender Muskulatur ist daher beim Großtier weniger wahrscheinlich als beim Hund. Darüber hinaus wurden die Untersuchungen am stehenden Tier in gewohnter Umgebung durchgeführt und die Probanden mussten nicht wie zur Untersuchung des Bizepssehnen- und Patellarsehnenreflexes beim Hund in Seitenlage verbracht werden. Andere Studien, die die Elektromyographie-gestützte Muskelaktivität am unsedierten Probanden maßen, konzentrierten sich auf die Analyse vordefinierter Bewegungsabläufe (Cullen et al., 2015; Cullen et al., 2017) und nicht auf die Quantifizierung der Reflexantwort per se.

5.4. Beurteilung der Reflexantwort anhand eines Videos

Bei der Versuchsdurchführung wurde zum einen Wert darauf gelegt, alle Sequenzen möglichst gleich und standardisiert anzufertigen, wobei sich an Lehrvideos auf *YouTube*®^{2,3} orientiert wurde. Zum anderen wurde die klinisch neurologische Untersuchung so realistisch wie möglich abgebildet.

Andere Studien haben zuvor bereits die Analyse videobasierter Evaluierungen zur Bestimmung des Interobserver Agreement genutzt (Borin et al., 2012; Carswell et al., 2012; Dafkin et al., 2013; Dafkin et al., 2016; Essex et al., 2003; Olby et al., 2014; Packer et al., 2015). Der Grad der Standardisierung variiert allerdings erheblich zwischen den Studien: während die meisten der genannten humanmedizinischen Studien ihren Beurteilern standardisierte Videos zur Beurteilung vorlegten, beschränkte sich die Studie von Packer et al. (2015) beispielsweise auf die Beurteilung nichtstandardisierter, von Hundebesitzern auf *YouTube*® zur Verfügung gestellter Videos. Nach Meinung des Autors ist die Standardisierung einer Untersuchung unter klinischen Bedingungen in der Veterinärmedizin schwieriger als in der Humanmedizin, da es

² <https://www.youtube.com/watch?v=days0qHUpGI> (abgerufen am 28.05.2018)

³ <https://www.youtube.com/watch?v=BqlvfBDloXw> (abgerufen am 28.05.2018)

nicht möglich ist den veterinärmedizinischen Probanden den Versuchsaufbau zu erklären und man sie nicht bitten kann, bestimmte Positionen einzunehmen oder sich kooperativ zu verhalten. Das unterschiedlich hohe Interobserver Agreement, das von den Autoren jeweils präsentiert wurde, war jedoch nicht ausschließlich auf den unterschiedlich zugrunde gelegten Standard der Videos zurückzuführen, sondern auch auf die unterschiedliche Anzahl der Beurteiler und Kategorien der Beurteilungsskala (vgl. Tabelle 1). Zudem unterstrichen Dafkin et al. (2016) die Tatsache, dass sich die Analyse eines standardisierten Videos nicht mit den alltäglich klinischen Gegebenheiten deckt und es eventuell sogar eher zu einer falsch-hohen Reliabilität kommen könne. Inwieweit die Videoqualität in dieser Arbeit das Interobserver Agreement trotz Standardisierung beeinflusst haben könnte, wurde nicht untersucht.

Die Telemedizin ist ein relativ modernes Gebiet der Medizin. Bereiche, in denen die Telemedizin bisher vorwiegend genutzt wird, sind die Radiologie (Johnston und Worrall, 2003), die Pathologie (Chordia et al., 2016; Farahani und Pantanowitz, 2015; Higgins, 2015; Houghton et al., 2014; Neil und Demetris, 2014; Rampy und Glassy, 2015) und auch die Neurologie (Wechsler, 2015; Wechsler et al., 2017; Yager et al., 2014). Der Telemedizin wird in einer zunehmend vernetzten und globalisierten Welt ein hoher Stellenwert beigemessen, was sich nicht zuletzt in der Tatsache ausdrückt, dass es schon eigene Fachzeitschriften für diesen Bereich gibt^{4,5}. Wesentliche Vorteile sind sicherlich die generelle, globale (Johnston und Worrall, 2003; Wechsler, 2015) und rasche (Yager et al., 2012) Verfügbarkeit von Spezialisten zum Beispiel bei akuten Schlaganfallpatienten (Johnston und Worrall, 2003).

Um eine adäquate Vergleichbarkeit zwischen den Evaluierungen zu garantieren, muss gewährleistet werden, dass die Voraussetzungen zur Beurteilung bei allen Sequenzen und für jeden Beurteiler gleich sind (Wechsler et al., 2017). Yager et al. (2014) kamen zu dem Ergebnis, dass ausgewählte neurologische Symptome zuverlässig mittels Teleneurologie beurteilt werden können. Davis et al. (2014) etablierten erfolgreich ein teleneurologisches Modell zur Verlaufskontrolle neurologischer Patienten. Nach Meinung des Autors profitiert der Kliniker im Rahmen der veterinärneurologischen Untersuchung wesentlich von der meist durch die

⁴ <https://www.hindawi.com/journals/ijta/> (abgerufen am 28.05.2018)

⁵ <http://journals.sagepub.com/home/jtt> (abgerufen am 28.05.2018)

Besitzer angefertigter Videodokumentation der Symptomatik ihres Tieres. Packer et al. (2015) fanden bei der videobasierten Evaluierung von epileptischen Anfällen beim Hund nur eine geringe Beurteilerübereinstimmung. Die Videos waren allerdings nicht standardisierte, von Besitzern auf *YouTube*© zur Verfügung gestellte Videos. Die in der vorliegenden Arbeit präsentierten Ergebnisse belegen, dass die Reflexantwort des Bizeps- und Patellarsehnenreflexes -in Abhängigkeit der Beurteilerexpertise- beim gesunden Hund zuverlässig anhand eines Videos beurteilt werden kann. Bei zukünftigen videobasierten Reliabilitätsanalysen wäre es sinnvoll eine Befragung der Beurteiler durchzuführen, wann eine Beurteilung schwer fiel (zum Beispiel „Hecheln des Hundes“, „schwarzes Fell“), um einer artifiziellen Fehlinterpretation der Symptome durch Standardisierung vorzubeugen. So ist es beispielsweise der subjektive Eindruck des Autors, dass bei hechelnden Hunden die Detektion der Reflexantwort des Bizepssehnenreflexes schwerer fiel.

Yager et al. (2012) entwickelten ein Modell, in dem ein Intensivmedizin-Oberarzt auch in der Nacht via Telemedizin verfügbar war und dokumentierten in drei Fällen eine durch den Oberarzt angeleitete Stabilisierung eines Intensivpatienten via Telefonkonferenz. Der in der vorliegenden Arbeit skizzierte Versuchsaufbau kann als Modell eines standardisierten videobasierten neurologischen Untersuchungsganges fungieren, um den Nutzen beziehungsweise die Reliabilität der Teleneurologie auch für die Veterinärmedizin zu untersuchen. Eigenen Erfahrungen folgend, kann die Interpretation neurologischer Symptome im Rahmen der klinischen Untersuchung schwierig sein. Die Anwendung teleneurologischer Konsultationen mit erfahrenen Neurologen in kritischen Situationen ist möglicher Weise für den veterinärmedizinischen Patienten sinnvoll.

5.5. Auswahl und Limitationen der verwendeten Reliabilitätskoeffizienten

Die in dieser Arbeit präsentierten Ergebnisse veranschaulichen die Fehleranfälligkeit der Reliabilitätskoeffizienten und unterstreichen die Notwendigkeit ihrer korrekten Interpretation. Bei der Präsentation der Ergebnisse wird somit Wert darauf gelegt, sie durch Anführung empfohlener Interpretationsparameter in den korrekten Kontext zu stellen (Feinstein und Cicchetti, 1990; Flight und Julious, 2015).

Die meisten Studien, die das Interobserver Agreement der Reflexevaluierung thematisierten, führten nur einen Reliabilitätskoeffizienten an (s. Tabelle 1). Die Präsentation verschiedener Koeffizienten ermöglicht es jedoch einerseits einen Trend abzubilden und zeigt ferner die Interpretationslimitationen der einzelnen Koeffizienten auf, warum sich der Autor dieser Arbeit dazu entschieden hat, verschiedene Koeffizienten in die Reliabilitätsanalyse einzubeziehen.

Obwohl für den *ICC* verschiedene Interpretationsmodelle existieren (vgl. Altman, 1991; Benito et al., 2017; Chinn, 1991; LeBreton und Senter, 2008; Lewis et al., 2012; Portney und Watkins, 2015; Vincent und Weir, 1999), wurde in dieser Arbeit ein strenges Modell gewählt (Vincent und Weir, 1999). Es ist somit zu erwarten, dass die *ICC*-Werte in dieser Arbeit unter der Verwendung liberalerer Modelle wesentlich besser interpretiert worden wären. Ein häufig angewendetes Modell zur Interpretation des *ICC* ist das von Altman (1991) wonach ein *ICC*-Wert von 1 - 0,81 als *sehr gut*, 0,80 - 0,61 als *gut*, 0,60 - 0,41 als *mittelmäßig*, 0,40 - 0,21 als *ausreichend* und Werte von < 0,20 als *schlecht* zu interpretieren ist. Zum Vergleich: in der vorliegenden Arbeit werden alle *ICC*-Werte < 0,70 als *schlecht* bewertet. Die *ICC*-Werte, die in dieser Arbeit somit als *zweifelhaft* interpretiert werden, wären Altman (1991) folgend, somit also zumindest als eine *gute* Übereinstimmung bewertet worden. Dies hebt hervor, dass sich die verschiedenen Studien hinsichtlich der Interpretation ihrer Reliabilitätsanalyse gravierend unterscheiden können. Die Interpretation von Kappa hingegen folgt nahezu ausschließlich dem von Landis und Koch (1977) empfohlenen Schema und ist somit gut zwischen den verschiedenen Studien im Hinblick auf seine Interpretation vergleichbar.

Das Paradoxon, dass Kappa mit steigender Prävalenz des untersuchten Merkmals sinkt, veranlasste Sargeant und Martin (1998) die Indices der durchschnittlichen positiven beziehungsweise negativen Übereinstimmungen zur Bestimmung der Reliabilität heranzuziehen. Andere Autoren präsentierten die K_w -Koeffizienten mit dem Verweis auf eine hohe Prävalenz des untersuchten Merkmals, nennen jedoch keine alternativen Koeffizienten beziehungsweise ergänzenden Interpretationsparameter (Rousing und Waiblinger, 2004). Die *Prozentuale Übereinstimmung* stellt in den Studien mit hohem Prävalenz-Index und dichotomer Beurteilung, wie bei der Reflexpräsenzanalyse in dieser Arbeit, den aussagekräftigsten Koeffizienten dar (Hansen et al., 1994), da eine Gewichtung hier nicht notwendig ist und $r\%_{Präs}$ nur die *vollständigen Übereinstimmungen* als prozentualen Anteil an allen Beurteilungen wiedergibt. *Cohen's Kappa* berücksichtigt zwar, dass die Übereinstimmungen auch teilweise rein zufällig hätten getroffen werden können, allerdings ist bekanntermaßen die Höhe von

Kappa durch die Prävalenz der untersuchten Kategorie beschränkt (Hansen et al., 1994), so dass es notwendig wird, Interpretationsparameter für eine solidere Interpretation der Ergebnisse heranzuziehen (Feinstein und Cicchetti, 1990; Flight und Julious, 2015). Für Manschot et al. (1998) liegt Kappa in klinischen Untersuchungen meist zwischen 0,40 - 0,70, wobei Werte zwischen 0,60 - 0,80 selten sind und eine *perfekte* Übereinstimmung kaum möglich ist. Die Ergebnisse dieser Arbeit zeigen, dass Kappa in den allermeisten Fällen in den drei Studien im Bereich zwischen 0,40 und 0,70 liegt, Werte $> 0,70$ werden nur selten erzielt und beschränken sich ausschließlich auf die Gruppen der Neurologen ($n = 7/24$) und praktizierenden Tierärzte ($n = 1/24$). Die Anzahl an Werten von $Kappa < 0,40$ ist umso niedriger, umso höher das Level der Beurteilerexpertise ist: Neurologen ($n = 2/24$), praktizierende Tierärzte ($n = 9/24$) und Studenten ($n = 16/24$).

Die von Sim und Wright (2005) vorgeschlagene Definition der *klinischen Akzeptanz* für Werte $K \geq 0,40$ schlugen auch Burn und Weir (2011) für Studien in der Veterinärmedizin vor. Die Reliabilitätsanalyse zeigt, dass die Kappa-Analyse umso häufiger als *klinisch akzeptabel* bewertet wird, je höher die Expertise der Beurteilergruppe ist. Interessanterweise muss Kappa bei der Reflexaktivitätsanalyse in den Studien 1 und 2a für die Gruppen der praktizierenden Tierärzte und Studenten in allen Fällen aufgrund der niedrigen *Prozentualen Übereinstimmung* als *klinisch inakzeptabel* interpretiert werden, während Neurologen in fast allen Fällen dieser beiden Studien *klinisch akzeptabel* beurteilen. Obwohl die ICC-Werte der Reflexaktivitätsanalyse in Studie 2b bei Neurologen und praktizierenden Tierärzte als *mittelmäßig* interpretiert werden, ist die Kappa-Analyse für beide Gruppen in jeweils nur einem Fall *klinisch akzeptabel* und ansonsten *klinisch inakzeptabel*. Die K_w -Analyse der Studenten ist in allen Studien *klinisch inakzeptabel*, jedoch beschränkt sich das von Burn und Weir (2011) vorgeschlagene Interpretationsmodell von Kappa auf dichotome Beurteilungen und berücksichtigt somit keine Gewichtung, wie bei K_w verwendet.

Durch die wie in dieser Arbeit angewendete gemeinsame Interpretation der *Prozentualen Übereinstimmung*, des Kappa-Koeffizienten und seiner Interpretationsparameter kann eine Aussage darüber getroffen werden, inwieweit ein hohes Interobserver Agreement auf eine tatsächliche Beurteilerübereinstimmung oder größtenteils auf eine inbalancierte Verteilung der Beurteilungen zurückzuführen ist (Burn und Weir, 2011). Leider nennt die Literatur keinen Richtwert des Prävalenz-Index, ab wann eine Beurteilungsverteilung als inbalanciert gilt. Generell heißt es, dass Kappa sinkt wenn die Prävalenz eines Merkmals gegen 1 tendiert

(Hansen et al., 1994; Sim und Wright, 2005), allerdings nicht mit welchem *PI* diese Tendenz definiert ist. Burn und Weir (2011) sprachen in ihrer Arbeit von „zu homogenen“ Untersuchungsgruppen ab einem *PI* von 0,9. Auffällig ist, dass *PI*-Werte von $\geq 0,9$ in allen drei Studien nur bei der Reflexpräsenzanalyse erreicht werden. Dies erklärt sich aus der Wahrscheinlichkeit, dass ein Großteil der Beurteilungen in eine Kategorie fällt, umso höher ist, je weniger Kategorien existieren. Wie bereits erwähnt, behandelten auch die zitierten Arbeiten nur Reliabilitätsanalysen mit dichotomer Beurteilungsskala.

Die Ergebnisse dieser Arbeit zeigen, dass die Anzahl von Beurteilerübereinstimmungen mit einem $PI \geq 0,9$ deutlich höher bei der Reflexpräsenzbeurteilung des Patellar- ($n = 6/9$, Studie 2b) als bei der des Bizepssehnenreflexes ($n = 1/9$, Studie 2a) ist. Während der genannte K_C -Wert in Studie 2a zwischen Beurteiler N1 und N3 als *klinisch akzeptabel* interpretiert werden kann, muss die Kappa-Analyse von vier der insgesamt sechs genannten Beurteilerübereinstimmungen in Studie 2b aufgrund eines $K_C < 0,40$ als *klinisch nicht eindeutig* bewertet werden. Das deutet darauf hin, dass die Beurteilungen in Studie 2b häufiger inbalanciert vorliegen, als in Studie 2a und K_C daher vermutlich „falsch-niedrig“ ist. Es ist offensichtlich, dass die Beurteilung des Patellarsehnenreflexes wesentlich homogener ist als die des Bizepssehnenreflexes, was in der inbalancierten Verteilung der Randverteilungen resultiert. Deutlich wird dies am ehesten an dem negativen K_C (-0,02) des Beurteilerpaars N1-N3, sowie den Beurteilungen der Beurteilerpaare T1-T3 und T2-T3 in Studie 2b, die jeweils $K_C = 0$ ergeben. In der Gesamtheit liegen die K_C -Werte von 12 Beurteilerkombinationen (33,33%) bei der Beurteilung der Patellarsehnenreflexpräsenz in Studie 2a $\leq 0,00$ (s. Anhang, Tabelle 44). Vor dem Hintergrund, dass jede dieser Beurteilungen eine *Prozentuale Übereinstimmung* von deutlich über 90% erreicht, ist davon auszugehen, dass die hohe Prävalenz der Kategorie „normal“ innerhalb der Beurteilungen den niedrigen Kappa-Werten zugrunde liegt.

Auf die Berechnung alternativer Kappa-Koeffizienten, die Prävalenz- und Bias-korrigiert sind („prevalence-adjusted and bias-adjusted kappa“, PABAK) (Byrt et al., 1993) oder das „free-marginal multirater K“ (Randolph, 2005) wird verzichtet, da Prävalenz und Bias wichtige Informationen bei der Interpretation und dem Vergleich zwischen Studien liefern und deren Korrektur daher sehr umstritten ist (Hoehler, 2000).

Der Einfluss der Beurteiler- und Kategorienanzahl der verwendeten Ordinalskala wird ebenfalls bei vergleichender Betrachtung der Literatur deutlich (s. Tabelle 1). Annaswamy et al. (2012)

untersuchten das Interobserver Agreement zwischen zwei Beurteilern bei der Beurteilung des *Hoffmann-Zeichens* mittels einer dichotomen Beurteilungsskala und dokumentierten eine *ordentliche* Übereinstimmung von $K_C = 0,65$. Thaller und Hughes (2014) ließen zwei Beurteiler eine neurologische Untersuchung durchführen und berechneten das Interobserver Agreement zwischen ihren Ergebnissen. Für die Reflexbeurteilung verwendeten sie eine 4-Punkteskala. Das Interobserver Agreement für die spinalen Reflexe lag bei $K_C = 0,54$ (= *mittelmäßig*). Litvan et al. (1996) präsentierten ebenfalls ein im Median *mittelmäßiges* Interobserver Agreement. Sie legten ihren Ergebnissen die Beurteilung der Reflexaktivität anhand einer 5-Punkte-Skala zugrunde und maßen das Interobserver Agreement zwischen vier Neurologen mit etwa gleicher Erfahrung. Die Ergebnisse sind somit vergleichbar mit denen der Reflexaktivitätsanalyse in Studie 1. In der Studie von Litvan et al. (1996) waren die Untersucher allerdings gleichzeitig Beurteiler und es ist bekannt, dass das Interobserver Agreement zwischen Studien, in denen der Untersucher gleichzeitig Beurteiler ist, stark variiert (vgl. Hansen et al., 1994; Manschot et al., 1998; Stam und van Crevel, 1990; Vogel, 1992). Stam und van Crevel (1990) untersuchten das Interobserver Agreement der Reflexaktivität beim Menschen zwischen drei Neurologen, die jeweils selbst die Untersuchungen durchführten. Sie verwendeten im Gegensatz zu der hier vorgelegten Studie eine 9-Punkte-Skala womit die Wahrscheinlichkeit einer *vollständigen Übereinstimmung* per se wesentlich geringer ist. Die Autoren fanden eine Übereinstimmung der drei Untersucher bei der Evaluierung des Bizepssehnenreflexes in 17,5% und bei der Evaluierung des Patellarsehnenreflexes in 35% der Fälle. In der hier präsentierten Studie ist der Anteil *vollständiger Übereinstimmungen* in der Neurologen-Gruppe bei der Evaluierung des Bizepssehnenreflex (73%) höher als beim Patellarsehnenreflex (61%), ferner ist die Differenz dieser Anteile zwischen den beiden Reflexen wesentlich geringer als bei Stam und van Crevel (1990). Der Anteil *partieller (Nicht-) Übereinstimmungen* und *vollständiger Nichtübereinstimmungen* war bei Stam und Crevel (1990) ebenfalls höher bei der Evaluierung des Bizepssehnenreflexes als bei der Evaluierung des Patellarsehnenreflexes. Berücksichtigt man, dass die drei Untersucher in der zitierten Studie allesamt Neurologen waren, so steht die Verteilung der Evaluierungen innerhalb der jeweiligen Übereinstimmungsstufen im klaren Gegensatz zu den Ergebnissen der in dieser Arbeit untersuchten Neurologengruppe nach Auszählung der Studien 2a und 2b. Ebenfalls sind die Differenzen zwischen der Anzahl der Beurteilungen innerhalb der jeweiligen Übereinstimmungsstufe zwischen den beiden Reflexen in der hier präsentierten Arbeit deutlich geringer als bei Stam und van Crevel (1990).

Die Tatsache, dass in der vorliegenden Studie ausschließlich gesunde Hunde eingeschlossen wurden, impliziert, dass die Prävalenz der Gliedmaßen, die von den Beurteilern als „normal“ klassifiziert werden, erwartungsgemäß hoch ist. Ferner führt der hohe Prävalenz-Index wie bereits erwähnt dazu, dass die Ergebnisse der Reliabilitätsanalyse unter Vorbehalt interpretiert werden müssen, da es durch eine inbalancierte Verteilung der Beurteilungen über den Kategorien zu einer Abnahme von Kappa und somit einer Unterschätzung der Konkordanz kommt (Hansen et al., 1994; Sim und Wright, 2005). Burn und Weir (2011) rieten daher dazu, im Vorhinein balancierte beziehungsweise nicht homogene Probandengruppen zu wählen. In einer hier anschließenden Arbeit könnte man die präsentierten Ergebnisse mit einer heterogenen Gruppe vergleichen. Die Probandengruppe sollte sich hierbei aus einer annähernd gleichen Anzahl an Hunden zusammensetzen, bei denen im Vorfeld eine Läsion innerhalb der für den Bizepssehnenreflex verantwortlichen Rückenmarkssegmenten C₆-T₂ mittels Bildgebung verifiziert und solchen, bei denen eine Läsion in dieser neuroanatomischen Lokalisation ausgeschlossen wurde.

5.6. Beeinflussung des Interobserver Agreement

Die Auswahl der in dieser Studie untersuchten Risikofaktoren auf das Interobserver Agreement orientiert sich an der existierenden veterinär- und humanmedizinischen Literatur. Grundsätzlich lassen sich die untersuchten Risikofaktoren in intrinsische (*Alter, Gewicht, Geschlecht* und *Felllänge*) und extrinsische (*Beurteilerexpertise* und *Körperhälfte*) unterteilen.

Untersuchungen aus der Humanmedizin zeigten, dass eine hohe Variabilität des Interobserver Agreement zwischen Studien existiert, in denen die Reflexantwort subjektiv beurteilt wurde (Annaswamy et al., 2012; Dafkin et al., 2013; Hansen et al., 1994; Manschot et al., 1998; Stam und van Crevel, 1990; Vogel, 1992). Es muss allerdings berücksichtigt werden, dass Studien zur Untersuchung des Interobserver Agreement sich hinsichtlich der Verwendung der Beurteilungsskala, der Anzahl der Beurteiler, des Versuchsaufbaus und der berechneten Reliabilitätskoeffizienten und ihrer Interpretation unterscheiden und somit nicht per se untereinander vergleichbar sind (Koran, 1975; Vroomen et al., 2000).

Es ist bekannt, dass das Interobserver Agreement der Beurteilung der Reflexantwort zwischen den verschiedenen Reflexen variiert. Aus verschiedenen humanmedizinischen Studien geht

hervor, dass die Reproduzierbarkeit der Evaluation besser in den unteren Extremitäten im Vergleich zu den oberen Extremitäten ist (Kuruoglu und Oh, 1993; Stam und van Crevel, 1990).

Vreeling et al. (1993) konnten das Interobserver Agreement zwischen zwei Neurologen deutlich verbessern, nachdem sie die Methode der Reflexprovokation standardisierten. Allerdings konstatierten sie ebenfalls, dass zwischen etablierten Untersuchern trotz Standardisierung der Methodik immer noch Unterschiede in der Durchführung und Interpretation der Reflexevaluierung bestünden. In der vorgelegten Arbeit wurde Untersucher 2 vor Beginn der Studie mit der Reflexprovokation des Bizeps- und Patellarsehnenreflexes durch Untersucher 1 eingewiesen. Die Untersuchungsmethode zwischen den Untersuchern unterscheidet sich dennoch geringfügig. Diese Unterschiede bilden allerdings den klinischen Alltag ab und erhöhen die Generalisierbarkeit der Ergebnisse (Singerman und Lee, 2008). In den Studien 2a und 2b beurteilen die Beurteiler ausschließlich die Versuchsdurchführung durch Untersucher 2. Die Provokation der Reflexantwort wurde somit konstant gleich durchgeführt. Dieser Umstand könnte neben der Reduzierung von Beurteilern und Untersuchungssequenzen ein Grund sein, warum das Interobserver Agreement in Studie 2a bei fast allen Gruppen höher ist, als in Studie 1. Interessanterweise zeigt die Regressionsanalyse dennoch keinen signifikanten Einfluss des Untersuchers auf das Interobserver Agreement.

Beurteiler mit einer größeren Erfahrung stimmen in ihrer Einschätzung öfter überein, als Beurteiler mit einer geringeren Expertise. Über diese von Koran (1975) aufgestellte Hypothese herrscht allerdings in der Literatur Uneinigkeit. Diverse Studien statuierten, dass sich das Interobserver Agreement durch Übung der Beurteiler in der Evaluierung des zu beurteilenden Merkmals erhöhen lässt (Koran, 1975; Kottner et al., 2009b; Meade et al., 1968; O'Keefe et al., 1994; Raftery und Holland, 1967; Raijmakers et al., 1991). Raijmakers et al. (1991) konnten eine zuverlässigere Evaluierung des Plantarreflexes durch Studenten nach Betrachten eines Lehrvideos feststellen. In ihrer Studie diente ähnlich wie in der vorgelegten Arbeit die Beurteilung durch einen Neurologen als Referenzkategorie. Dafkin et al. (2013) zeigten hingegen, dass die Beurteilerexpertise keinen Einfluss auf die Genauigkeit der Evaluierung des Patellarsehnenreflexes hat. Die Autoren unterteilten die Beurteiler in drei Gruppen (Studenten, praktizierende Ärzte und Neurologen), allerdings waren die Gruppen wesentlich heterogener als in der vorliegenden Arbeit. So hatten beispielsweise die Neurologen klinische Erfahrung mit der Evaluierung des Patellarsehnenreflexes zwischen 5 - 17 Jahren, die praktizierenden Ärzte sogar zwischen 6 - 41 Jahren. Litvan et al. (1996) resümierten, dass Übung in der

Beurteilung der Reflexantwort ebenfalls keinen Einfluss auf die Beurteilung der Reflexantwort hat, allerdings untersuchten in ihrer Studie vier bereits etablierte Neurologen die Reflexe an jeweils 40 Patienten vor und nach einer Trainingseinheit. Der Effekt einer Trainingseinheit sollte bei etablierten Neurologen nach Ansicht des Autors ohnehin gering sein, was die Interpretation der Ergebnisse fraglich erscheinen lässt. Beim Versuchsaufbau der vorliegenden Studie wurde sich explizit gegen die Einarbeitung der Versuchsteilnehmer entschieden, um die hypothetisch vorhandenen Unterschiede im Interobserver Agreement bei der Reflexevaluierung zwischen den Gruppen deutlich abbilden zu können.

Dafkin et al. (2013) konnten herausarbeiten, dass die Beurteiler in ihrer Studie die Reflexantwort des Patellarsehnenreflexes umso höher skalierten, je höher die maximale Reflexamplitude und je kürzer die Reflexzeit waren. Dabei schienen praktizierende Ärzte und Studenten ihre Beurteilung hauptsächlich auf dem Ausmaß der Kniegelenkwinkeländerung zu gründen. Interessanterweise konnten die Autoren der erwähnten Studie für die Gruppe mit der am höchsten zu erwartenden Expertise, den Neurologen, keinen statistischen Zusammenhang zwischen Interobserver Agreement und kinematischen Messdaten herstellen und kamen zu dem Schluss, dass Neurologen die Reflexantwort des Patellarsehnenreflexes anhand nicht gemessener Kriterien quantifizierten. Nach welchen Kriterien die Beurteiler in der hier vorgelegten Arbeit jeweils die Reflexantwort bewerteten, wurde nicht untersucht. Die Tatsache, dass sich das Interobserver Agreement zwischen den Gruppen bei der Beurteilung des Bizepssehnenreflexes mehr unterscheidet als bei der des Patellarsehnenreflexes, deutet allerdings daraufhin, dass die Kriterien bei der Beurteilung des Bizepssehnenreflexes entweder weniger gut detektierbar oder weniger gut vertraut sind, je niedriger die Beurteilerexpertise ist (Singerman und Lee, 2008). Die Ergebnisse der Regressionsanalyse belegen, dass die Detektion einer Kontraktur des *M. biceps brachii* bei Hunden mit langem Fell schwieriger ist, als bei Hunden mit kurzem Fell, erst recht wenn die Beurteilerexpertise niedrig ist. Des Weiteren ist der Rückschluss zulässig, dass es basierend auf den Anteilen der *vollständigen Übereinstimmungen* Kriterien gibt, die in Abhängigkeit der Beurteilerexpertise besser oder weniger gut bekannt und bewertet werden können. Ein signifikanter Einfluss der Felllänge auf das Interobserver Agreement kann hingegen nicht für den Patellarsehnenreflex gezeigt werden. Dies würde die vergleichsweise bessere Reliabilität bei Evaluierung des Patellarsehnenreflexes bei den weniger erfahrenen Beurteilern (praktizierende Tierärzte und Studenten) erklären.

Annaswamy et al. (2012) diskutierten, ob die Händigkeit des Beurteilers eine Auswirkung auf die suffiziente Provokation des Hoffmann-Zeichens in ihrer Studie hätte haben können. In den in der Literatur vorliegenden Studien wird die Reflexantwort zwischen beiden Körperseiten vergleichend evaluiert, um eine etwaige Asymmetrie der Reflexantwort zu detektieren. In der hier vorliegenden Arbeit erfolgte diese Unterteilung mit dem Ziel den Einfluss der hierdurch veränderten Positionierung der Hand des Untersuchers zu untersuchen. Es wird also in dieser Arbeit mit der Untersuchung des Einflusses der Körperseite auf das Interobserver Agreement der Effekt einer durch den Untersucher herbeigeführten veränderten Positionierung untersucht, womit der Risikofaktor *Körperseite* in dieser Arbeit als extrinsischer Faktor definiert wurde. Die erwähnte geringgradig veränderte Positionierung erhöht den eigenen Ergebnissen folgend nicht signifikant die Wahrscheinlichkeit einer diskordanten Beurteilung. Es wäre interessant zu untersuchen, ob die Händigkeit des Untersuchers die Wahrscheinlichkeit einer Diskordanz erhöht. Dieser Faktor wurde nach Kenntnis des Autors bislang nicht untersucht und in der vorgelegten Arbeit waren beide Untersucher in Studie 1 rechtshändig.

Auf die Analyse eines hypothetisch positionsbedingten Einflusses auf das Interobserver Agreement bei der Evaluierung der Reflexantwort des Bizepssehnenreflexes wird aufgrund der zweifelhaften zuverlässigen Beurteilbarkeit anhand eines Videos verzichtet. De Lahunta und Glass (2009a) und auch Thomas und Dewey (2008) empfahlen zwar die vergleichende bodennahe Evaluierung des Patellar- beziehungsweise Bizepssehnenreflexes, allerdings ist anzunehmen, dass beim hier skizzierten Versuchsaufbau eine am Boden liegende Gliedmaße durch die bodenferne überdeckt würde und so eine videobasierte Beurteilung nicht gewährleistet werden könnte. Die mangelhafte Beurteilbarkeit würde das Ergebnis somit erwartungsgemäß verfälschen.

Die Ergebnisse liefern das statistische Fundament für die von Thomas und Dewey (2008) vermutete Schwierigkeit der Beurteilung des Bizepssehnenreflexes bei Hunden mit langem Fell. In Studie 2a wird deutlich, dass langes Fell die Wahrscheinlichkeit einer diskordanten Beurteilung um den Faktor 2,6 erhöht. Für den Patellarsehnenreflex kann kein Einfluss der Felllänge auf das Interobserver Agreement berechnet werden. Nach Meinung des Autors wird die Sichtbarkeit der Extension des Kniegelenkes weniger durch langes Fell gestört, als die oftmals dezente Kontraktion des Musculus biceps brachii oder die leichte Flexion des Ellbogengelenkes, welche die Kriterien der Bizepssehnenreflexantwort bilden. Bezogen auf die Videoanalyse ist noch anzunehmen, dass die Fellfarbe auch einen Einfluss auf die Detektion

der Reflexantwort des Bizepssehnenreflexes haben könnte. Zwar wurde eine Kamera mit automatischem Weißabgleich verwendet und bei kontrollierten, standardisierten Lichtverhältnissen gearbeitet, allerdings fiel dem Autor dieser Arbeit die Evaluierung des Bizepssehnenreflexes in Studie 1 bei Hunden mit dunklem Fell subjektiv schwerer, als bei Hunden mit hellem Fell. Auf eine statistische Auswertung wurde aufgrund der geringen Anzahl an Gliedmaßen mit dunklem Fell ($n = 8$) und der subjektiven Definition von „dunkel“ verzichtet.

Studien aus der Human- (Chandrasekhar et al., 2013; Dick, 2003) und Veterinärmedizin (Levine et al., 2002) belegten, dass die Reflexaktivität mit zunehmenden Alter abnimmt. Die Ursachen dafür sind divers und Gegenstand gegenwärtiger Diskussion (Aminoff, 2012; Balagopal et al., 1997; Deschenes, 2004a und b). Levine et al. (2002) detektierten bei 6 von insgesamt 86 untersuchten Hunden eine abwesende Reflexantwort des Patellarsehnenreflexes in einer oder beiden Hinterextremitäten. Zwei dieser Beurteilungen fielen in die Gruppe der Hunde mit einem Alter von < 10 (Gruppe 1), vier in die Gruppe ≥ 10 Jahre (Gruppe 2). Auf diesen proportionalen Anteilen gründeten sie die Schlussfolgerung, dass bei Hunden ab zehn Jahren signifikant häufiger eine abwesende Reflexantwort des Patellarsehnenreflexes vorliegt. Es muss jedoch erwähnt werden, dass sich die beiden verglichenen Gruppen hinsichtlich ihrer Verteilung stark unterschieden (Gruppe 1 = 72 Hunde, Gruppe 2 = 14 Hunde). Bei der Kategorisierung der Hunde in der vorliegenden Arbeit wird sich am Median des Alters der Probanden orientiert. Wie aus den Ergebnissen hervorgeht, resultiert diese Unterteilung in einer wesentlich gleichmäßigeren Verteilung über den Unterkategorien „jung“ ($n = 32$; 50%) und „alt“ ($n = 32$; 50%) in Studie 2b. Dadurch liegt der Trennwert (6 Jahre) jedoch auch deutlich niedriger als bei Levine et al. (2002) (10 Jahre). Ein eventueller Einfluss des Alters, wie Levine et al. (2002) es definierten, könnte in dieser Studie also übersehen worden sein, da Hunde von über 10 Jahren nicht separat betrachtet wurden. Anders als Levine et al. (2002), die allein den Einfluss des Alters auf die Präsenz des Patellarsehnenreflexes untersuchten, ist es ein Ziel der vorliegenden Arbeit ebenfalls die Beurteilerübereinstimmung der Patellarsehnenreflexaktivitätsevaluierung zu quantifizieren. Es kann in der vorliegenden Arbeit statistisch nicht belegt werden, dass das Alter des Hundes die Wahrscheinlichkeit einer diskordanten Beurteilung erhöht. Chandrasekhar et al. (2013) konnten belegen, dass beim Menschen die altersbedingte Reduktion der Reflexantwort am deutlichsten bei einem Schlagwinkel von 45° ist. Die Untersuchung des Schlagwinkels beziehungsweise der

Impulssetzung ist jedoch nicht Gegenstand der vorliegenden Arbeit. Auch hier würde eine Standardisierung den klinischen Alltag nicht korrekt abbilden.

Stam und van Crevel (1990) statuierten, dass die im klinischen Alltag stattfindende Kommunikation über die Anamnese eines Patienten zwischen verschiedenen Untersuchern das Interobserver Agreement erhöht. Vogel (1992) konnte nachweisen, dass sich das Interobserver Agreement zwischen zwei Beurteilern bei der ordinalen Beurteilung der Reflexaktivität signifikant erhöhen ließ, wenn beide Beurteiler die Krankengeschichte des Patienten kannten (Vogel, 1992). Auf die dichotome Beurteilung der Reflexpräsenz hatte die zusätzliche Information laut Vogel (1992) keinen Einfluss. Auch bei der Evaluierung von Thoraxröntgenbildern konnten Doubilet und Herman (1981) die richtig-positive Interpretation durch die Beurteiler erhöhen, nachdem sie zusätzliche Informationen des Patienten ergänzten. Vogel (1992) diskutierte die Gefahr des Primings, was bedeutet, dass durch die Kenntnis der Krankengeschichte die Symptome selektiv anhand der Vorgeschichte interpretiert werden könnten und so in verfälschten Ergebnissen resultieren.

Im klinischen Alltag ist die Untersuchung durch drei oder vier Untersucher eher von untergeordneter Bedeutung (Koran, 1975). Mit Koran (1975) übereinstimmend, ist es auch aus eigener Erfahrung meist so, dass die neurologische Untersuchung des Patienten im klinischen Alltag von maximal zwei Beurteilern durchgeführt wird. Vor dem Hintergrund, dass die Wahrscheinlichkeit diskordanter Beurteilungen mit der Anzahl der Beurteiler steigt, ist davon auszugehen, dass die präsentierten Reliabilitätskoeffizienten K_F und ICC im klinischen Alltag von geringerer klinischer Relevanz sind. Dennoch liegen bei der Reflexpräsenzanalyse aller drei Studien $\bar{X}K_C$ und $K_{F\text{ Präs}}$ im gleichen Übereinstimmungsmaß beziehungsweise unterscheiden sich in einem Fall nur marginal (praktizierende Tierärzte in Studie 2b). Ein Vergleich zwischen $\bar{X}K_w$ und $K_{F\text{ Akt}}$ scheitert an der fehlenden Gewichtung von K_F .

Im Rahmen der neurologischen Untersuchung ist es zudem meist so, dass der Untersucher auch gleichzeitig der Beurteiler ist. Zwar ist bekannt, dass das Interobserver Agreement zwischen Studien, in denen Untersucher und Beurteiler dieselbe Person ist, stark variiert (vgl. Hansen et al., 1994; Manschot et al., 1998; Stam und van Crevel, 1990; Vogel, 1992), allerdings ist keine Studie bekannt, in der untersucht wurde, inwieweit sich das Interobserver Agreement verändert hätte, wenn der Beurteiler selbst untersucht hätte. Es ist somit unklar, inwieweit sich das Interobserver Agreement verändert hätte, wenn die Beurteiler der vorgelegten Arbeit die

Reflexprovokation selbst hätten durchführen müssen. Vor dem Hintergrund, dass mehrere Autoren die Durchführung der Reflexprovokation des Bizepssehnenreflexes als schwierig beurteilten (Braund und Sharp, 2002; Oliver et al., 1997; Schatzberg et al., 2012), können die hier präsentierten Ergebnisse diese Aussage weder bestätigen, noch dementieren, da sich in dieser Arbeit ausschließlich auf die Beurteilung und nicht die Durchführung fokussiert wird. Es wird gezeigt, dass der Bizepssehnenreflex zuverlässig ausgelöst werden kann und in weiterer Folge zu beurteilbaren Reflexantworten führt. Die Regressionsanalyse in Studie 1 zeigt darüber hinaus, dass der Untersucher keinen Einfluss auf die Güte der Beurteilerübereinstimmung nimmt. Bemerkenswert ist noch, dass K_C für die beiden Untersucher in der Reflexpräsenzanalyse in Studie 1 innerhalb ihrer Gruppe am niedrigsten ist ($K_C = 0,49$; s. Anhang, Tabelle 34). Der Unterschied zu den anderen Beurteilerkombinationen ist allerdings nicht signifikant.

Die übereinstimmende Beurteilung der Bizepssehnenreflexantwort ist wesentlich von der Expertise des Beurteilers abhängig. Die Ergebnisse der vorgelegten Arbeit bilden diesen Einfluss gut ab. Während Neurologen bei der Beurteilung des Bizepssehnenreflexes mit einer hohen Wahrscheinlichkeit übereinstimmend urteilen, sollte die Interpretation von Beurteilern ohne veterinärneurologische Spezialisierung mit Vorbehalt betrachtet werden. Dies untermauert die häufig in der Fachliteratur erwähnte angezweifelte zuverlässige Beurteilbarkeit. Dass die Beurteilerexpertise einen höheren Einfluss auf das Interobserver Agreement der Evaluierung des Bizepssehnenreflexes, als auf die des Patellarsehnenreflexes nimmt, ist sicherlich dem Umstand geschuldet, dass die Durchführung der Untersuchung des Bizepssehnenreflexes und damit einhergehend seine Interpretation in der Literatur eher marginal beschrieben beziehungsweise der neurologischen Fachliteratur vorbehalten bleibt und somit kaum im Rahmen der neurologischen Untersuchung untersucht wird. Bei Neurologen, die mit dieser Literatur vertraut sind, ist die Wahrscheinlichkeit somit höher übereinstimmend zu evaluieren. Des Weiteren ist die Interpretation des Bizepssehnenreflexes, wie bereits erwähnt, bei Hunden mit langem Fell schwieriger als bei Hunden mit kurzem Fell und erfordert daher mehr Erfahrung.

5.7. Ausblick

Inwieweit die Evaluierung des Bizepssehnenreflexes klinische Relevanz besitzt beziehungsweise einen zuverlässigen Parameter in der Detektion von Läsionen innerhalb des unteren motorischen Neurons der Vorderextremität darstellt, muss in folgenden Studien untersucht werden. Interessant wäre in diesem Zusammenhang auch, ob sich seine Sensitivität beziehungsweise Spezifität von der des Flexorreflexes unterscheidet. Vor dem Hintergrund, dass die Sensitivität jeder der in der Vorderextremität beschriebenen Spinalreflexe von verschiedenen Autoren auch angezweifelt wird, könnte der Bizepssehnenreflex ein notwendiger zusätzlicher Parameter sein, die Integrität des unteren motorischen Neurons der Vorderextremität beim Hund zu überprüfen.

Auf der Basis der hier ermittelten Reliabilitätsanalyse sollten sich zukünftige Studien darauf konzentrieren, die Genauigkeit des Bizepssehnenreflexes beim Hund zu untersuchen, um dessen medizinischen Nutzen für den klinischen Alltag zu erörtern. Der Versuchsaufbau sollte in folgenden Studien sowohl gesunde Hunde als Kontrollgruppe, als auch Hunde mit einer diagnostizierten Läsion des verantwortlichen Reflexbogens inkludieren. Ein weiterer interessanter Ansatz wäre die Gegenüberstellung der subjektiven Interpretation der Reflexantwort verschiedener spinaler Reflexe und kinematischer Messdaten, wie zum Beispiel die Gelenkwinkeländerung während der Reflexantwort. Modelle, die die theoretische Grundlage einer solchen Untersuchung bilden könnten, liegen bereits aus der Humanmedizin vor und wurden in dieser Arbeit hinreichend zitiert. Eine Reliabilitätsanalyse des Flexorreflexes liegt in der veterinärmedizinischen Literatur nach Kenntnis des Autors ebenfalls nicht vor. Seine Untersuchung kann nicht anhand eines Videos stattfinden, da die Kraft der Gliedmaßenbeugung nicht nachvollzogen werden könnte. Ergänzend muss berücksichtigt werden, dass die Kraft probandenspezifisch bewertet werden muss, da ein kleiner Hund erwartungsgemäß weniger absolute Kraft aufbringt, als ein großer Hund. In Anlehnung an bereits existierende Modelle könnte man die Stärke der resultierenden Flexorenkontraktion mittels Oberflächenelektromyographie quantifizieren.

Eine Reliabilitätsanalyse der neurologischen Untersuchung des Hundes anhand von Videodokumentation könnte die Grundlage bilden, die Telemedizin auf klinischer Ebene für die Veterinärmedizin zu etablieren. Ein gutes Ergebnis könnte im klinischen Alltag die

Flexibilität der Mitarbeiter fördern und die Anzahl der durch Personen mit geringer neurologischer Expertise gestellter Fehldiagnosen reduzieren.

6. Zusammenfassung

Das Ziel dieser Arbeit ist erstens die vergleichende Quantifizierung des Interobserver Agreements (IA) bei der Beurteilung des Bizeps- (BSR) und Patellarsehnenreflexes (PSR) gesunder Hunde und zweitens die Evaluierung des Einflusses intrinsischer (*Geschlecht, Alter, Felllänge, Gewicht*) und extrinsischer Faktoren (*Beurteilerexpertise, Körperseite, Untersucher*) auf das IA beider Reflexe.

Die Beurteiler werden gemäß ihrer zu erwartenden Expertise in drei Gruppen unterteilt (Neurologen = höchste Expertise, praktizierende Tierärzte = mittlere Expertise, Studenten = niedrigste Expertise). Jeder Beurteiler evaluiert unabhängig voneinander die Reflexpräsenz (RP(BSR), RP(PSR)) und Reflexaktivität (RA(BSR), RA(PSR)) der beiden Reflexe anhand eines Videos mittels einer 5-Punkte-Ordinalskala. In Studie 1 werden hierzu insgesamt 103 anonymisierte Videosequenzen von 61 Vordergliedmaßen evaluiert, von denen jeweils 42 Sequenzen von zwei Untersuchern (U1, U2) und 19 von jeweils einem der beiden Untersucher untersucht werden. In Studie 2 werden ausschließlich die Untersuchungen aus Studie 1 von U2 (n = 54) ausgewertet und der Evaluierung des PSR an 64 Hintergliedmaßen, die ebenfalls durch U2 untersucht werden, gegenübergestellt.

Berechnet werden verschiedene Reliabilitätskoeffizienten sowie für jede Beurteilergruppe die Anzahl *vollständiger Übereinstimmungen, partieller (Nicht-)Übereinstimmungen* mit einer Differenz von einem, respektive zwei Skaleneinheiten und *vollständiger Nichtübereinstimmungen*. Der Einfluss der untersuchten Faktoren wird mittels Regressionsanalyse berechnet und die Höhe in Form der Odds Ratio wiedergegeben.

Die RP(BSR) wird umso übereinstimmender evaluiert, je höher die Expertise der jeweiligen Beurteilergruppe ist und liegt für Neurologen signifikant am höchsten. Bei der RA(BSR) ist ebenfalls das IA und die Anzahl konkordanter Beurteilungen umso höher, je höher die Expertise der Beurteiler ist und liegt für Studenten signifikant am niedrigsten. Die Beurteilerexpertise besitzt einen signifikanten Einfluss auf das IA bei der Beurteilung des BSR: für die RP haben praktizierende Tierärzte eine 3,8-fach, Studenten eine 5,7-fach höhere Wahrscheinlichkeit als Neurologen diskordant zu beurteilen. Für die RA(BSR) haben praktizierende Tierärzte eine 4,4-fach und Studenten eine 7,6-fach höhere Wahrscheinlichkeit einer Nichtübereinstimmung. Bei Hunden mit langem Fell ist die Wahrscheinlichkeit die RA(BSR) nicht übereinstimmend zu evaluieren 2,6-fach höher als bei Hunden mit kurzem Fell. Dieser Effekt ist signifikant. Die

RP(PSR) wird ebenfalls umso übereinstimmender beurteilt, je höher die Expertise der Beurteilergruppe ist. Das IA ist für Studenten bei der Beurteilung der RA(PSR) signifikant am niedrigsten. Der Anteil *vollständiger Übereinstimmungen* ist bei Neurologen höher für den Bizepssehnenreflex als für den Patellarsehnenreflex, während praktizierende Tierärzte und Studenten jeweils häufiger konkordant die RA(PSR) beurteilen als die des RA(BSR). Neurologen und praktizierende Tierärzte zeigen für die Evaluierung der RA(PSR) eine identische Verteilung an Übereinstimmungen und Nichtübereinstimmungen und urteilen innerhalb der Gruppe mit einer Differenz von maximal einer Skaleneinheit. Keiner der untersuchten Einflussfaktoren nimmt einen signifikanten Einfluss auf die Höhe des IA bei der Evaluierung der RP(PSR), hingegen haben Studenten eine 4,3fach höhere Wahrscheinlichkeit die RA(PSR) diskordant zu beurteilen als die beiden anderen Beurteilergruppen. Dieser Effekt ist signifikant ($p < 0,001$).

Der Bizepssehnenreflex kann von Neurologen zuverlässig beurteilt werden. Die Beurteilerexpertise des Beurteilers und die Felllänge des Hundes nehmen einen Einfluss auf das IA bei der Beurteilung des Bizepssehnenreflexes. Hierbei nimmt die Beurteilerexpertise einen größeren Einfluss auf das IA der Beurteilung des BSR als auf das der PSR-Beurteilung.

7. Summary

Reliability-analysis of the subjective assessment of the bicipital reflex in healthy dogs

The goals of this thesis are two-folded: first the comparative quantification of the Interobserver Agreement (IA) on the evaluation of the canine bicipital (BTR) and patellar tendon (PTR) reflex in healthy dogs and second the identification of intrinsic (*sex, age, fur length, weight*) and extrinsic (*observer's expertise, body side, examiner*) risk factors on the IA of both reflexes.

The observers are subdivided in three groups based on their expected level of expertise (neurologists = highest -, practitioners = middle – and veterinary students = lowest level of expertise). Each observer evaluates independently the reflex presence (RP(BTR), RP(PTR)) and the reflex activity (RA(BTR), RA(PTR)) of both reflexes on a videotape using a 5-point-ordinalscale. For study 1, 103 anonymized sequences of 61 forelimbs are evaluated, of which 42 sequences have been examined by both examiners (E1, E2) and 19 by one of both examiners (E1 or E2) each. For study 2 only the sequences examined by E2 in study 1 (n = 54) were analyzed and compared to the evaluation of the patellar tendon-reflex on 64 hindlimbs, which have been examined only by E2 as well.

Multiple reliability coefficients were calculated. Additionally, the amount of *complete agreement*-, *partial (dis)agreement* with difference of 1 or ≥ 2 points of the ordinalscale- and *complete disagreement*-evaluations are counted for all groups. The influence of the risk factors mentioned above are calculated using a mixed regression-model. The Odds Ratio for each factor was presented.

The higher the level of expertise the higher is the IA for RP(BTR) and is significant highest for neurologists. For the RA(BTR), the amount of *complete agreement*-evaluations was the higher the higher the level of expertise of the observer was. The IA of the RA(BTR) is significantly lowest for students. The level of expertise has a significant impact on the height of the IA in the evaluation of the bicipital tendon-reflex: for the RP(BTR) practitioners have a 3.8-way, students a 5.7-way higher risk to judge discordantly compared to the neurologists. For the RA(BTR), practitioners have a 4.4-way, students a 7.6-way higher risk of discordance. In longhaired dogs the likelihood of disagreement is 2.6-way higher compared to shorthaired dogs in the evaluation of RA(BTR). This impact is significant. Likewise, the RP(PTR) is the higher the higher the observers' expertise is. Students have the significant lowest IA of the evaluation of the

RP(PTR). The RA(PTR) is evaluated significantly more consistent by practitioners and students than the RA(BTR). For practitioners this difference is significant. For neurologists there is no difference between the IA of the RA(BTR) and RA(PTR). Neurologists have a higher amount of *complete agreement*-evaluations in RA(BTR) compared to the RA(PTR), while practitioners and students have a higher amount of *complete agreement*-evaluations in the RA(PTR). Neurologists and practitioners have the identical distribution of *complete agreement*- and *disagreement*-evaluations of RA(PTR) and show no evaluation with a difference of more than one scale-point. None of the examined risk factors has a significant impact on the height of IA in the evaluation of RP(PTR), while students have a 4.3-way higher risk of discordance when evaluating the RA(PTR) compared to the other groups. This effect is significant ($p < 0,001$).

Neurologists can reliably assess the bicipital tendon-reflex. The level of the observer's expertise and the fur length of the dog take an impact on the height of the IA of RA(BTR). The influence of the observer's expertise is higher on the evaluation of the BTR than on the evaluation of the PTR.

8. Literaturverzeichnis

1. Abdelhakiem, M., Asai, Y., Kamishina, H., Katayama, M., Uzuka, Y. (2015):
The accuracy of the patellar reflex for localization of the site of a single level thoracolumbar disc herniation in dogs.
Turkish Journal of Veterinary & Animal Sciences 39: 589-593.
 2. Aghighi, S. A., Tipold, A., Piechotta, M., Lewczuk, P., Kaestner, S. B. R. (2012):
Assessment of the effects of adjunctive gabapentin on postoperative pain after intervertebral disc surgery in dogs.
Veterinary Anaesthesia and Analgesia 39: 636-646.
 3. Aideyan, U. O., Berbaum, K., Smith, W. L. (1995):
Influence of prior radiologic information on the interpretation of radiographic examinations.
Academic Radiology 2: 205-208.
 4. Allam, M. W., Lee, D. G., Nulsen, F. E., Fortune, E. A. (1952):
The Anatomy of the Brachial Plexus of the Dog.
Anatomical Record 114: 173-179.
 5. Altman D. G. (1991):
Some common problems in medical research.
In: Altman, D. G. (Hrsg.): *Practical statistics for medical research*.
Chapman and Hall, 1. Aufl., London. S.: 396-440.
 6. Aminoff M. J. (2012):
Chapter 11 - Clinical Electromyography.
In: Aminoff, M. J. (Hrsg.): *Aminoff's Electrodiagnosis in Clinical Neurology*.
W. B. Saunders, 6. Aufl., London. S.: 233-259.
 7. Andersen, O. K., Sonnenborg, F. A., Arendt-Nielsen, L. (2001):
Reflex receptive fields for human withdrawal reflexes elicited by nonpainful and painful electrical stimulation of the foot sole.
Clinical Neurophysiology 112: 641-649.
 8. Andersen, O. K., Sonnenborg, F. A., Arendt-Nielsen, T. (1999):
Modular organization of human leg withdrawal reflexes elicited by electrical stimulation of the foot sole.
Muscle & Nerve 22: 1520-1530.
 9. Annaswamy, T. M., Sakai, T., Goetz, L. L., Pacheco, F. M., Ozarkar, T. (2012):
Reliability and repeatability of the Hoffmann sign.
PM&R: The Journal of Injury, Function, and Rehabilitation 4: 498-503.
 10. Balagopal, P., Rooyackers, O. E., Adey, D. B., Ades, P. A., Nair, K. S. (1997):
Effects of aging on in vivo synthesis of skeletal muscle myosin heavy-chain and sarcoplasmic protein in humans.
The American Journal of Physiology 273: E790-800.
-

-
11. Barbeau, H., Marchand-Pauvert, V., Meunier, S., Nicolas, G., Pierrot-Deseilligny, E. (2000):
Posture-related changes in heteronymous recurrent inhibition from quadriceps to ankle muscles in humans.
Experimental Brain Research 130: 345-361.
 12. Bathien, N. (1971):
Reflexes spinaux chez l'homme et niveaux d'attention.
Electroencephalography and Clinical Neurophysiology 30: 32-37.
 13. Baumgartner, R., Stauber, P., McHugh, D., Koehler, K., Garry, P. (1995):
Cross-Sectional Age-Differences in Body-Composition in Persons 60+ Years of Age.
Journals of Gerontology Series A-Biological Sciences and Medical Sciences 50: M307-M316.
 14. Bear M. F., Connors B. W. und Paradiso M. A. (2009a):
Kapitel 13: Spinale Kontrolle der Motorik.
In: Engel, A. K. (Hrsg.): *Neurowissenschaften; ein grundlegendes Lehrbuch für Biologie, Medizin und Psychologie*.
Spektrum Akad. Verl., 3. Aufl., Heidelberg. S.: 469-500.
 15. Bear M. F., Connors B. W. und Paradiso M. A. (2009b):
Kapitel 14: Bewegungskontrolle durch das Gehirn.
In: Engel, A. K. (Hrsg.): *Neurowissenschaften; Ein grundlegendes Lehrbuch für Biologie, Medizin und Psychologie*.
Spektrum Akad. Verl., 3. Aufl., Heidelberg. S.: 501-533.
 16. Benito, J., Monteiro, B. P., Beauchamp, G., Lascelles, B. D. X., Steagall, P. V. (2017):
Evaluation of interobserver agreement for postoperative pain and sedation assessment in cats.
Journal of the American Veterinary Medical Association 251: 544-551.
 17. Bergadano, A., Andersen, O. K., Arendt-Nielsen, L., Schatzmann, U., Spadavecchia, C. (2006):
Quantitative assessment of nociceptive processes in conscious dogs by use of the nociceptive withdrawal reflex.
American Journal of Veterinary Research 67: 882-889.
 18. Blendis, L. M., McNeilly, W. J., Sheppard, L., Williams, R., Laws, J. W. (1970): Observer variation in the clinical and radiological assessment of hepatosplenomegaly.
British Medical Journal 1: 727-730.
 19. Boes, C. J. (2014):
The history of examination of reflexes.
Journal of Neurology 261: 2264-2274.
-

20. Böhme G., Nickel R., Schummer A. und Seiferle E. (2004a):
Armgeflecht, Plexus brachialis, und Nerven der Vordergliedmaße.
In: Nickel, R., Schummer, A. und Seiferle, E. (Hrsg.): Lehrbuch der Anatomie der
Haustiere. 4. Nervensystem, Sinnesorgane, endokrine Drüsen.
Parey, 4. Aufl., Berlin. S. 241-266.
 21. Böhme G., Nickel R., Schummer A. und Seiferle E. (2004b):
Leitungslehre des Rückenmarks.
In: Nickel, R., Schummer, A. und Seiferle, E. (Hrsg.): Lehrbuch der Anatomie der
Haustiere. 4. Nervensystem, Sinnesorgane, endokrine Drüsen.
Parey, 4. Aufl., Berlin. S. 51-61.
 22. Böhme G., Nickel R., Schummer A. und Seiferle E. (2004c):
Lendengeflecht, Plexus lumbalis, und seine Nerven für die Hintergliedmaße.
In: Nickel, R., Schummer, A. und Seiferle, E. (Hrsg.): Lehrbuch der Anatomie der
Haustiere. 4. Nervensystem, Sinnesorgane, endokrine Drüsen.
Parey, 4. Aufl., Berlin. S. 270-279.
 23. Böhme G., Nickel R., Schummer A. und Seiferle E. (2004d):
Peripheres Nervensystem.
In: Nickel, R., Schummer, A. und Seiferle, E. (Hrsg.): Lehrbuch der Anatomie der
Haustiere. 4. Nervensystem, Sinnesorgane, endokrine Drüsen.
Parey, 4. Aufl., Berlin. S. 228-298.
 24. Böhme G., Nickel R., Schummer A. und Seiferle E. (2004e):
Zentralnervensystem, Rückenmark.
In: Nickel, R., Schummer, A. und Seiferle, E. (Hrsg.): Lehrbuch der Anatomie der
Haustiere. 4. Nervensystem, Sinnesorgane, endokrine Drüsen.
Parey, 4. Aufl., Berlin. S. 28-51.
 25. Bolton, C. F., Sawa, G. M., Carter, K. (1981):
The effects of temperature on human compound action potentials.
Journal of Neurology, Neurosurgery, and Psychiatry 44: 407-413.
 26. Borin, A., Mello, L. E., Neiva, F. C., Testa, J. R., Cruz, O. L. (2012):
Experimental video analysis of eye blink reflex in a primate model.
Otolology & Neurotology 33: 1625-1629.
 27. Born, J., Hans, P., Albert, A., Bonnal, J. (1987):
Interobserver agreement in assessment of motor response and brain stem reflexes.
Neurosurgery 20: 513-517.
 28. Bortz J., Lienert G. A. (1998):
Übereinstimmungsmaße für subjektive Merkmalsbeurteilungen.
In: Bortz, J. und Lienert, G. A. (Hrsg.): Kurzgefaßte Statistik für die klinische Forschung.
Springer, 1. Aufl., Berlin. S. 265-283.
-

29. Braund K. G., Sharp J. H. (2002):
Neurological examination and localization.
In: Slatter, D. (Hrsg.): Textbook of Small Animal Surgery.
Saunders, 3. Aufl., Philadelphia. S.: 1100-1102.
 30. Bronisch, F. W. (Hrsg.) (1979):
Die Reflexe und ihre Untersuchung in Klinik und Praxis.
Thieme, 5. Aufl., Stuttgart. S. 3.
 31. Brott, T., Adams, H. P., Olinger, C. P., Marler, J. R., Barsan, W. G., Biller, J., et al. (1989):
Measurements of acute cerebral infarction: a clinical examination scale.
Stroke 20: 864-870.
 32. Buchthal, F., Pinelli, P., Rosenfalck, P. (1954):
Action Potential Parameters in Normal Human Muscle and their Physiological Determinants.
Acta Physiologica Scandinavica 32: 219-229.
 33. Budras K. D., Fricke W. und Richter R. (2004):
Autonomes Nervensystem, Aorta abdominalis, V. cava caudalis, innere Lendenmuskeln und Plexus lumbalis.
In: Budras, K. D. (Hrsg.): Atlas der Anatomie des Hundes: Lehrbuch für Tierärzte und Studierende.
Schlütersche, 7. Aufl., Hannover. S. 60.
 34. Burke, M. J., Colter, S. B. (1990):
A practical review of canine and feline spinal cord anatomy.
Progress in Veterinary Neurology 1: 358-370.
 35. Burke, D., Gandevia, S. C., McKeon, B. (1983):
The afferent volleys responsible for spinal proprioceptive reflexes in man.
The Journal of Physiology 339: 535-552.
 36. Burn, C. C., Weir, A. A. (2011):
Using prevalence indices to aid interpretation and comparison of agreement ratings between two or more observers.
The Veterinary Journal 188: 166-170.
 37. Byrt, T., Bishop, J., Carlin, J. B. (1993):
Bias, prevalence and kappa.
Journal of Clinical Epidemiology 46: 423-429.
 38. Cales, P., Burtin, P., Person, B., Ruget, O., Charlois, T., Ferrero, P., et al. (1993):
Interobserver Agreement on Endoscopic Diagnosis of Diffuse Gastric-Mucosal Features in Patients with Cirrhosis.
European Journal of Gastroenterology & Hepatology 5: 263-268.
-

39. Carlstedt, T., Misra, V. P., Papadaki, A., McRobbie, D., Anand, P. (2012):
Return of spinal reflex after spinal cord surgery for brachial plexus avulsion injury.
Journal of Neurosurgery 116: 414-417.
 40. Carswell, C., Ranopa, M., Pal, S., Macfarlane, R., Siddique, D., Thomas, D., et al. (2012):
Video Rating in Neurodegenerative Disease Clinical Trials: The Experience of PRION-1.
Dementia and Geriatric Cognitive Disorders EXTRA 2: 286-297.
 41. Chandrasekhar, A., Abu Osman, N. A., Tham, L. K., Lim, K. S., Abas, W. A. B. W. (2013):
Influence of Age on Patellar Tendon Reflex Response.
Plos One 8: e80799.
 42. Chinn, S. (1991):
Statistics in respiratory medicine. 2. Repeatability and method comparison.
Thorax 46: 454-456.
 43. Chordia, T., Vikey, A., Choudhary, A., Samdariya, Y., Chordia, D. (2016):
Current status and future trends in telepathology and digital pathology.
Journal of Oral and Maxillofacial Pathology 20: 178-182.
 44. Clarke, A. M. (1965):
Relationship between the electromyogram and the force of the isometric reflex response of normal human subjects.
Nature 208: 551-552.
 45. Clarke E., O'Malley C. D. (1996):
VI. The Reflex.
In: Clarke, E. und O'Malley, C. D. (Hrsg.): *The Human Brain and Spinal Cord: A historical study illustrated by writings from antiquity to the twentieth century.*
Norman Publishing, 2. Aufl., San Francisco. S. 323-382.
 46. Clarke, R. W., Harris, J. (2004):
The organization of motor responses to noxious stimuli.
Brain Research Reviews 46: 163-172.
 47. Clarke, R. W., Wych, B. E., Harris, J. (2001):
Adaptive changes in withdrawal reflexes after noxious stimulation at the heel and the toes in the decerebrated rabbit.
Neuroscience Letters 304: 120-122.
 48. Cohen, J. (1960):
A Coefficient of Agreement for Nominal Scales.
Educational and Psychological Measurement 20: 37-46.
-

-
49. Condliffe, E., Clark, D., Patten, C. (2005):
Reliability of elbow stretch reflex assessment in chronic post-stroke hemiparesis. *Clinical Neurophysiology* 116: 1870-1878.
 50. Cook, C., Roman, M., Stewart, K. M., Leithe, L. G., Isaacs, R. (2009):
Reliability and diagnostic accuracy of clinical special tests for myelopathy in patients seen for cervical dysfunction.
The Journal of Orthopaedic and Sports Physical Therapy 39: 172-178.
 51. Cozens, J., Miller, S., Chambers, I., Mendelow, A. (2000):
Monitoring of head injury by myotatic reflex evaluation.
Journal of Neurology, Neurosurgery, and Psychiatry 68: 581-588.
 52. Cullen, K. L., Dickey, J. P., Bent, L. R., Thomason, J. J., Moens, N. M. M., Brown, S., et al. (2015):
A validated approach for collecting fine-wire electromyographic recordings in four canine shoulder muscles during highly dynamic tasks.
Comparative Exercise Physiology 11: 65-74.
 53. Cullen, K. L., Dickey, J. P., Brown, S. H. M., Nykamp, S. G., Bent, L. R., Thomason, J. J., et al. (2017):
The magnitude of muscular activation of four canine forelimb muscles in dogs performing two agility-specific tasks.
BMC Veterinary Research 13: 68.
 54. Dafkin, C., Green, A., Kerr, S., McKinon, W. (2012):
The patellar reflex: does activity of quadriceps femoris muscles reflect leg movement?
Neurological Research 34: 623-626.
 55. Dafkin, C., Green, A., Kerr, S., Raymond, A., Veliotes, D., Elvin, A., et al. (2014):
Kinematic and kinetic analysis of the inter- and intra-applicator assessment of the Babinski reflex.
Neurophysiologie Clinique/Clinical Neurophysiology 44: 471-477.
 56. Dafkin, C., Green, A., Kerr, S., Veliotes, D., McKinon, W. (2013):
The accuracy of subjective clinical assessments of the patellar reflex.
Muscle & Nerve 47: 81-88.
 57. Dafkin, C., Green, A., Kerr, S., Veliotes, D., Olivier, B., McKinon, W. (2016):
The Interrater Reliability of Subjective Assessments of the Babinski Reflex.
Journal of Motor Behavior 48: 116-121.
 58. Davies, L. G. (1958):
Observer variation in reports on electrocardiograms.
British Heart Journal 20: 153-161.
 59. Davis, C. M., Beale, D. K. (1970):
An apparatus for isometric study of the human knee jerk.
Journal of Applied Physiology 28: 705-706.
-

60. Davis, C. M., Beaton, R. D. (1968):
Facilitation and adaptation of the human quadriceps stretch reflex produced by auditory stimulation.
Journal of Comparative and Physiological Psychology 66: 483-487.
 61. Davis, L. E., Coleman, J., Harnar, J., King, M. K. (2014):
Teleneurology: successful delivery of chronic neurologic care to 354 patients living remotely in a rural state.
Telemedicine and e-Health 20: 473-477.
 62. De Luca, C. (1997):
The use of surface electromyography in biomechanics.
Journal of Applied Biomechanics 13: 135-163.
 63. De Lahunta A., Glass E. (2009a):
Lower Motor Neuron: Spinal Nerve, General Somatic Efferent System.
In: De Lahunta, A. und Glass, E. (Hrsg.): *Veterinary Neuroanatomy and Clinical Neurology*.
Saunders Elsevier, 3. Aufl., St. Louis. S. 77-167.
 64. De Lahunta A., Glass E. (2009b):
The Neurological Examination.
In: De Lahunta, A. und Glass, E. (Hrsg.): *Veterinary Neuroanatomy and Clinical Neurology*.
Saunders Elsevier, 3. Aufl., St. Louis. S. 487-501.
 65. De Lahunta A., Glass E. (2009c):
Upper Motor Neuron.
In: De Lahunta, A. und Glass, E. (Hrsg.): *Veterinary Neuroanatomy and Clinical Neurology*.
Saunders Elsevier, 3. Aufl., St. Louis. S. 192-220.
 66. Delwaide, P. J., Toulouse, P. (1981):
Facilitation of monosynaptic reflexes by voluntary contraction of muscle in remote parts of the body. Mechanisms involved in the Jendrassik Manoeuvre.
Brain 104: 701-709.
 67. Denys, E. (1991):
Aaem Minimonograph #14: The Influence of Temperature in Clinical Neurophysiology.
Muscle & Nerve 14: 795-811.
 68. Dept. of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario (1980):
Clinical disagreement: I. How often it occurs and why.
Canadian Medical Association Journal 123: 499-504.
 69. Deschenes, M. (2004):
Effects of aging on muscle fibre type and size.
Sports Medicine 34: 809-824.
-

70. Dick, J. P. (2003):
The deep tendon and the abdominal reflexes.
Journal of Neurology, Neurosurgery, and Psychiatry 74: 150-153.
 71. Dietz, V., Grillner, S., Trepp, A., Hubli, M., Bolliger, M. (2009):
Changes in spinal reflex and locomotor activity after a complete spinal cord injury: a common mechanism?
Brain 132: 2196-2205.
 72. Dietz, V., Trippel, M., Berger, W. (1991):
Reflex Activity and Muscle Tone during Elbow Movements in Patients with Spastic Paresis.
Annals of Neurology 30: 767-779.
 73. Doubilet, P., Herman, P. G. (1981):
Interpretation of radiographs: effect of clinical history.
American Journal of Roentgenology 137: 1055-1058.
 74. Douglas, W., Malcolm, J. (1955):
The Effect of Localized Cooling on Conduction in Cat Nerves.
Journal of Physiology 130: 53-71.
 75. Dyck, P. J., Boes, C. J., Mulder, D., Millikan, C., Windebank, A. J., Dyck, P. J. B., et al. (2005):
History of standard scoring, notation, and summation of neuromuscular signs. A current survey and recommendation.
Journal of the Peripheral Nervous System 10: 158-173.
 76. Emery, F. E. (1931):
The influence of mental activity on the height of the knee-jerk.
American Journal of Physiology 97: 658-661.
 77. Emery, F. E. (1929):
Effect of the heart beat on the tonus of skeletal muscle.
American Journal of Physiology 88: 529-533.
 78. Erb, W. (1875):
Ueber Sehnenreflexe bei Gesunden und Rückenmarkskranken.
Archiv für Psychiatrie und Nervenkrankheiten 5: 792-802.
 79. Essex, M. J., Goldsmith, H. H., Smider, N. A., Dolski, I., Sutton, S. K., Davidson, R. J. (2003):
Comparison of video- and EMG-based evaluations of the magnitude of children's emotion-modulated startle response.
Behavior Research Methods 35: 590-598.
 80. Farahani, N., Pantanowitz, L. (2015):
Overview of Telepathology.
Surgical Pathology Clinics 8: 223-231.
-

81. Feinstein, A. R., Cicchetti, D. V. (1990):
High agreement but low Kappa: I. the problems of two paradoxes.
Journal of Clinical Epidemiology 43: 543-549.
 82. Fleiss, J. (1971):
Measuring Nominal Scale Agreement among Many Raters.
Psychological Bulletin 76: 378-382.
 83. Fleiss, J., Cohen, J. (1973):
Equivalence of Weighted Kappa and Intraclass Correlation Coefficient as Measures of Reliability.
Educational and Psychological Measurement 33: 613-619.
 84. Flight, L., Julious, S. A. (2015):
The disagreeable behaviour of the kappa statistic.
Pharmaceutical Statistics 14: 74-78.
 85. Forterre, F., Konar, M., Tomek, A., Doherr, M., Howard, J., Spreng, D., et al. (2008):
Accuracy of the withdrawal reflex for localization of the site of cervical disk herniation in dogs: 35 cases (2004-2007).
Journal of the American Veterinary Medical Association 232: 559-563.
 86. Fowkes, F. G. R. (1986):
Diagnostic vigilance.
The Lancet 327: 493-494.
 87. Frijns, C., Laman, D., van Duijn, M., van Duijn, H. (1997):
Normal values of patellar and ankle tendon reflex latencies.
Clinical Neurology and Neurosurgery 99: 31-36.
 88. Full, A. M., Heller, H. L. B., Mercier, M. (2016):
Prevalence, clinical presentation, prognosis, and outcome of 17 dogs with spinal shock and acute thoracolumbar spinal cord disease.
Journal of Veterinary Emergency and Critical Care 26: 412-418.
 89. Garosi L. (2013):
Lesion localization and differential diagnosis.
In: Platt, S. R. und Olby, N. J. (Hrsg.): *BSAVA Manual of Canine and Feline Neurology*.
British Small Animal Veterinary Association, 4. Aufl., Gloucester. S. 25-35.
 90. Garosi L., Lowrie M. (2013):
The neurological examination.
In: Platt, S. R. und Olby, N. J. (Hrsg.): *BSAVA Manual of Canine and Feline Neurology*.
British Small Animal Veterinary Association, 4. Aufl., Gloucester. S. 1-24.
 91. Garosi, L., Dennis, R., Penderis, J., Lamb, C., Targett, M., Cappello, R., et al. (2001):
Results of magnetic resonance imaging in dogs with vestibular disorders: 85 cases (1996-1999).
Journal of the American Veterinary Medical Association 218: 385-391.
-

92. Giebels, F., Kohn, B., Shihab, N., Volk, H., Loderstedt, S. (2014):
Evaluation of the biceps tendon reflex in dogs.
Journal of Veterinary Medicine and Research 1: 1013.
 93. Goetz, C. G., Bonduelle, M. und Gelfand, T. (Hrsg.) (1995):
Charcot: Constructing Neurology.
Oxford University Press Inc., 1. Aufl., New York.
 94. Griffiths, I. R., Duncan, I. D. (1975):
Age changes in the dorsal and ventral lumbar nerve roots of dogs.
Acta Neuropathologica 32: 75-85.
 95. Grouven, U., Bender, R., Ziegler, A., Lange, S. (2007):
Der Kappa-Koeffizient.
Deutsche Medizinische Wochenschrift 132: e65-e68.
 96. Gwet K. L. (2014):
Introduction.
In: Gwet K. L. (Hrsg.): *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters.*
Advanced Analytics, LLC, 4. Aufl., Gaithersburg. S. 3-24.
 97. Haines, D. E., Manto, M., Glickstein, M. (2007):
Clinical symptoms of cerebellar disease and their interpretation.
The Cerebellum 6: 141-156.
 98. Hansen, M., Christensen, P., Sindrup, S., Olsen, N., Kristensen, O., Friis, M. (1994):
Interobserver Variation in the Evaluation of Neurological Signs - Patient-Related Factors.
Journal of Neurology 241: 492-496.
 99. Higgins, C. (2015):
Applications and challenges of digital pathology and whole slide imaging.
Biotechnic & Histochemistry 90: 341-347.
 100. Hodges, P., Gandevia, S. (2000):
Pitfalls of intramuscular electromyographic recordings from the human costal diaphragm.
Clinical Neurophysiology 111: 1420-1424.
 101. Hoehler, F. K. (2000):
Bias and prevalence effects on kappa viewed in terms of sensitivity and specificity.
Journal of Clinical Epidemiology 53: 499-503.
 102. Holsti, O. R. (Hrsg.) (1969):
Content analysis for the social sciences and humanities.
Addison-Wesley Pub. Co., 1. Aufl., Reading, Mass.
-

103. Hongo, T., Jankowska, E., Lundberg, A. (1969):
Rubrospinal Tract .I. Effects on Alpha-Motoneurones Innervating Hindlimb Muscles in Cats.
Experimental Brain Research 7: 344-364.
 104. Hongo, T., Jankowska, E., Lundberg, A. (1972):
The Rubrospinal Tract. III. Effects on Primary Afferent Terminals.
Experimental Brain Research 15: 39-53.
 105. Houghton, J. P., Ervine, A. J., Kenny, S. L., Kelly, P. J., Napier, S. S., McCluggage, W. G., et al. (2014):
Concordance between digital pathology and light microscopy in general surgical pathology: a pilot study of 100 cases.
Journal of Clinical Pathology 67: 1052-1055.
 106. Hubbard, J. I., Jones, S. F., Landau, E. M. (1971):
The effect of change upon transmitter release, facilitation and post-tetanic potentiation.
Journal of Physiology 216: 591-609.
 107. Isaza Jaramillo, S. P., Uribe Uribe, C. S., Garcia Jimenez, F. A., Cornejo-Ochoa, W., Alvarez Restrepo, J. F., Roman, G. C. (2014):
Accuracy of the Babinski sign in the identification of pyramidal tract dysfunction. *Journal of the Neurological Sciences* 343: 66-68.
 108. Jaggy A., Spiess B. (2007):
Neurologischer Status.
In: Jaggy, A. (Hrsg.): *Atlas und Lehrbuch der Kleintierneurologie*. Schlütersche, 2. Aufl., Hannover. S. 2-34.
 109. Jang, D., Sung, I. Y., Kang, Y. J. (2013):
Usefulness of the Tendon Reflex for Assessing Spasticity After Botulinum Toxin-A Injection in Children With Cerebral Palsy.
Journal of Child Neurology 28: 21-26.
 110. Johnston, K. C., Worrall, B. B. (2003):
Teleradiology Assessment of Computerized Tomographs Online Reliability Study (TRACTORS) for Acute Stroke Evaluation.
Telemedicine Journal and e-Health 9: 227-233.
 111. Kheder, A., Nair, K. P. S. (2012):
Spasticity: pathophysiology, evaluation and management.
Practical Neurology 12: 289-298.
-

112. Kiernan, M. C., Mogyoros, I., Burke, D. (1999):
Value of homonymous and heteronymous monosynaptic reflexes in the diagnosis and follow-up of cervical spinal injuries.
Journal of Clinical Neuroscience 6: 24-26.
 113. Klein B. G., Cunningham J. G. (2012):
Chapter 7: The Concept of a Reflex.
In: Cunningham, J. G. (Hrsg.): *Cunningham's textbook of Veterinary Physiology*. Elsevier Saunders, 5. Aufl., St. Louis, Mo. S. 77-80.
 114. Koran, L. M. (1975):
The reliability of clinical methods, data and judgments (first of two parts).
The New England Journal of Medicine 293: 642-646.
 115. Kornberg, M., Bley, T. und von Klopmann, T. (2012):
Prüfung der spinalen Reflexe.
In: Suter, P. F., Kohn, B. und Schwarz, G. (Hrsg.): *Praktikum der Hundeklinik*. Enke, 11. Aufl., Stuttgart. S. 933-935.
 116. Kottner, J., Halfens, R., Dassen, T. (2009a):
An interrater reliability study of the assessment of pressure ulcer risk using the Braden scale and the classification of pressure ulcers in a home care setting.
International Journal of Nursing Studies 46: 1307-1312.
 117. Kottner, J., Tannen, A., Halfens, R., Dassen, T. (2009b):
Does the number of raters influence the pressure ulcer prevalence rate?
Applied Nursing Research 22: 68-72.
 118. Landis, J. R., Koch, G. G. (1977):
The Measurement of Observer Agreement for Categorical Data.
Biometrics 33: 159-174.
 119. Lanska, D. (1989):
The History of Reflex Hammers.
Neurology 39: 1542-1549.
 120. Lanska, D. (1995a):
The Dana Reflex Hammer (C.1892).
Journal of Child Neurology 10: 367-368.
 121. Lanska, D. (1995b):
The Stookey Reflex Hammer.
Journal of Child Neurology 10: 23-24.
 122. Lanska, D. (1999):
The Babinski reflex hammer.
Neurology 53: 655-655.
-

123. Lanska, D., Lanska, M. (1990):
John Madison Taylor (1855-1931) and the First Reflex Hammer.
Journal of Child neurology 5: 38-39.
 124. Larsson, L. (1978):
Morphological and functional characteristics of the ageing skeletal muscle in man. A cross-sectional study.

Acta Physiologica Scandinavica. Supplementum 457: 1-36.
 125. Lebedowska, M., Fisk, J. (2003):
Quantitative evaluation of reflex and voluntary activity in children with spasticity.
Archives of Physical Medicine and Rehabilitation 84: 828-837.
 126. LeBreton, J. M., Senter, J. L. (2008):
Answers to 20 Questions About Interrater Reliability and Interrater Agreement.
Organizational Research Methods 11: 815-852.
 127. Lee, R. A., Tay, L. L., Lim, K. S., Tan, C. (2011):
Variability of plantar response in normal population.
Neurology Asia 16: 143-147.
 128. Lehmann, G. J., McGill, S. M. (1999):
The importance of normalization in the interpretation of surface electromyography: A proof of principle.
Journal of Manipulative and Physiological Therapeutics 22: 444-446.
 129. Lemoyne, R., Dabiri, F., Jafari, R. (2008):
Quantified deep tendon reflex device, second generation.
Journal of Mechanics in Medicine and Biology 8: 75-85.
 130. Lervik, A., Haga, H. A., Ranheim, B., Spadavecchia, C. (2012):
The influence of a continuous rate infusion of dexmedetomidine on the nociceptive withdrawal reflex and temporal summation during isoflurane anaesthesia in dogs.
Veterinary Anaesthesia and Analgesia 39: 414-425.
 131. Levine, J. M., Hillman, R. B., Hollis, N. E., De Lahunta, A. (2002):
The Influence of Age on Patellar Reflex Response in the Dog.
Journal of Veterinary Internal Medicine 16: 244-246.
 132. Lewis, G. N., Rice, D. A., Jourdain, K., McNair, P. J. (2012):
Influence of stimulation location and posture on the reliability and comfort of the nociceptive flexion reflex.
Pain Research and Management 17: 110-114.
 133. Liddell, E. G. T., Sherrington, C. S. (1925):
Recruitment and some other Features of Reflex Inhibition.
Proceedings of the Royal Society of London. Series B 97: 488-518.
-

-
134. Lin, F., Sabbahi, M. (1999):
Correlation of spasticity with hyperactive stretch reflexes and motor dysfunction in hemiplegia.
Archives of Physical Medicine and Rehabilitation 80: 526-530.
 135. Litvan, I., Mangone, C. A., Werden, W., Bueri, J. A., Estol, C. J., Garcea, D. O., et al. (1996):
Reliability of the NINDS Myotatic Reflex Scale.
Neurology 47: 969-972.
 136. Maher, J., Reilly, M., Daly, L., Hutchinson, M. (1992):
Plantar power: reproducibility of the plantar response.
British Medical Journal 304: 482-482.
 137. Mahler, S. P., Adogwa, A. O. (2008):
Anatomical and experimental studies of brachial plexus, sciatic, and femoral nerve-location using peripheral nerve stimulation in the dog.
Veterinary Anaesthesia and Analgesia 35: 80-89.
 138. Manschot, S., van Passel, L., Buskens, E., Algra, A., van Gijn, J. (1998):
Mayo and NINDS scales for assessment of tendon reflexes: between observer agreement and implications for communication.
Journal of Neurology, Neurosurgery, and Psychiatry 64: 253-255.
 139. Martin, W. R., Eades, C. G., Fraser, H. F., Wikler, A. (1964):
Use of Hindlimb Reflexes of the Chronic Spinal Dog for Comparing Analgesics.
The Journal of Pharmacology and Experimental Therapeutics 144: 8-11.
 140. Meade, T. W., Gardner, M. J., Cannon, P., Richardson, P. C. (1968):
Observer variability in recording the peripheral pulses.
British Heart Journal 30: 661-665.
 141. Mihelic, D., Mihelic, D., Slavica, A., Dezdek, D., Trbojevic-Vukicevic, T., Dzaja, P., et al. (2007):
N. femoralis, N. obturatorius and N. ischiadicus: deviation in creation in the dogs.
Anatomia, Histologia, Embryologia 36: 401-407.
 142. Miller, T. A., Mogyoros, I., Burke, D. (1995a):
Homonymous and Heteronymous Monosynaptic Reflexes in Biceps Brachii.
Muscle & Nerve 18: 585-592.
 143. Miller, T. A., Mogyoros, I., Kiernan, M., Burke, D. (1995b):
Reproducibility of a heteronymous monosynaptic reflex in biceps brachii.
Electroencephalography and Clinical Neurophysiology/Electromyography and Motor Control 97: 318-325.
-

144. Miller M. E., Christensen G. C. und Evans H. E. (1979):
Chapter 11. The Spinal Nerves.
In: Evans, H. E. und De Lahunta, A. (Hrsg.): Miller's Anatomy of the dog.
Saunders, 2. Aufl., Philadelphia. S. 572-625.
 145. Miller, T. M., Johnston, S. C. (2005):
Should the Babinski sign be part of the routine neurologic examination?
Neurology 65: 1165-1168.
 146. Mitchell, S. W., Lewis, M. J. (1886):
The tendon-jerk and muscle-jerk in disease, and especially in posterior sclerosis.
The American Journal of the Medical Sciences 184: 363-372.
 147. Monticelli, P., Drozdzyńska, M., Stathopoulou, T., Neilson, D., Gregori, T., Viscasillas, J. (2016):
A description of a technique for ultrasound-guided lumbar plexus catheter in dogs: cadaveric study.
Veterinary Anaesthesia and Analgesia 43: 453-456.
 148. Muguet-Chanoit, A., Babb, K., Olby, N. J. (2010):
The Cutaneous Trunci Muscle Reflex in Normal Dogs.
Journal of Veterinary Internal Medicine 24: 744-744.
 149. Muguet-Chanoit, A. C., Olby, N. J., Lim, J., Gallagher, R., Niman, Z., Dillard, S., et al. (2012):
The Cutaneous Trunci Muscle Reflex: A Predictor of Recovery in Dogs with Acute Thoracolumbar Myelopathies Caused by Intervertebral Disc Extrusions.
Veterinary Surgery 41: 200-206.
 150. Murakami, T., Feeney, D. A., Willey, J. L., Carlin, B. P. (2014):
Evaluation of the accuracy of neurologic data, survey radiographic results, or both for localization of the site of thoracolumbar intervertebral disk herniation in dogs. American Journal of Veterinary Research 75: 251-259.
 151. Neil, D. A., Demetris, A. J. (2014):
Digital pathology services in acute surgical situations.
The British Journal of Surgery 101: 1185-1186.
 152. Nickel R., Schummer A. und Seiferle E. (2004a):
Eigenmuskulatur der Beckengliedmaße der Fleischfresser.
In: Frewein, J., Willes, K. und Wilkens, H. (Hrsg.): Lehrbuch der Anatomie der Haustiere. 1. Bewegungsapparat.
Parey, 8. Aufl., Berlin. S. 517-530.
 153. Nickel R., Schummer A. und Seiferle E. (2004b):
Eigenmuskulatur der Schultergliedmaße.
In: Frewein, J., Willes, K. und Wilkens, H. (Hrsg.): Lehrbuch der Anatomie der Haustiere. 1. Bewegungsapparat.
Parey, 8. Aufl., Berlin. S. 438-492.
-

-
154. Nielsen, J. F., Sinkjaer, T. (1996):
A comparison of clinical and laboratory measures of spasticity.
Multiple sclerosis (Houndmills, Basingstoke, England) 1: 296-301.
 155. Nuzzo, J. L., Trajano, G. S., Barry, B. K., Gandevia, S. C., Taylor, J. L. (2016):
Arm posture-dependent changes in corticospinal excitability are largely spinal in origin.
Journal of Neurophysiology 115: 2076-2082.
 156. Nyberg-Hansen, R., Brodal, A. (1963):
Sites of termination of corticospinal fibers in the cat. An experimental study with silver impregnation methods.
The Journal of Comparative Neurology 120: 369-391.
 157. Okada, M., Kitagawa, M., Ito, D., Itou, T., Kanayama, K., Sakai, T. (2010):
Magnetic resonance imaging features and clinical signs associated with presumptive and confirmed progressive myelomalacia in dogs: 12 cases (1997-2008).
Journal of the American Veterinary Medical Association 237: 1160-1165.
 158. O'Keeffe, S. T., Smith, T., Valacio, R., Jack, C. I. A., Playfer, J. R., Lye, M. (1994):
A comparison of two techniques for ankle jerk assessment in elderly subjects.
The Lancet 344: 1619-1620.
 159. Olby, N. J., Lim, J., Babb, K., Bach, K., Domaracki, C., Williams, K., et al. (2014):
Gait scoring in dogs with thoracolumbar spinal cord injuries when walking on a treadmill.
BMC Veterinary Research 10: 58.
 160. Oliver J. E., Lorenz M. D. und Kornegay J. N. (1997):
Fundamentals.
In: Oliver, J. E., Lorenz, M. D. und Kornegay, J. N. (Hrsg.): Handbook of Veterinary Neurology.
Saunders, 3. Aufl., Philadelphia. S.: 24-25.
 161. Packer, R. M. A., Berendt, M., Bhatti, S., Charalambous, M., Cizinauskas, S., De Risio, L., et al. (2015):
Interobserver agreement of canine and feline paroxysmal event semiology and classification by veterinary neurology specialists and non-specialists.
BMC Veterinary Research 11: 39.
 162. Pagliaro, P., Zamparo, P. (1999):
Quantitative evaluation of the stretch reflex before and after hydro kinesy therapy in patients affected by spastic paresis.
Journal of Electromyography and Kinesiology 9: 141-148.
 163. Pascoe, P. J. (2016):
The effects of lidocaine or a lidocaine-bupivacaine mixture administered into the infraorbital canal in dogs.
American Journal of Veterinary Research 77: 682-687.
-

164. Pearson K., Jordan G. (2000):
Spinal Reflexes.
In: Kandel, E. R., Schwartz, J. H. und Jessell, T. M. (Hrsg.): Principles of Neural Science. McGraw-Hill, 4. Aufl., New York, NY. S. 713-735.
165. Pereira, J. R. (1925):
Physiological studies on the patellar reflex. I.-The Patellometer.
Brain 48: 255-258.
166. Péréon, Y., Tich, S., Fournier, E., Genet, R., Guiheneuc, P. (2004):
Electrophysiological recording of deep tendon reflexes: normative data in children and in adults.
Neurophysiologie Clinique/Clinical Neurophysiology 34: 131-139.
167. Perez, M. A., Field-Fote, E. C. (2003):
Impaired posture-dependent modulation of disynaptic reciprocal Ia inhibition in individuals with incomplete spinal cord injury.
Neuroscience Letters 341: 225-228.
168. Peterson, C. L., Rogers, L. M., Mogk, J. P., Bednar, M. S., Bryden, A. M., Keith, M. W., et al. (2014):
Posture-dependent changes in corticomotor excitability of the biceps after spinal cord injury and tendon transfer.
Conference proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference 2014: 4302-4305.
169. Petrie A., Watson P. (2013):
Appendix C.
In: Petrie A., Watson P. (Hrsg.): Statistics for Veterinary and Animal Science. Wiley-Blackwell, 1. Aufl., West Sussex. S. 343.
170. Pisano, F., Miscio, G., Del Conte, C., Pianca, D., Candeloro, E., Colombo, R. (2000):
Quantitative measures of spasticity in post-stroke patients.
Clinical Neurophysiology 111: 1015-1022.
171. Portney, L. G., Watkins, M. P. (Hrsg.) (2015):
Foundations of clinical research: applications to practice.
FA Davis, 3. Aufl., Philadelphia.
172. Pschyrembel, W., Dornblüth, O. (Hrsg.) (2002):
Klinisches Wörterbuch.
Walter de Gruyter, 259. Aufl., Berlin.
173. Raftery, E. B., Holland, W. W. (1967):
Examination of the heart: an investigation into variation.
American Journal of Epidemiology 85: 438-444.
-

-
174. Raijmakers, P. G. H. M., Cabezas, M. C., Smal, J. A., Van Gijn, J. (1991):
Teaching the Plantar Reflex.
Clinical Neurology and Neurosurgery 93: 201-204.
 175. Rampy, B. A., Glassy, E. F. (2015):
Pathology Gross Photography: The Beginning of Digital Pathology.
Surgical Pathology Clinics 8: 195-211.
 176. Randolph, J. J. (2005):
Free-Marginal Multirater Kappa (multirater K_{free}): An Alternative to Fleiss' Fixed-Marginal Multirater Kappa.
Proceedings of the Joensuu Learning and Instruction Symposium. S. 1-20.
 177. Reetz, J. A., Caceres, A. V., Suran, J. N., Oura, T. J., Zwingenberger, A. L., Mai, W. (2013):
Sensitivity, positive predictive value, and interobserver variability of computed tomography in the diagnosis of bullae associated with spontaneous pneumothorax in dogs: 19 cases (2003-2012).
Journal of the American Veterinary Medical Association 243: 244-251.
 178. Robinson, P., Wilson, D., Coral, A., Murphy, A., Verow, P. (1999):
Variation between experienced observers in the interpretation of accident and emergency radiographs.
British Journal of Radiology 72: 323-330.
 179. Rohrbach, H., Andersen, O. K., Zeiter, S., Wieling, R., Spadavecchia, C. (2015):
Repeated electrical stimulations as a tool to evoke temporal summation of nociceptive inputs in healthy, non-medicated experimental sheep.
Physiology & Behaviour 142: 85-89.
 180. Rohrbach, H., Korpivaara, T., Schatzmann, U., Spadavecchia, C. (2009):
Comparison of the effects of the alpha-2 agonists detomidine, romifidine and xylazine on nociceptive withdrawal reflex and temporal summation in horses.
Veterinary Anaesthesia and Analgesia 36: 384-395.
 181. Rohrbach, H., Zeiter, S., Andersen, O. K., Wieling, R., Spadavecchia, C. (2014):
Quantitative assessment of the nociceptive withdrawal reflex in healthy, non-medicated experimental sheep.
Physiology & Behaviour 129: 181-185.
 182. Rossi, A., Mazzocchio, R., Scarpini, C. (1988):
Changes in Ia reciprocal inhibition from the peroneal nerve to the soleus alpha-motoneurons with different static body positions in man.
Neuroscience Letters 84: 283-286.
-

183. Rousing, T., Waiblinger, S. (2004):
Evaluation of on-farm methods for testing the human–animal relationship in dairy herds with cubicle loose housing systems-test-retest and inter-observer reliability and consistency to familiarity of test person.
Applied Animal Behaviour Science 85: 215-231.
184. Rutkove, S. (2001):
Effects of temperature on neuromuscular electrophysiology.
Muscle & Nerve 24: 867-882.
185. Salomon F., Geyer H. (2007):
Hintergliedmaßen des Hundes.
In: Salomon, F. und Geyer, H. (Hrsg.): *Atlas der angewandten Anatomie der Haustiere*. Enke Verlag, 3. Aufl., Stuttgart. S. 238-259.
186. Sargeant, J. M., Martin, S. W. (1998):
The dependence of kappa on attribute prevalence when assessing the repeatability of questionnaire data.
Preventive Veterinary Medicine 34: 115-123.
187. Sawaya, S. G., Combet, D., Chanoit, G., Thiebault, J., Levine, D., Marcellin-Little, D. J. (2008):
Assessment of impulse duration thresholds for electrical stimulation of muscles (chronaxy) in dogs.
American Journal of Veterinary Research 69: 1305-1309.
188. Schatzberg S. J. (2010):
Neurological examination and neuroanatomic diagnosis.
In: Ettinger, S. J. und Feldman, E. C. (Hrsg.): *Textbook of veterinary internal medicine; Diseases of the dog and the cat*. Saunders, 7. Aufl., St. Louis. S. 1404-1405.
189. Schatzberg S. J., Kent M. und Platt S. R. (2012):
Neurologic Examination and Neuroanatomic Diagnosis.
In: Tobias, K. M. und Johnston, S. A. (Hrsg.): *Veterinary Surgery: Small Animal*. 1. Elsevier Saunders, 1. Aufl., St. Louis. S. 325-339.
190. Schmid, A. B., Brunner, F., Luomajoki, H., Held, U., Bachmann, L. M., Kunzer, S., et al. (2009):
Reliability of clinical tests to evaluate nerve function and mechanosensitivity of the upper limb peripheral nervous system.
BMC Musculoskeletal Disorders 10: 11-2474-10-11.
191. Schomburg, E. D. (1990):
Spinal functions in sensorimotor control of movements.
Neurosurgical Review 13: 179-185.
-

-
192. Seim H. B. (2009):
Allgemeine Neurochirurgie.
In: Fossum, T. W. und Duprey, L. P. (Hrsg.): Chirurgie der Kleintiere.
Elsevier, Urban & Fischer, 2. Aufl., München. S. 1423-1425.
 193. Sharp, J. W., Bailey, C. S., Johnson, R. D., Kitchell, R. L. (1990):
Spinal Nerve Root Origin of the Median, Ulnar and Musculocutaneous Nerves and their
Muscle Nerve Branches to the Canine Forelimb.
Anatomia, Histologia, Embryologia 19: 359-368.
 194. Sharp N. J. H., Wheeler S. J. (2005):
Patient examination.
In: Sharp, N. J. H. und Wheeler, S. J. (Hrsg.): Small animal spinal disorders.
Elsevier, 2. Aufl., Edinburgh. S. 19-35.
 195. Sim, J., Wright, C. (2005):
The kappa statistic in reliability studies: Use, interpretation, and sample size
requirements.
Physical Therapy 85: 257-268.
 196. Simons, D. G., Dimitrijevic, M. R. (1972):
Quantitative variations in the force of quadriceps responses to serial patellar tendon taps
in normal man.
American Journal of Physical Medicine 51: 240-263.
 197. Singerman, J., Lee, L. (2008):
Consistency of the Babinski reflex and its variants.
European Journal of Neurology 15: 960-964.
 198. Smith, P., Jeffery, N. (2005):
Spinal shock-comparative aspects and clinical relevance.
Journal of Veterinary Internal Medicine 19: 788-793.
 199. Spadavecchia, C., Arendt-Nielsen, L., Andersen, O. K., Spadavecchia, L., Doherr, M.,
Schatzmann, U. (2003):
Comparison of nociceptive withdrawal reflexes and recruitment curves between the
forelimbs and hind limbs in conscious horses.
American Journal of Veterinary Research 64: 700-707.
 200. Spadavecchia, C., Spadavecchia, L., Andersen, O. K., Arendt-Nielsen, L., Leandri, M.,
Schatzmann, U. (2002):
Quantitative assessment of nociception in horses by use of the nociceptive withdrawal
reflex evoked by transcutaneous electrical stimulation.
American Journal of Veterinary Research 63: 1551-1556
 201. Stam, J., Tan, K. M. (1987):
Tendon reflex variability and method of stimulation.
Electroencephalography and Clinical Neurophysiology 67: 463-467.
-

203. Stam, J., van Crevel, H. (1989):
Measurement of Tendon Reflexes by Surface Electromyography in Normal Subjects.
Journal of Neurology 236: 231-237.
 204. Stam, J., van Crevel, H. (1990):
Reliability of the Clinical and Electromyographic Examination of Tendon Reflexes.
Journal of Neurology 237: 427-431.
 205. Stam, J., Vanleeuwen, J. (1984):
A Simple Measurement Hammer for Quantitative Reflex Studies.
Electroencephalography and Clinical Neurophysiology 58: 282-284.
 206. Stoffel M. H. (2011a):
Das somatische Nervensystem: Wechselwirkung mit der Außenwelt.
In: Kocher, M. und Guldemann, C. (Hrsg.): *Funktionelle Neuroanatomie für die Tiermedizin*.
Enke, 1. Aufl., Stuttgart. S. 22-25.
 207. Stoffel M. H. (2011b):
Somatisches Nervensystem und höhere Sinne.
In: Kocher, M. und Guldemann, C. (Hrsg.): *Funktionelle Neuroanatomie für die Tiermedizin*.
Enke, 1. Aufl., Stuttgart. S. 122-165.
 208. Thaller, M., Hughes, T. (2014):
Inter-rater agreement of observable and elicitable neurological signs.
Clinical Medicine (London, England) 14: 264-267.
 209. Thomas W. B., Dewey C. W. (2008):
Performing the neurological examination.
In: Dewey, C. W. (Hrsg.): *A Practical Guide to Canine & Feline Neurology*.
Wiley Blackwell, 2. Aufl., Ames, Iowa. S. 66-68.
 210. Tomasello, F., Mariani, F., Fieschi, C., Argentino, C., Bono, G., De Zanche, L., et al.
(1982):
Assessment of inter-observer differences in the Italian multicenter study on reversible
cerebral ischemia.
Stroke 13: 32-35.
 211. Torbørgsen, T., Stalberg, E. (1986):
Effects of Temperature on Muscle Twitch.
Muscle & Nerve 9: 574-574.
 212. Tudor, G., Finlay, D., Taub, N. (1997):
An assessment of inter-observer agreement and accuracy when reporting plain
radiographs.
Clinical Radiology 52: 235-238.
-

-
213. Tudury, E. A., Araújo B. M., Fernandes, T. H. T., et al. (2013):
Carpi radialis and cranial tibial reflexes: myotatic or muscular response?
Proceedings of the 78th International SCIVAC Congress; 2013; 31. Mai - 2. Juni; S. 576.
 214. Tudury, E. A., de Figueiredo, M. L., Tavares Fernandes, T. H., Araujo, B. M., Bonelli, M. d. A., Diogo, C. C., et al. (2017):
Evaluation of cranial tibial and extensor carpi radialis reflexes before and after anesthetic block in cats.
Journal of Feline Medicine and Surgery 19: 105-109.
 215. Tuttle, W. (1924):
The Effect of Attention or Mental Activity on the Patellar Tendon Reflex.
Journal of Experimental Psychology 7: 401.
 216. Van Gijn, J., Bonke, B. (1977):
Interpretation of Plantar Reflexes - Biasing Effect of Other Signs and Symptoms. Journal of Neurology, Neurosurgery, and Psychiatry 40: 787-789.
 217. Villamonte-Chevalier, A., Soler, M., Sarria, R., Agut, A., Latorre, R. (2012): Anatomical study of fibrous structures of the medial aspect of the canine elbow joint. Veterinary Record 171: 596.
 218. Vincent, W., Weir, J. (Hrsg.) (1999):
Statistics in kinesiology.
Human Kinetics, 2. Aufl., Champaign, USA.
 219. Voerman, G., Gregoric, M., Hermens, H. (2005):
Neurophysiological methods for the assessment of spasticity: The Hoffmann reflex, the tendon reflex, and the stretch reflex.
Disability and Rehabilitation 27: 33-68.
 220. Vogel, H. (1992):
Influence of Additional Information on Interrater Reliability in the Neurologic Examination.
Neurology 42: 2076-2081.
 221. Von Engelhardt, W. (Hrsg.) (2005):
Physiologie der Haustiere.
Enke, 2. Aufl., Stuttgart.
 222. Vreeling, F. W., Jolles, J., Verhey, F. R. J., Houx, P. J. (1993):
Primitive Reflexes in Healthy, Adult Volunteers and Neurological Patients - Methodological Issues.
Journal of Neurology 240: 495-504.
-

223. Vroomen, P. C., de Krom, M. C., Knottnerus, J. A. (2000):
Consistency of history taking and physical examination in patients with suspected lumbar nerve root involvement.
Spine 25: 91-97.
224. Wechsler, L. R. (2015):
Advantages and limitations of teleneurology.
JAMA Neurology 72: 349-354.
225. Wechsler, L. R., Demaerschalk, B. M., Schwamm, L. H., Adeoye, O. M., Audebert, H. J., Fanale, C. V., et al. (2017):
Telemedicine Quality and Outcomes in Stroke: A Scientific Statement for Healthcare Professionals From the American Heart Association/American Stroke Association.
Stroke 48: e3-e25.
226. Westphal, C. (1875):
Ueber einige Bewegungs-Erscheinungen an gelähmten Gliedern.
Archiv für Psychiatrie und Nervenkrankheiten 5: 803-834.
227. Wilson, J. (1978):
Relationship of Patellar Tendon Reflex to Ventral Branch of 5th Lumbar Spinal Nerve in Dog.
American Journal of Veterinary Research 39: 1774-1777.
228. Yager, P. H., Clark, M. E., Dapul, H. R., Murphy, S., Zheng, H., Noviski, N. (2014):
Reliability of circulatory and neurologic examination by telemedicine in a pediatric intensive care unit.
The Journal of Pediatrics 165: 962-966.e5.
229. Yager, P. H., Cummings, B. M., Whalen, M. J., Noviski, N. (2012):
Nighttime telecommunication between remote staff intensivists and bedside personnel in a pediatric intensive care unit: a retrospective study.
Critical Care Medicine 40: 2700-2703.
230. Ziff, M., Stark, R. J. (2017):
How to assess tendon reflexes of the lower limb in the elderly.
Journal of the Neurological Sciences 372: 196-200.
-

9. Anhang

9.1. Evaluierungsbögen

Tabelle 29. Evaluierungsbogen für Studie 1 und 2a.

Code	Presence		Grading				
	1 = yes	0 = no	0 (absent)	1 (reduced)	2 (normal)	3 (increased)	4 (tonic)
1B							
2A							
2B							
3A							
3B							
4A							
4B							
10A							
10B							
11A							
11B							
13A							
13B							
18A							
18B							
21A							
22B							
23A							
23B							
24B							
25B							
27A							
27B							
31A							
31B							
32A							
32B							
33A							
33B							
34B							
36A							
36B							
37A							
37B							
41A							

Code	Presence		Grading				
	1 = yes	0 = no	0 (absent)	1 (reduced)	2 (normal)	3 (increased)	4 (tonic)
41B							
42A							
42B							
44B							
45A							
45B							
47B							
48A							
48B							
49B							
51A							
51B							
54A							
58A							
59A							
59B							
60A							
60B							
63A							
63B							
65A							
65B							
66A							
66B							
68B							
69A							
69B							
70A							
70B							
71A							
71B							
72A							
72B							
73A							
73B							
74A							
74B							
77A							
77B							
81A							
81B							

Code	Presence		Grading				
	1 = yes	0 = no	0 (absent)	1 (reduced)	2 (normal)	3 (increased)	4 (tonic)
82B							
83B							
84A							
84B							
85A							
86A							
86B							
87A							
87B							
89A							
89B							
90A							
90B							
91A							
93A							
93B							
94A							
94B							
95B							
96B							
97A							
97B							
98B							
99A							
99B							
100A							
100B							

Tabelle 30. Evaluierungsbogen für Studie 2b.

Code	Presence		Grading				
	1 = yes	0 = no	0 (absent)	1 (reduced)	2 (normal)	3 (increased)	4 (tonic)
1							
2							
3							
4							
5							
7							
9							
11							
12							
14							
15							
17							
18							
19							
23							
24							
25							
27							
28							
29							
31							
32							
33							
34							
35							
36							
37							
38							
39							
41							
42							
47							
48							
50							
51							
52							
54							
55							
57							

Code	Presence		Grading				
	1 = yes	0 = no	0 (absent)	1 (reduced)	2 (normal)	3 (increased)	4 (tonic)
58							
62							
63							
66							
67							
68							
70							
72							
73							
74							
75							
76							
77							
78							
79							
84							
85							
87							
89							
92							
93							
94							
96							
98							
100							

9.2. Evaluierungen der Beurteiler in Studie 1

Tabelle 31. Evaluierungen der Reflexpräsenz und -aktivität durch die vier Neurologen (N 1-4) in Studie 1. (Reflexpräsenz: 0 = abwesend, 1 = vorhanden; Reflexaktivität: 0 = abwesend, 1 = reduziert, 2 = normal, 3 = gesteigert, 4 = tonisch).

Nr.	Code	Reflexpräsenz				Reflexaktivität			
		N 1	N 2	N 3	N 4	N 1	N 2	N 3	N 4
1.	1B	1	1	1	1	2	2	2	2
2.	2A	1	1	1	1	2	2	2	2
3.	2B	1	1	1	1	2	1	1	2
4.	3A	1	1	1	1	2	2	1	2
5.	3B	1	1	1	1	3	3	3	3
6.	4A	1	1	0	1	2	1	0	2
7.	4B	1	1	1	1	2	2	1	2
8.	10A	1	1	1	1	3	3	3	3
9.	10B	1	1	1	1	2	1	1	2
10.	11A	1	0	1	1	3	0	2	3
11.	11B	1	0	1	1	2	0	1	2
12.	13A	1	1	1	1	3	2	3	2
13.	13B	1	1	1	1	3	1	2	2
14.	18A	1	1	1	1	2	2	1	2
15.	18B	1	1	1	1	2	2	1	2
16.	21A	1	1	1	1	1	1	1	2
17.	22B	1	1	1	1	3	3	2	4
18.	23A	1	1	0	1	2	2	0	2
19.	23B	1	1	0	1	1	2	0	1
20.	24B	1	1	1	1	2	3	2	2
21.	25B	1	1	0	1	3	3	0	2
22.	27A	1	1	1	1	2	2	1	2
23.	27B	1	1	1	1	2	2	1	2
24.	31A	1	1	1	1	2	2	2	2
25.	31B	1	1	1	1	2	2	2	2
26.	32A	1	0	1	1	2	0	3	2

Nr.	Code	Reflexpräsenz				Reflexaktivität			
		N 1	N 2	N 3	N 4	N 1	N 2	N 3	N 4
57.	65B	1	1	1	1	2	1	1	2
58.	66A	1	1	1	1	2	3	2	2
59.	66B	1	1	1	1	2	2	2	2
60.	68B	1	1	1	1	1	4	1	2
61.	69A	1	1	1	1	2	3	2	2
62.	69B	1	1	1	1	2	3	2	2
63.	70A	1	1	1	1	2	2	2	2
64.	70B	1	1	1	1	2	2	2	2
65.	71A	1	1	1	1	1	1	2	2
66.	71B	1	1	1	1	1	2	1	2
67.	72A	1	1	1	1	2	2	2	2
68.	72B	1	1	1	1	3	2	2	4
69.	73A	1	1	1	1	2	3	2	4
70.	73B	1	1	1	1	2	3	2	4
71.	74A	1	1	1	1	2	3	2	2
72.	74B	1	1	1	1	2	3	1	2
73.	77A	1	1	0	1	2	4	0	2
74.	77B	1	1	1	1	2	3	2	2
75.	81A	1	1	1	1	3	2	2	2
76.	81B	1	1	1	1	3	2	1	2
77.	82B	1	1	1	1	2	3	3	3
78.	83B	1	1	1	1	2	2	2	2
79.	84A	1	1	1	1	2	2	2	2
80.	84B	1	1	1	1	2	2	1	2
81.	85A	1	1	1	1	2	2	2	1
82.	86A	1	1	1	1	2	2	2	1
83.	86B	1	1	1	1	2	2	2	2
84.	87A	1	1	1	1	2	2	3	2
85.	87B	1	1	1	1	3	3	3	2
86.	89A	1	1	1	1	2	2	3	2

Tabelle 32. Evaluierungen der Reflexpräsenz und -aktivität durch die vier praktizierenden Tierärzte (T 1-4) in Studie 1. (Reflexpräsenz: 0 = abwesend, 1 = vorhanden; Reflexaktivität: 0 = abwesend, 1 = reduziert, 2 = normal, 3 = gesteigert, 4 = tonisch).

Nr.	Code	Reflexpräsenz				Reflexaktivität			
		T 1	T 2	T 3	T 4	T 1	T 2	T 3	T 4
1.	1B	1	1	1	1	2	2	2	2
2.	2A	1	1	1	1	2	2	2	2
3.	2B	1	1	1	1	2	1	1	2
4.	3A	1	1	1	1	2	2	1	2
5.	3B	1	1	1	1	3	3	3	3
6.	4A	1	1	0	1	2	1	0	2
7.	4B	1	1	1	1	2	2	1	2
8.	10A	1	1	1	1	3	3	3	3
9.	10B	1	1	1	1	2	1	1	2
10.	11A	1	0	1	1	3	0	2	3
11.	11B	1	0	1	1	2	0	1	2
12.	13A	1	1	1	1	3	2	3	2
13.	13B	1	1	1	1	3	1	2	2
14.	18A	1	1	1	1	2	2	1	2
15.	18B	1	1	1	1	2	2	1	2
16.	21A	1	1	1	1	1	1	1	2
17.	22B	1	1	1	1	3	3	2	4
18.	23A	1	1	0	1	2	2	0	2
19.	23B	1	1	0	1	1	2	0	1
20.	24B	1	1	1	1	2	3	2	2
21.	25B	1	1	0	1	3	3	0	2
22.	27A	1	1	1	1	2	2	1	2
23.	27B	1	1	1	1	2	2	1	2
24.	31A	1	1	1	1	2	2	2	2
25.	31B	1	1	1	1	2	2	2	2
26.	32A	1	0	1	1	2	0	3	2
27.	32B	1	1	1	1	2	3	3	2

Nr.	Code	Reflexpräsenz				Reflexaktivität			
		T 1	T 2	T 3	T 4	T 1	T 2	T 3	T 4
28.	33A	1	1	0	1	2	1	0	2
29.	33B	0	0	0	1	0	0	0	2
30.	34B	1	1	1	1	3	3	2	2
31.	36A	1	1	1	1	1	1	1	1
32.	36B	1	1	1	1	2	2	2	2
33.	37A	1	1	1	1	2	2	3	2
34.	37B	1	1	1	1	2	3	3	4
35.	41A	1	1	1	1	2	3	1	2
36.	41B	1	0	0	1	1	0	0	2
37.	42A	1	0	0	0	1	0	0	2
38.	42B	1	1	1	1	2	1	2	2
39.	44B	1	0	1	1	1	0	1	2
40.	45A	1	1	1	1	2	2	2	2
41.	45B	1	1	1	1	2	2	3	2
42.	47B	0	0	0	0	0	0	0	0
43.	48A	1	1	1	1	2	2	2	2
44.	48B	1	1	1	1	3	3	3	3
45.	49B	1	0	1	1	2	0	1	2
46.	51A	1	1	1	1	2	2	2	2
47.	51B	1	1	1	1	2	2	1	2
48.	54A	1	1	1	1	2	1	1	2
49.	58A	1	1	1	1	2	2	2	2
50.	59A	1	1	0	1	2	1	0	2
51.	59B	1	1	1	1	2	2	2	2
52.	60A	1	1	1	1	3	3	3	3
53.	60B	1	1	1	1	3	2	2	2
54.	63A	1	1	1	1	2	2	2	4
55.	63B	1	0	1	1	1	0	1	4
56.	65A	1	1	1	1	1	1	2	2
57.	65B	1	1	1	1	2	1	1	2

Nr.	Code	Reflexpräsenz				Reflexaktivität			
		T 1	T 2	T 3	T 4	T 1	T 2	T 3	T 4
58.	66A	1	1	1	1	2	3	2	2
59.	66B	1	1	1	1	2	2	2	2
60.	68B	1	1	1	1	1	4	1	2
61.	69A	1	1	1	1	2	3	2	2
62.	69B	1	1	1	1	2	3	2	2
63.	70A	1	1	1	1	2	2	2	2
64.	70B	1	1	1	1	2	2	2	2
65.	71A	1	1	1	1	1	1	2	2
66.	71B	1	1	1	1	1	2	1	2
67.	72A	1	1	1	1	2	2	2	2
68.	72B	1	1	1	1	3	2	2	4
69.	73A	1	1	1	1	2	3	2	4
70.	73B	1	1	1	1	2	3	2	4
71.	74A	1	1	1	1	2	3	2	2
72.	74B	1	1	1	1	2	3	1	2
73.	77A	1	1	0	1	2	4	0	2
74.	77B	1	1	1	1	2	3	2	2
75.	81A	1	1	1	1	3	2	2	2
76.	81B	1	1	1	1	3	2	1	2
77.	82B	1	1	1	1	2	3	3	3
78.	83B	1	1	1	1	2	2	2	2
79.	84A	1	1	1	1	2	2	2	2
80.	84B	1	1	1	1	2	2	1	2
81.	85A	1	1	1	1	2	2	2	1
82.	86A	1	1	1	1	2	2	2	1
83.	86B	1	1	1	1	2	2	2	2
84.	87A	1	1	1	1	2	2	3	2
85.	87B	1	1	1	1	3	3	3	2
86.	89A	1	1	1	1	2	2	3	2
87.	89B	1	1	1	1	1	2	1	2

Tabelle 33. Evaluierungen der Reflexpräsenz und –aktivität durch die vier Studenten (S 1-4) in Studie 1. (Reflexpräsenz: 0 = abwesend, 1 = vorhanden; Reflexaktivität: 0 = abwesend, 1 = reduziert, 2 = normal, 3 = gesteigert, 4 = tonisch).

Nr.	Code	Reflexpräsenz				Reflexaktivität			
		S 1	S 2	S 3	S 4	S 1	S 2	S 3	S 4
1.	1B	1	1	1	1	2	2	2	2
2.	2A	1	1	1	1	2	2	2	2
3.	2B	1	1	1	1	2	1	1	2
4.	3A	1	1	1	1	2	2	1	2
5.	3B	1	1	1	1	3	3	3	3
6.	4A	1	1	0	1	2	1	0	2
7.	4B	1	1	1	1	2	2	1	2
8.	10A	1	1	1	1	3	3	3	3
9.	10B	1	1	1	1	2	1	1	2
10.	11A	1	0	1	1	3	0	2	3
11.	11B	1	0	1	1	2	0	1	2
12.	13A	1	1	1	1	3	2	3	2
13.	13B	1	1	1	1	3	1	2	2
14.	18A	1	1	1	1	2	2	1	2
15.	18B	1	1	1	1	2	2	1	2
16.	21A	1	1	1	1	1	1	1	2
17.	22B	1	1	1	1	3	3	2	4
18.	23A	1	1	0	1	2	2	0	2
19.	23B	1	1	0	1	1	2	0	1
20.	24B	1	1	1	1	2	3	2	2
21.	25B	1	1	0	1	3	3	0	2
22.	27A	1	1	1	1	2	2	1	2
23.	27B	1	1	1	1	2	2	1	2
24.	31A	1	1	1	1	2	2	2	2
25.	31B	1	1	1	1	2	2	2	2
26.	32A	1	0	1	1	2	0	3	2
27.	32B	1	1	1	1	2	3	3	2

Nr.	Code	Reflexpräsenz				Reflexaktivität			
		S 1	S 2	S 3	S 4	S 1	S 2	S 3	S 4
28.	33A	1	1	0	1	2	1	0	2
29.	33B	0	0	0	1	0	0	0	2
30.	34B	1	1	1	1	3	3	2	2
31.	36A	1	1	1	1	1	1	1	1
32.	36B	1	1	1	1	2	2	2	2
33.	37A	1	1	1	1	2	2	3	2
34.	37B	1	1	1	1	2	3	3	4
35.	41A	1	1	1	1	2	3	1	2
36.	41B	1	0	0	1	1	0	0	2
37.	42A	1	0	0	0	1	0	0	2
38.	42B	1	1	1	1	2	1	2	2
39.	44B	1	0	1	1	1	0	1	2
40.	45A	1	1	1	1	2	2	2	2
41.	45B	1	1	1	1	2	2	3	2
42.	47B	0	0	0	0	0	0	0	0
43.	48A	1	1	1	1	2	2	2	2
44.	48B	1	1	1	1	3	3	3	3
45.	49B	1	0	1	1	2	0	1	2
46.	51A	1	1	1	1	2	2	2	2
47.	51B	1	1	1	1	2	2	1	2
48.	54A	1	1	1	1	2	1	1	2
49.	58A	1	1	1	1	2	2	2	2
50.	59A	1	1	0	1	2	1	0	2
51.	59B	1	1	1	1	2	2	2	2
52.	60A	1	1	1	1	3	3	3	3
53.	60B	1	1	1	1	3	2	2	2
54.	63A	1	1	1	1	2	2	2	4
55.	63B	1	0	1	1	1	0	1	4
56.	65A	1	1	1	1	1	1	2	2
57.	65B	1	1	1	1	2	1	1	2

Nr.	Code	Reflexpräsenz				Reflexaktivität			
		S 1	S 2	S 3	S 4	S 1	S 2	S 3	S 4
58.	66A	1	1	1	1	2	3	2	2
59.	66B	1	1	1	1	2	2	2	2
60.	68B	1	1	1	1	1	4	1	2
61.	69A	1	1	1	1	2	3	2	2
62.	69B	1	1	1	1	2	3	2	2
63.	70A	1	1	1	1	2	2	2	2
64.	70B	1	1	1	1	2	2	2	2
65.	71A	1	1	1	1	1	1	2	2
66.	71B	1	1	1	1	1	2	1	2
67.	72A	1	1	1	1	2	2	2	2
68.	72B	1	1	1	1	3	2	2	4
69.	73A	1	1	1	1	2	3	2	4
70.	73B	1	1	1	1	2	3	2	4
71.	74A	1	1	1	1	2	3	2	2
72.	74B	1	1	1	1	2	3	1	2
73.	77A	1	1	0	1	2	4	0	2
74.	77B	1	1	1	1	2	3	2	2
75.	81A	1	1	1	1	3	2	2	2
76.	81B	1	1	1	1	3	2	1	2
77.	82B	1	1	1	1	2	3	3	3
78.	83B	1	1	1	1	2	2	2	2
79.	84A	1	1	1	1	2	2	2	2
80.	84B	1	1	1	1	2	2	1	2
81.	85A	1	1	1	1	2	2	2	1
82.	86A	1	1	1	1	2	2	2	1
83.	86B	1	1	1	1	2	2	2	2
84.	87A	1	1	1	1	2	2	3	2
85.	87B	1	1	1	1	3	3	3	2
86.	89A	1	1	1	1	2	2	3	2
87.	89B	1	1	1	1	1	2	1	2

Tabelle 34. Analyse der Reflexpräsenz. Übersicht über die Reliabilitätskoeffizienten *Cohen's Kappa* (K_C) mit dem dazugehörigen 95%igen Konfidenzintervall (KI95%) und Standardfehler (SF) sowie *Prozentuale Übereinstimmung* ($r\%_{Präs}$) aller möglichen Beurteilerkombinationen (Kombi) in Studie 1. Die Sortierung erfolgt absteigend.

	Kombi	K_C	KI95%		SF	Kombi	$r\%_{Präs}$
1.	T1-S1	1,00	1,000	1,000	0,000	T1-S1	100,00
2.	N4-N3	0,85	0,567	1,000	0,145	N1-N3	99,03
3.	N1-N3	0,80	0,404	1,000	0,199	N3-N4	99,03
4.	N4-N2	0,79	0,509	1,000	0,144	N1-N4	98,06
5.	N4-T2	0,79	0,509	1,000	0,144	N1-S1	98,06
6.	N4-T3	0,79	0,509	1,000	0,144	N1-S4	98,06
7.	N4-T1	0,66	0,217	1,000	0,225	N1-T1	98,06
8.	N4-S1	0,66	0,217	1,000	0,225	N2-N4	98,06
9.	N1-N4	0,66	0,217	1,000	0,225	N4-S1	98,06
10.	N2-N3	0,65	0,291	1,000	0,185	N4-T1	98,06
11.	N3-T2	0,65	0,291	1,000	0,653	N4-T2	98,06
12.	N3-T3	0,65	0,291	1,000	0,185	N4-T3	98,06
13.	N2-T3	0,65	0,324	0,968	0,164	S1-S4	98,06
14.	T2-T3	0,65	0,324	0,968	0,164	T1-S4	98,06
15.	N2-T2	0,65	0,324	0,968	0,164	N2-N3	97,09
16.	N2-S2	0,60	0,305	0,887	0,148	N3-S1	97,09
17.	N4-S2	0,55	0,232	0,860	0,160	N3-S4	97,09
18.	N1-T1	0,49	-0,118	1,000	0,310	N3-T1	97,09
19.	N1-S1	0,49	-0,118	1,000	0,310	N3-T2	97,09
20.	S1-S4	0,49	-0,118	1,000	0,310	N3-T3	97,09
21.	T1-S4	0,49	-0,118	1,000	0,310	N1-N2	96,12
22.	N1-S4	0,49	-0,118	1,000	0,310	N1-T2	96,12
23.	N1-N2	0,49	0,061	0,909	0,216	N1-T3	96,12
24.	T1-T2	0,49	0,061	0,909	0,216	N1-T4	96,12
25.	T1-T3	0,49	0,061	0,909	0,216	N2-S1	96,12
26.	T2-S1	0,49	0,061	0,909	0,216	N2-S4	96,12
27.	N2-T1	0,49	0,061	0,909	0,216	N2-T1	96,12
28.	N2-S1	0,49	0,061	0,909	0,216	N2-T2	96,12

	Kombi	K_C	KI95%		SF	Kombi	r%Präs
29.	N1-T2	0,49	0,061	0,909	0,216	N2-T3	96,12
30.	N1-T3	0,49	0,061	0,909	0,216	N4-S4	96,12
31.	T3-S1	0,49	0,061	0,909	0,216	T1-T2	96,12
32.	N1-T4	0,49	0,061	0,909	0,216	T1-T3	96,12
33.	N2-S4	0,49	0,061	0,909	0,216	T2-S1	96,12
34.	T2-S2	0,46	0,148	0,773	0,160	T2-T3	96,12
35.	T4-S2	0,46	0,148	0,773	0,160	T3-S1	96,12
36.	T3-S2	0,46	0,148	0,773	0,160	N3-T4	95,15
37.	N3-S2	0,44	0,103	0,769	0,170	N2-S2	94,17
38.	N2-S3	0,43	0,123	0,732	0,155	N4-S2	94,17
39.	T4-S3	0,43	0,123	0,732	0,155	N4-T4	94,17
40.	N3-T4	0,42	0,011	0,833	0,210	T1-T4	94,17
41.	N3-T1	0,39	-0,165	0,936	0,281	T2-S4	94,17
42.	N3-S1	0,39	-0,165	0,936	0,281	T3-S4	94,17
43.	N3-S4	0,39	-0,165	0,936	0,281	T4-S1	94,17
44.	N4-T4	0,37	-0,024	0,765	0,201	T4-S4	94,17
45.	N4-S4	0,32	-0,176	0,807	0,251	N3-S2	93,20
46.	S1-S2	0,31	-0,021	0,643	0,311	N1-S2	92,23
47.	T1-S2	0,31	-0,021	0,643	0,170	N2-T4	92,23
48.	S2-S3	0,31	0,028	0,594	0,145	S1-S2	92,23
49.	N1-S2	0,31	-0,021	0,643	0,170	S2-S4	92,23
50.	S2-S4	0,31	-0,021	0,643	0,170	T1-S2	92,23
51.	T3-S3	0,30	-0,002	0,602	0,154	T2-S2	92,23
52.	T2-T4	0,29	-0,065	0,650	0,182	T2-T4	92,23
53.	N2-T4	0,29	-0,065	0,650	0,182	T3-S2	92,23
54.	T3-T4	0,29	-0,065	0,650	0,182	T3-T4	92,23
55.	S1-S3	0,28	-0,162	0,731	0,228	T4-S2	92,23
56.	T1-S3	0,28	-0,028	0,596	0,159	N2-S3	91,26
57.	S3-S4	0,28	-0,163	0,731	0,228	S1-S3	91,26
58.	T1-T4	0,23	-0,169	0,624	0,202	S3-S4	91,26
59.	T4-S1	0,23	-0,169	0,624	0,202	T1-S3	91,26

	Kombi	K_C	KI95%		SF	Kombi	r%Präs
60.	T4-S4	0,23	-0,169	0,624	0,202	T4-S3	91,26
61.	T3-S4	0,23	-0,169	0,624	0,202	N1-S3	89,32
62.	T2-S4	0,23	-0,169	0,624	0,202	N4-S3	89,32
63.	N4-S3	0,22	-0,073	0,518	0,151	T3-S3	89,32
64.	T2-S3	0,17	-0,104	0,450	0,141	N3-S3	88,35
65.	N1-S3	0,13	-0,130	0,381	0,130	S2-S3	87,38
66.	N3-S3	0,10	-0,144	0,348	0,126	T2-S3	87,38

Tabelle 35. Analyse der Reflexaktivität. Übersicht über die Reliabilitätskoeffizienten *gewichtetes Kappa* (K_w) mit dem dazugehörigen 95%igen Konfidenzintervall (KI95%) und Standardfehler (SF) sowie *Prozentuale Übereinstimmung* ($r\%_{Akt}$) aller möglichen Beurteilerkombinationen (Kombi) in Studie 1. Die Sortierung erfolgt absteigend.

	Kombi	K_w	KI95%		SF	Kombi	$r\%_{Akt}$
1.	T1-S1	0,61	0,463	0,757	0,075	N2-N3	85,44
2.	N2-N3	0,60	0,343	0,849	0,129	N1-N3	82,52
3.	N4-T1	0,59	0,419	0,759	0,087	N2-N4	82,52
4.	N4-N2	0,58	0,346	0,815	0,119	N1-N2	79,61
5.	N4-N3	0,57	0,360	0,787	0,109	N3-N4	79,61
6.	N4-T2	0,57	0,419	0,722	0,077	N1-N4	75,73
7.	N2-T3	0,56	0,410	0,715	0,078	N1-S4	75,73
8.	N1-N3	0,56	0,312	0,811	0,127	N4-S4	71,84
9.	N1-N2	0,55	0,333	0,775	0,113	N4-T1	71,84
10.	T1-T2	0,55	0,395	0,702	0,078	N3-S4	70,87
11.	T1-T3	0,52	0,329	0,711	0,097	N4-S1	69,90
12.	N4-S1	0,50	0,297	0,706	0,104	T1-T3	69,90
13.	T2-S1	0,50	0,344	0,654	0,079	T1-S1	68,93
14.	T2-T3	0,49	0,258	0,724	0,119	N2-T1	67,96
15.	T1-T4	0,49	0,334	0,645	0,079	N2-S4	66,99
16.	N2-T2	0,47	0,242	0,694	0,115	N2-T3	66,99
17.	N3-T2	0,47	0,273	0,663	0,100	N4-T4	66,99
18.	N3-T1	0,45	0,285	0,622	0,086	N2-S1	66,02
19.	N2-T1	0,45	0,246	0,658	0,105	N2-T2	66,02
20.	N1-N4	0,45	0,218	0,671	0,116	T2-T3	66,02
21.	T2-T4	0,44	0,275	0,606	0,085	T4-S1	66,02
22.	T4-S1	0,44	0,226	0,649	0,108	N1-T1	65,05
23.	T3-S3	0,43	0,273	0,595	0,082	N2-T4	65,05
24.	N2-S1	0,43	0,224	0,632	0,104	N3-S1	65,05
25.	N3-S1	0,42	0,240	0,606	0,093	N3-T1	65,05
26.	N2-S3	0,42	0,262	0,580	0,081	N1-S1	64,08
27.	S1-S3	0,42	0,258	0,583	0,083	N3-T2	64,08
28.	T2-S2	0,42	0,242	0,593	0,090	S1-S4	64,08

	Kombi	K_w	KI95%		SF	Kombi	r%_{Akt}
29.	N1-T2	0,41	0,233	0,592	0,091	T1-T4	63,11
30.	N1-T1	0,41	0,230	0,591	0,092	N1-T2	62,14
31.	S1-S2	0,41	0,212	0,600	0,099	N1-T3	62,14
32.	N3-T3	0,40	0,173	0,619	0,114	N1-T4	62,14
33.	T1-S2	0,39	0,231	0,552	0,082	N3-T3	62,14
34.	N2-T4	0,39	0,185	0,591	0,104	N4-T2	62,14
35.	N4-T4	0,38	0,189	0,578	0,099	T1-T2	62,14
36.	N3-S2	0,38	0,196	0,565	0,094	N3-T4	61,17
37.	S2-S3	0,38	0,181	0,580	0,102	T4-S4	60,19
38.	T3-T4	0,38	0,210	0,543	0,085	T2-S1	58,25
39.	N1-S2	0,38	0,227	0,525	0,076	T1-S3	57,28
40.	N1-T3	0,37	0,192	0,555	0,093	N4-T3	56,31
41.	N2-S2	0,37	0,160	0,579	0,107	T1-S4	55,34
42.	N4-T3	0,36	0,152	0,574	0,108	T3-S3	55,34
43.	T3-S1	0,36	0,188	0,526	0,086	N2-S3	54,37
44.	T4-S2	0,36	0,122	0,591	0,120	N2-S2	53,40
45.	T3-S2	0,35	0,198	0,510	0,080	N4-S3	53,40
46.	T1-S3	0,35	0,152	0,553	0,102	S1-S2	53,40
47.	N1-T4	0,33	0,101	0,557	0,116	T2-T4	53,40
48.	N4-S2	0,33	0,146	0,510	0,093	T3-S1	52,43
49.	N1-S1	0,33	0,098	0,551	0,116	T3-T4	52,43
50.	N3-T4	0,32	0,090	0,540	0,115	N1-S3	51,46
51.	T2-S3	0,30	0,087	0,502	0,106	S1-S3	51,46
52.	S1-S4	0,29	0,065	0,510	0,113	T2-S4	51,46
53.	N1-S3	0,28	0,118	0,439	0,082	T4-S2	51,46
54.	T4-S3	0,27	0,067	0,463	0,101	N3-S3	50,49
55.	S3-S4	0,23	0,101	0,361	0,066	N3-S2	49,51
56.	T1-S4	0,22	0,055	0,387	0,085	N4-S2	49,51
57.	N4-S3	0,22	0,019	0,421	0,102	S2-S3	47,57
58.	T4-S4	0,21	0,033	0,395	0,093	S2-S4	47,57
59.	N4-S4	0,21	0,015	0,406	0,100	T3-S4	47,57

	Kombi	K_w	KI95%		SF	Kombi	r%_{Akt}
60.	N3-S3	0,20	0,020	0,379	0,092	N1-S2	46,60
61.	N3-S4	0,19	-0,034	0,410	0,113	T4-S3	46,60
62.	N1-S4	0,18	-,0049	0,404	0,116	T2-S3	45,63
63.	S2-S4	0,17	-0,029	0,365	0,100	S3-S4	43,69
64.	N2-S4	0,15	-0,025	0,321	0,088	T1-S2	43,69
65.	T3-S4	0,15	0,017	0,277	0,066	T2-S2	43,69
66.	T2-S4	0,08	-0,093	0,253	0,088	T3-S2	40,78

9.3. Evaluierungen der Beurteiler in Studie 2a

Tabelle 36. Evaluierungen der Reflexpräsenz und –aktivität durch die drei Neurologen (N 1-3) in Studie 2a. (Reflexpräsenz: 0 = abwesend, 1 = vorhanden; Reflexaktivität: 0 = abwesend, 1 = reduziert, 2 = normal, 3 = gesteigert, 4 = tonisch).

Nr.	Code	Reflexpräsenz			Reflexaktivität		
		N 1	N 2	N 3	N 1	N 2	N 3
1.	1B	1	1	1	2	2	2
2.	2B	1	1	1	1	2	1
3.	3B	1	1	1	2	2	2
4.	4B	1	1	1	1	2	2
5.	10B	1	1	1	2	1	1
6.	11B	1	1	1	1	1	1
7.	13B	1	1	1	2	2	2
8.	18B	1	1	1	2	2	2
9.	22B	1	1	1	2	2	2
10.	23B	1	1	1	1	2	2
11.	24B	1	1	1	2	2	2
12.	25B	1	1	1	2	1	3
13.	27B	1	1	1	2	2	2
14.	31B	1	1	1	2	2	2
15.	32B	1	1	1	2	2	2
16.	33B	1	0	1	1	0	1
17.	34B	1	1	1	2	2	2
18.	36B	1	1	1	2	2	2
19.	37B	1	1	1	2	2	2
20.	41B	1	1	1	2	1	1
21.	42B	1	1	1	2	2	2
22.	44B	0	0	0	0	0	0
23.	45B	1	1	1	2	2	2
24.	47B	0	0	0	0	0	0
25.	48B	1	1	1	2	2	2
26.	49B	1	0	0	1	0	0

Nr.	Code	Reflexpräsenz			Reflexaktivität		
		N 1	N 2	N 3	N 1	N 2	N 3
27.	51B	1	1	1	2	2	2
28.	59B	1	1	1	2	2	2
29.	60B	1	1	1	2	2	2
30.	63B	1	1	1	1	1	1
31.	65B	1	1	1	2	2	2
32.	66B	1	1	1	2	2	2
33.	68B	1	1	1	2	1	1
34.	69B	1	1	1	2	2	2
35.	70B	1	1	1	2	2	2
36.	71B	1	1	1	2	2	2
37.	72B	1	1	1	2	2	3
38.	73B	1	1	1	2	2	2
39.	74B	1	1	1	2	2	2
40.	77B	1	1	1	2	2	2
41.	81B	1	1	1	2	2	2
42.	82B	1	1	1	2	2	2
43.	83B	1	1	1	2	1	2
44.	84B	1	1	1	2	2	2
45.	86B	1	1	1	2	2	2
46.	87B	1	1	1	2	2	2
47.	89B	1	1	1	2	1	1
48.	90B	1	1	1	2	2	2
49.	93B	1	1	1	1	2	2
50.	94B	1	1	1	2	2	2
51.	95B	1	1	1	2	1	2
52.	96B	1	1	1	1	2	2
53.	97B	1	1	1	2	2	2
54.	98B	1	1	1	2	2	2
55.	99B	1	1	1	2	2	2
56.	100B	1	1	1	2	2	2

Tabelle 37. Evaluierungen der Reflexpräsenz und -aktivität durch die drei praktizierenden Tierärzte (T 1-3) in Studie 2a. (Reflexpräsenz: 0 = abwesend, 1 = vorhanden; Reflexaktivität: 0 = abwesend, 1 = reduziert, 2 = normal, 3 = gesteigert, 4 = tonisch).

Nr.	Code	Reflexpräsenz			Reflexaktivität		
		T 1	T 2	T 3	T 1	T 2	T 3
1.	1B	1	1	1	1	2	2
2.	2B	1	1	1	2	1	1
3.	3B	1	1	1	3	2	2
4.	4B	1	1	1	1	1	1
5.	10B	1	1	1	2	3	1
6.	11B	1	1	1	1	1	1
7.	13B	1	1	1	3	3	2
8.	18B	1	1	1	3	2	1
9.	22B	1	1	1	2	2	2
10.	23B	1	1	1	1	2	1
11.	24B	1	1	1	2	3	2
12.	25B	1	1	0	3	4	0
13.	27B	1	1	1	1	2	1
14.	31B	1	1	1	2	2	2
15.	32B	1	1	1	1	2	1
16.	33B	0	0	0	0	0	0
17.	34B	1	1	1	2	2	2
18.	36B	1	1	1	2	2	2
19.	37B	1	1	1	2	1	1
20.	41B	1	1	1	1	1	1
21.	42B	1	1	1	2	2	2
22.	44B	1	0	0	1	0	0
23.	45B	1	1	1	2	2	2
24.	47B	0	0	0	0	0	0
25.	48B	1	1	1	2	2	2
26.	49B	1	0	0	1	0	0
27.	51B	1	1	1	1	1	1

Nr.	Code	Reflexpräsenz			Reflexaktivität		
		T 1	T 2	T 3	T 1	T 2	T 3
28.	59B	1	1	1	2	2	2
29.	60B	1	1	1	2	3	2
30.	63B	1	1	1	1	1	1
31.	65B	1	1	1	2	2	2
32.	66B	1	1	1	2	3	1
33.	68B	1	1	1	1	1	1
34.	69B	1	1	1	2	2	1
35.	70B	1	1	1	2	3	2
36.	71B	1	1	1	1	1	1
37.	72B	1	1	1	2	2	2
38.	73B	1	1	1	2	2	2
39.	74B	1	1	1	2	2	2
40.	77B	1	1	1	2	2	2
41.	81B	1	1	1	2	2	1
42.	82B	1	1	1	2	3	2
43.	83B	1	1	1	3	1	1
44.	84B	1	1	1	2	3	2
45.	86B	1	1	1	2	2	2
46.	87B	1	1	1	2	3	2
47.	89B	1	1	0	1	1	0
48.	90B	1	1	1	2	2	2
49.	93B	1	1	1	2	2	1
50.	94B	1	1	1	2	2	3
51.	95B	1	1	1	1	3	1
52.	96B	1	1	1	2	1	2
53.	97B	1	1	1	2	1	1
54.	98B	1	1	1	2	3	2
55.	99B	1	1	1	2	2	2
56.	100B	1	1	1	1	2	1

Tabelle 38. Evaluierungen der Reflexpräsenz und –aktivität durch die drei Studenten (S 1-3) in Studie 2a. (Reflexpräsenz: 0 = abwesend, 1 = vorhanden; Reflexaktivität: 0 = abwesend, 1 = reduziert, 2 = normal, 3 = gesteigert, 4 = tonisch).

Nr.	Code	Reflexpräsenz			Reflexaktivität		
		S 1	S 2	S 3	S 1	S 2	S 3
1.	1B	1	1	1	2	2	2
2.	2B	1	1	1	2	1	1
3.	3B	1	1	1	3	3	3
4.	4B	1	1	1	2	2	1
5.	10B	1	1	1	2	1	1
6.	11B	1	0	1	2	0	1
7.	13B	1	1	1	3	1	2
8.	18B	1	1	1	2	2	1
9.	22B	1	1	1	3	3	2
10.	23B	1	1	0	1	2	0
11.	24B	1	1	1	2	3	2
12.	25B	1	1	0	3	3	0
13.	27B	1	1	1	2	2	1
14.	31B	1	1	1	2	2	2
15.	32B	1	1	1	2	3	3
16.	33B	0	0	0	0	0	0
17.	34B	1	1	1	3	3	2
18.	36B	1	1	1	2	2	2
19.	37B	1	1	1	2	3	3
20.	41B	1	0	0	1	0	0
21.	42B	1	1	1	2	1	2
22.	44B	1	0	1	1	0	1
23.	45B	1	1	1	2	2	3
24.	47B	0	0	0	0	0	0
25.	48B	1	1	1	3	3	3
26.	49B	1	0	1	2	0	1
27.	51B	1	1	1	2	2	1

Nr.	Code	Reflexpräsenz			Reflexaktivität		
		S 1	S 2	S 3	S 1	S 2	S 3
28.	59B	1	1	1	2	2	2
29.	60B	1	1	1	3	2	2
30.	63B	1	0	1	1	0	1
31.	65B	1	1	1	2	1	1
32.	66B	1	1	1	2	2	2
33.	68B	1	1	1	1	4	1
34.	69B	1	1	1	2	3	2
35.	70B	1	1	1	2	2	2
36.	71B	1	1	1	1	2	1
37.	72B	1	1	1	3	2	2
38.	73B	1	1	1	2	3	2
39.	74B	1	1	1	2	3	1
40.	77B	1	1	1	2	3	2
41.	81B	1	1	1	3	2	1
42.	82B	1	1	1	2	3	3
43.	83B	1	1	1	2	2	2
44.	84B	1	1	1	2	2	1
45.	86B	1	1	1	2	2	2
46.	87B	1	1	1	3	3	3
47.	89B	1	1	1	1	2	1
48.	90B	1	1	1	2	4	2
49.	93B	1	1	1	3	2	3
50.	94B	1	1	1	2	3	3
51.	95B	1	1	1	3	3	2
52.	96B	1	1	1	2	2	2
53.	97B	1	1	1	2	2	2
54.	98B	1	1	1	3	3	2
55.	99B	1	1	1	2	3	3
56.	100B	1	1	1	1	1	1

Tabelle 39. Analyse der Reflexpräsenz. Übersicht über die Reliabilitätskoeffizienten *Cohen's Kappa* (K_C) mit dem dazugehörigen 95%igen Konfidenzintervall (KI95%) und Standardfehler (SF) sowie *Prozentuale Übereinstimmung* ($r\%_{Präs}$) aller möglichen Beurteilerkombinationen (Kombi) in Studie 2a. Die Sortierung erfolgt absteigend.

	Kombi	K_C	KI95%		SF	Kombi	$r\%_{Präs}$
1.	N2-T2	1,00	1,000	1,000	0,000	N2-T2	100,00
2.	T1-S1	1,00	1,000	1,000	0,000	T1-S1	100,00
3.	N2-N3	0,85	0,556	1,000	0,149	N1-N3	98,21
4.	N3-T2	0,85	0,556	1,000	0,149	N2-N3	98,21
5.	N1-N3	0,79	0,394	1,000	0,203	N3-T2	98,21
6.	N2-T3	0,78	0,491	1,000	0,148	N1-N2	96,43
7.	T2-T3	0,78	0,491	1,000	0,148	N1-S1	96,43
8.	N2-S2	0,70	0,385	1,000	0,161	N1-T1	96,43
9.	T2-S2	0,70	0,385	1,000	0,161	N1-T2	96,43
10.	N1-N2	0,65	0,204	1,000	0,228	N2-S1	96,43
11.	N1-T2	0,65	0,204	1,000	0,228	N2-T1	96,43
12.	N2-S1	0,65	0,204	1,000	0,228	N2-T3	96,43
13.	N2-T1	0,65	0,204	1,000	0,228	T1-T2	96,43
14.	T1-T2	0,65	0,204	1,000	0,228	T2-S1	96,43
15.	T2-S1	0,65	0,204	1,000	0,228	T2-T3	96,43
16.	N3-T3	0,64	0,272	1,000	0,188	N2-S2	94,64
17.	N3-S2	0,57	0,199	0,936	0,188	N3-S1	94,64
18.	T3-S2	0,57	0,225	0,905	0,174	N3-T1	94,64
19.	S1-S3	0,55	0,105	0,992	0,226	N3-T3	94,64
20.	T1-S3	0,55	0,105	0,992	0,226	S1-S3	94,64
21.	T3-S3	0,50	0,116	0,877	0,194	T1-S3	94,64
22.	N1-S1	0,48	-0,134	1,000	0,314	T2-S2	94,64
23.	N1-T1	0,48	-0,134	1,000	0,314	N1-T3	92,86
24.	N1-T3	0,47	0,048	0,895	0,216	N3-S2	92,86
25.	T1-T3	0,47	0,048	0,895	0,216	T1-T3	92,86
26.	T3-S1	0,47	0,048	0,895	0,216	T3-S1	92,86
27.	S2-S3	0,44	0,071	0,813	0,189	N1-S2	91,07
28.	N1-S2	0,41	0,014	0,810	0,203	N1-S3	91,07

	Kombi	K_C	KI95%		SF	Kombi	r%Präs
29.	S1-S2	0,41	0,014	0,810	0,203	N2-S3	91,07
30.	T1-S2	0,41	0,014	0,810	0,203	S1-S2	91,07
31.	N2-S3	0,40	-0,030	0,823	0,218	T1-S2	91,07
32.	T2-S3	0,40	-0,030	0,823	0,218	T2-S3	91,07
33.	N3-S1	0,37	-0,183	0,930	0,284	T3-S2	91,07
34.	N3-T1	0,37	-0,183	0,930	0,284	T3-S3	91,07
35.	N1-S3	0,25	-0,194	0,689	0,225	N3-S3	89,29
36.	N3-S3	0,20	-0,214	0,607	0,209	S2-S3	89,29

Tabelle 40. Analyse der Reflexaktivität. Übersicht über die Reliabilitätskoeffizienten *gewichtetes Kappa* (K_w) mit dem dazugehörigen 95%igen Konfidenzintervall (KI95%) und Standardfehler (SF) sowie *Prozentuale Übereinstimmung* ($r\%_{Akt}$) aller möglichen Beurteilerkombinationen (Kombi) in Studie 2a. Die Sortierung erfolgt absteigend.

	Kombi	K_w	KI95%		SF	Kombi	$r\%_{Akt}$
1.	N2-N3	0,77	0,549	0,998	0,115	N2-N3	89,29
2.	N1-N3	0,67	0,446	0,896	0,115	N1-N3	80,36
3.	N2-T3	0,65	0,485	0,822	0,086	N1-N2	75,00
4.	T1-S1	0,62	0,431	0,817	0,099	N2-T1	71,43
5.	N3-T2	0,62	0,420	0,813	0,100	N2-T3	67,86
6.	T2-S1	0,60	0,436	0,772	0,086	T1-T3	67,86
7.	N1-N2	0,59	0,334	0,845	0,130	N2-T2	66,07
8.	T3-S3	0,58	0,403	0,757	0,090	N3-T1	66,07
9.	N3-T1	0,57	0,392	0,755	0,093	T1-S1	66,07
10.	N3-S1	0,55	0,361	0,745	0,098	N1-T1	64,29
11.	T1-T2	0,55	0,348	0,750	0,103	N3-S1	64,29
12.	N1-T2	0,53	0,338	0,719	0,097	N2-S1	62,50
13.	N3-S2	0,52	0,295	0,740	0,114	N3-T2	62,50
14.	N2-T1	0,51	0,231	0,789	0,143	N3-T3	62,50
15.	S2-S3	0,51	0,293	0,719	0,109	N1-T3	60,71
16.	N2-S3	0,50	0,321	0,680	0,091	T3-S3	60,71
17.	N1-T1	0,49	0,268	0,707	0,112	T1-S3	58,93
18.	N2-T2	0,49	0,182	0,789	0,155	T2-T3	58,93
19.	T2-S2	0,48	0,255	0,712	0,117	N1-T2	57,14
20.	N3-T3	0,47	0,183	0,760	0,147	N1-S1	55,36
21.	N2-S2	0,46	0,216	0,694	0,122	T1-T2	55,36
22.	N2-S1	0,45	0,193	0,708	0,132	T2-S1	53,57
23.	T1-S3	0,45	0,176	0,725	0,140	N3-S3	51,79
24.	T1-T3	0,45	0,172	0,731	0,143	N1-S3	50,00
25.	S1-S2	0,45	0,205	0,693	0,124	N2-S3	50,00
26.	N1-S2	0,45	0,268	0,629	0,092	S1-S3	50,00
27.	N1-T3	0,43	0,192	0,657	0,119	N3-S2	48,21
28.	T2-T3	0,42	0,116	0,731	0,157	S1-S2	48,21

	Kombi	K_w	KI95%		SF	Kombi	r%_{Akt}
29.	N1-S1	0,40	0,148	0,647	0,127	S2-S3	48,21
30.	N3-S3	0,39	0,137	0,643	0,129	N2-S2	46,43
31.	T1-S2	0,37	0,141	0,594	0,115	T3-S1	46,43
32.	T3-S1	0,35	0,133	0,571	0,112	T2-S3	44,64
33.	T3-S2	0,35	0,148	0,553	0,103	N1-S2	39,29
34.	N1-S3	0,35	0,124	0,569	0,113	T2-S2	39,29
35.	T2-S3	0,33	0,029	0,636	0,155	T1-S2	33,93
36.	S1-S3	0,27	0,265	0,730	0,119	T3-S2	33,93

9.4. Evaluierungen der Beurteiler in Studie 2b

Tabelle 41. Evaluierungen der Reflexpräsenz und -aktivität durch die drei Neurologen (N 1-3) in Studie 2b. (Reflexpräsenz: 0 = abwesend, 1 = vorhanden; Reflexaktivität: 0 = abwesend, 1 = reduziert, 2 = normal, 3 = gesteigert, 4 = tonisch).

Nr.	Code	Reflexpräsenz			Reflexaktivität		
		N 1	N 2	N 3	N 1	N 2	N 3
1.	1	1	1	1	1	2	2
2.	2	1	1	1	2	2	2
3.	3	1	1	1	2	2	3
4.	4	1	1	1	2	2	2
5.	5	1	1	1	2	2	1
6.	7	1	1	1	2	2	2
7.	9	1	1	1	2	2	2
8.	11	1	1	1	1	1	1
9.	12	1	1	1	2	2	2
10.	14	1	1	1	2	2	2
11.	15	1	1	1	2	2	2
12.	17	1	1	1	1	1	1
13.	18	1	1	1	2	2	1
14.	19	1	1	1	2	2	2
15.	23	1	1	1	2	2	1
16.	24	1	1	1	2	3	3
17.	25	1	1	1	1	1	1
18.	27	1	1	1	3	3	2
19.	28	0	0	1	0	0	1
20.	29	1	1	1	2	2	2
21.	31	1	1	1	1	1	1
22.	32	1	1	1	3	3	2
23.	33	1	1	1	1	1	1
24.	34	1	1	1	2	2	2
25.	35	1	1	1	1	2	1
26.	36	1	1	1	2	2	2

Nr.	Code	Reflexpräsenz			Reflexaktivität		
		N 1	N 2	N 3	N 1	N 2	N 3
27.	37	1	1	1	2	2	2
28.	38	1	1	1	2	2	2
29.	39	1	1	1	1	1	2
30.	41	1	1	1	2	2	2
31.	42	1	1	1	2	2	2
32.	47	1	1	1	2	3	3
33.	48	1	1	1	2	2	2
34.	50	1	1	1	2	2	2
35.	51	1	1	1	2	2	2
36.	52	1	1	1	2	2	2
37.	54	1	1	1	2	2	2
38.	55	1	1	1	1	1	1
39.	57	1	1	1	2	2	2
40.	58	1	1	1	2	2	2
41.	62	1	1	1	2	2	3
42.	63	1	1	1	2	2	2
43.	66	1	1	1	2	2	1
44.	67	1	1	1	2	2	2
45.	68	1	1	1	1	1	1
46.	70	1	1	1	2	2	2
47.	72	1	1	1	3	2	2
48.	73	1	1	1	2	2	2
49.	74	1	1	1	2	2	2
50.	75	1	1	1	2	2	1
51.	76	1	1	1	1	2	1
52.	77	1	1	1	1	1	1
53.	78	1	1	1	1	1	1
54.	79	1	1	1	2	2	1
55.	84	1	1	1	2	2	1
56.	85	1	1	1	1	2	1

Nr.	Code	Reflexpräsenz			Reflexaktivität		
		N 1	N 2	N 3	N 1	N 2	N 3
57.	87	1	1	1	1	2	1
58.	89	1	1	1	1	1	1
59.	92	1	1	1	2	2	2
60.	93	1	1	1	1	1	2
61.	94	1	1	1	2	2	2
62.	96	1	0	0	1	0	0
63.	98	1	1	1	2	3	2
64.	100	1	1	1	2	2	1

Tabelle 42. Evaluierungen der Reflexpräsenz und -aktivität durch die drei praktizierenden Tierärzte (T 1-3) in Studie 2b. (Reflexpräsenz: 0 = abwesend, 1 = vorhanden; Reflexaktivität: 0 = abwesend, 1 = reduziert, 2 = normal, 3 = gesteigert, 4 = tonisch).

Nr.	Code	Reflexpräsenz			Reflexaktivität		
		T 1	T 2	T 3	T 1	T 2	T 3
1.	1	1	1	1	2	2	2
2.	2	1	1	1	3	2	3
3.	3	1	1	1	3	3	3
4.	4	1	1	1	2	2	2
5.	5	1	1	1	2	2	2
6.	7	1	1	1	3	3	2
7.	9	1	1	1	3	2	2
8.	11	1	1	1	1	1	1
9.	12	1	1	1	2	2	2
10.	14	1	1	1	2	2	2
11.	15	1	1	1	3	2	2
12.	17	1	1	1	1	1	1
13.	18	1	1	1	2	2	2
14.	19	1	1	1	2	2	2
15.	23	1	1	1	2	2	2
16.	24	1	1	1	3	2	2
17.	25	1	1	1	1	1	1
18.	27	1	1	1	3	3	3
19.	28	1	1	1	1	1	1
20.	29	1	1	1	2	2	2
21.	31	1	1	1	1	2	2
22.	32	1	1	1	3	3	3
23.	33	1	1	1	1	1	1
24.	34	1	1	1	3	2	2
25.	35	1	1	1	2	2	2
26.	36	1	1	1	2	2	2
27.	37	1	1	1	2	2	2

Nr.	Code	Reflexpräsenz			Reflexaktivität		
		T 1	T 2	T 3	T 1	T 2	T 3
28.	38	1	1	1	2	3	3
29.	39	1	1	1	1	1	2
30.	41	1	1	1	2	2	2
31.	42	1	1	1	2	2	2
32.	47	1	1	1	3	3	3
33.	48	1	1	1	2	2	2
34.	50	1	1	1	1	2	2
35.	51	1	1	1	2	2	2
36.	52	1	1	1	3	2	2
37.	54	1	1	1	3	3	3
38.	55	1	1	1	1	1	1
39.	57	1	1	1	2	2	2
40.	58	1	1	1	2	3	2
41.	62	1	1	1	2	2	2
42.	63	1	1	1	2	2	2
43.	66	1	1	1	1	2	2
44.	67	1	1	1	2	2	2
45.	68	1	1	1	1	2	2
46.	70	1	1	1	3	2	2
47.	72	1	1	1	3	2	3
48.	73	1	1	1	3	3	2
49.	74	1	1	1	2	2	2
50.	75	1	1	1	2	1	2
51.	76	1	1	1	1	1	1
52.	77	1	1	1	1	1	2
53.	78	1	1	1	1	1	1
54.	79	1	1	1	1	1	1
55.	84	1	1	1	2	2	2
56.	85	1	1	1	1	1	2
57.	87	1	1	1	1	1	1

Nr.	Code	Reflexpräsenz			Reflexaktivität		
		T 1	T 2	T 3	T 1	T 2	T 3
58.	89	1	1	1	1	1	1
59.	92	1	1	1	3	2	2
60.	93	0	1	1	0	1	1
61.	94	1	1	1	3	2	2
62.	96	0	0	1	0	0	1
63.	98	1	1	1	3	2	2
64.	100	1	1	1	1	1	1

Tabelle 43. Evaluierungen der Reflexpräsenz und –aktivität durch die drei Studenten (S 1-3) in Studie 2b. (Reflexpräsenz: 0 = abwesend, 1 = vorhanden; Reflexaktivität: 0 = abwesend, 1 = reduziert, 2 = normal, 3 = gesteigert, 4 = tonisch).

Nr.	Code	Reflexpräsenz			Reflexaktivität		
		S 1	S 2	S 3	S 1	S 2	S 3
1.	1	1	1	1	1	1	2
2.	2	1	1	1	2	2	3
3.	3	1	1	1	3	3	3
4.	4	1	1	1	2	2	2
5.	5	1	1	1	1	1	1
6.	7	1	1	1	2	2	3
7.	9	1	1	1	2	2	2
8.	11	1	1	1	1	1	1
9.	12	1	1	1	2	1	2
10.	14	1	1	1	2	2	3
11.	15	1	1	1	3	3	3
12.	17	1	1	1	1	1	1
13.	18	1	1	1	2	2	2
14.	19	1	1	1	3	2	2
15.	23	1	1	1	2	3	4
16.	24	1	1	1	3	4	3
17.	25	1	1	1	1	1	1
18.	27	1	1	1	3	4	3
19.	28	1	0	0	1	0	0
20.	29	1	1	1	2	2	2
21.	31	1	1	1	2	1	1
22.	32	1	1	1	3	4	2
23.	33	1	0	1	1	0	1
24.	34	1	1	1	3	3	3
25.	35	1	1	1	2	1	2
26.	36	1	1	1	2	2	1
27.	37	1	1	1	2	2	1

Nr.	Code	Reflexpräsenz			Reflexaktivität		
		S 1	S 2	S 3	S 1	S 2	S 3
28.	38	1	1	1	3	3	3
29.	39	1	0	1	1	0	3
30.	41	1	1	1	2	2	2
31.	42	1	1	1	2	3	2
32.	47	1	1	1	3	3	3
33.	48	1	1	1	2	3	2
34.	50	1	1	1	2	1	1
35.	51	1	1	1	2	1	2
36.	52	1	1	1	3	3	3
37.	54	1	1	1	3	3	3
38.	55	1	0	1	1	0	1
39.	57	1	1	1	2	2	2
40.	58	1	1	1	2	2	2
41.	62	1	1	1	2	3	2
42.	63	1	1	1	3	3	2
43.	66	1	1	1	2	1	1
44.	67	1	1	1	2	2	2
45.	68	1	1	1	1	1	1
46.	70	1	1	1	3	2	2
47.	72	1	1	1	3	3	3
48.	73	1	1	1	2	4	3
49.	74	1	1	1	2	3	2
50.	75	1	1	1	2	2	2
51.	76	1	1	0	1	1	0
52.	77	1	1	0	1	1	0
53.	78	1	1	0	1	1	0
54.	79	1	1	1	1	1	1
55.	84	1	0	1	4	0	1
56.	85	1	1	1	1	1	1
57.	87	1	1	1	2	1	1

Nr.	Code	Reflexpräsenz			Reflexaktivität		
		S 1	S 2	S 3	S 1	S 2	S 3
58.	89	1	0	1	2	0	0
59.	92	1	1	1	2	4	2
60.	93	1	1	1	1	1	1
61.	94	1	1	1	3	4	3
62.	96	0	0	0	0	0	0
63.	98	1	1	1	3	3	3
64.	100	1	0	1	1	0	2

Tabelle 44. Analyse der Reflexpräsenz. Übersicht über die Reliabilitätskoeffizienten *Cohen's Kappa* (K_C) mit dem dazugehörigen 95%igen Konfidenzintervall (KI95%) und Standardfehler (SF) sowie *Prozentuale Übereinstimmung* ($r\%_{Präs}$) aller möglichen Beurteilerkombinationen (Kombi) in Studie 2b. Die Sortierung erfolgt absteigend.

	Kombi	Kappa	KI95%		SF	Kombi	r%Präs
1.	N3-S1	1,00	1,000	1,000	0,000	N3-S1	100,00
2.	N3-T2	1,00	1,000	1,000	0,000	N3-T2	100,00
3.	T2-S1	1,00	1,000	1,000	0,000	T2-S1	100,00
4.	N1-N2	0,66	0,318	1,000	0,037	N1-N2	98,44
5.	N2-N3	0,66	0,318	1,000	0,037	N1-T3	98,44
6.	N2-S1	0,66	0,318	1,000	0,037	N2-N3	98,44
7.	N2-T2	0,66	0,318	1,000	0,037	N2-S1	98,44
8.	N3-T1	0,66	0,318	1,000	0,037	N2-T2	98,44
9.	T1-S1	0,66	0,318	1,000	0,037	N3-T1	98,44
10.	T1-T2	0,66	0,318	1,000	0,037	N3-T3	98,44
11.	N2-S3	0,55	0,108	0,994	0,226	T1-S1	98,44
12.	N2-T1	0,48	-0,129	1,000	0,313	T1-T2	98,44
13.	N2-S2	0,37	-0,005	0,741	0,190	T2-T3	98,44
14.	N1-S3	0,32	-0,158	0,789	0,242	T3-S1	98,44
15.	N3-S3	0,32	-0,158	0,789	0,242	N1-3	96,88
16.	S3-S1	0,32	-0,158	0,789	0,242	N1-S1	96,88
17.	T2-S3	0,32	-0,158	0,789	0,242	N1-T2	96,88
18.	T1-S3	0,25	-0,189	0,694	0,225	N2-T1	96,88
19.	S2-S3	0,23	-0,110	0,578	0,176	N2-T3	96,88
20.	N1-S2	0,20	-0,136	0,536	0,171	T1-T3	96,88
21.	N3-S2	0,20	-0,136	0,536	0,171	N1-T1	95,31
22.	S2-S1	0,20	-0,136	0,536	0,171	N2-S3	95,31
23.	T2-S2	0,20	-0,136	0,536	0,171	N1-S3	93,75
24.	T1-S2	0,16	-0,166	0,482	0,165	N3-S3	93,75
25.	N1-T3	0,00	0,000	0,000	0,000	S1-S3	93,75
26.	N2-T3	0,00	0,000	0,000	0,000	T2-S3	93,75
27.	N3-T3	0,00	0,000	0,000	0,000	T1-S3	92,19
28.	T1-T3	0,00	0,000	0,000	0,000	T3-S3	92,19

	Kombi	Kappa	KI95%		SF	Kombi	r%Präs
29.	T2-T3	0,00	0,000	0,000	0,000	N2-S2	90,63
30.	T3-S1	0,00	0,000	0,000	0,000	N1-S2	89,06
31.	T3-S2	0,00	0,000	0,000	0,000	N3-S2	89,06
32.	T3-S3	0,00	0,000	0,000	0,000	S1-S2	89,06
33.	N1-N3	-0,02	-0,038	0,006	0,011	T2-S2	89,06
34.	N1-S1	-0,02	-0,038	0,006	0,011	T1-S2	87,50
35.	N1-T1	-0,02	-0,038	0,006	0,011	T3-S2	87,50
36.	N1-T2	-0,02	-0,038	0,006	0,011	S2-S3	85,94

Tabelle 45. Analyse der Reflexaktivität. Übersicht über die Reliabilitätskoeffizienten *gewichtetes Kappa* (K_w) mit dem dazugehörigen 95%igen Konfidenzintervall (KI95%) und Standardfehler (SF) sowie *Prozentuale Übereinstimmung* ($r\%_{Akt}$) aller möglichen Beurteilerkombinationen (Kombi) in Studie 2b. Die Sortierung erfolgt absteigend.

	Kombi	K_w	KI95%		SF	Kombi	$r\%_{Akt}$
1.	T2-T3	0,80	0,308	0,739	0,110	N1-N2	84,38
2.	N1-N2	0,77	0,505	0,797	0,075	T2-T3	84,38
3.	T1-S 1	0,77	0,530	0,819	0,074	N1-T3	76,56
4.	T1-S3	0,76	0,471	0,759	0,073	N1-T2	73,44
5.	T1-S2	0,75	0,649	0,885	0,060	N2-T2	73,44
6.	T1-T2	0,74	0,629	0,853	0,057	N2-T3	73,44
7.	S2-S3	0,69	0,413	0,767	0,090	T1-S1	73,44
8.	T1-T3	0,68	0,479	0,790	0,079	N3-T2	70,31
9.	N2-T1	0,68	0,515	0,825	0,079	T1-S3	70,31
10.	T2-S1	0,68	0,640	0,907	0,068	T1-T2	70,31
11.	N2-T2	0,67	0,385	0,730	0,088	N1-3	68,75
12.	N1-T2	0,66	0,550	0,805	0,065	N2-N3	68,75
13.	N1-T3	0,66	0,403	0,689	0,073	T2-S1	68,75
14.	N3-T2	0,65	0,393	0,751	0,092	T3-S1	68,75
15.	N1-T1	0,64	0,476	0,759	0,072	N2-S1	67,19
16.	N2-S1	0,63	0,405	0,689	0,072	N3-T3	67,19
17.	T3-S1	0,63	0,410	0,792	0,097	T1-T3	67,19
18.	N2-T3	0,62	0,650	0,876	0,058	N2-T1	65,63
19.	T2-S3	0,62	0,554	0,830	0,071	N1-S1	64,06
20.	N3-T1	0,62	0,391	0,829	0,112	N1-T1	62,50
21.	S2-S1	0,61	0,415	0,669	0,065	S1-S3	62,50
22.	N1-S1	0,60	0,470	0,712	0,062	N3-T1	60,94
23.	S3-S1	0,60	0,409	0,670	0,066	N3-S1	59,38
24.	T2-S2	0,59	0,677	0,829	0,039	S1-S2	57,81
25.	N2-N3	0,59	0,379	0,643	0,067	N2-S3	56,25
26.	T3-S3	0,59	0,486	0,782	0,076	T2-S3	56,25
27.	N3-S3	0,57	0,451	0,721	0,069	T3-S3	56,25
28.	N3-T3	0,57	0,429	0,715	0,073	N3-S3	54,69

	Kombi	K_w	KI95%		SF	Kombi	r%_{Akt}
29.	N2-S3	0,56	0,563	0,801	0,061	S2-S3	54,69
30.	N1-N3	0,55	0,471	0,773	0,077	T1-S2	53,13
31.	N1-S3	0,55	0,688	0,913	0,058	N1-S3	51,56
32.	N2-S2	0,55	0,509	0,806	0,076	N3-S2	50,00
33.	N3-S2	0,54	0,419	0,689	0,127	N1-S2	48,44
34.	N1-S2	0,54	0,538	0,749	0,054	N2-S2	43,75
35.	N3-S1	0,52	0,477	0,729	0,064	T2-S2	43,75
36.	T3-S2	0,51	0,531	0,785	0,065	T3-S2	43,75

10. Vorabveröffentlichungen

Teile dieser Dissertation wurden bereits vorab nach Genehmigung veröffentlicht:

Scientific paper:

Giebels, F.; Kohn, B.; Shihab, N.; Volk, H. A.; Loderstedt, S. (2014):
Evaluation of the biceps tendon reflex in dogs
J Vet Med Res 1(3): 1013.

Vorträge:

Giebels, F.; Kohn, B.; Pieper, L.; Loderstedt, S.:
Vergleichende Untersuchung des Interobserver Agreement in der Beurteilung des Patellar- und Bizepssehnenreflexes beim Hund
62. Jahreskongress der Deutschen Veterinärmedizinischen Gesellschaft für Kleintiermedizin,
Berlin, 27. – 30.10.2016.

Giebels, F.; Kohn, B.; Loderstedt, S.:
Comparative reliability-analysis in the evaluation of the canine patellar tendon- and bicipital reflex
9. Doktorandensymposium und DRS Präsentationsseminar “Biomedical Sciences”,
Berlin, 16.09.2016.

Giebels, F.; Kohn, B.; Loderstedt, S.:
Evaluierung des Bizepssehnenreflexes beim Hund
59. Jahreskongress der Deutschen Veterinärmedizinischen Gesellschaft für Kleintiermedizin,
Berlin, 06.-10.11.2013.

Giebels, F.; Kohn, B.; Loderstedt, S.:
Subjective evaluation of the bicipital reflex in the dog
ACVIM Forum, Seattle, Washington, 12. - 15.06.2013.

Poster-Präsentationen:

Giebels, F.; Kohn, B.; Pieper, L.; Loderstedt, S.:
Comparison of the interobserver agreement in the evaluation of the patellar tendon- and bicipital reflex in dogs
ECVN Symposium, Edinburgh, 15. - 17.09.2016.

Giebels, F.; Kohn, B.; Loderstedt, S.:
Differences in interobserver agreement of the bicipital reflex evaluation in dogs based on level of expertise
ECVN Symposium, Paris, 26. – 28.09.2013.

11. Danksagung

Ich möchte mich zunächst herzlich bedanken bei Frau Prof. Barbara Kohn für die hervorragende und geduldige Betreuung dieser Doktorarbeit. Des Weiteren bedanke ich mich sehr bei Herrn Dr. Shenja Loderstedt für die Vergabe des Themas, seine Unterstützung während der recht langen Korrekturphase und darüber hinaus. Ein großes Dankeschön gilt selbstverständlich auch Frau Dr. Laura Pieper vom Institut für Epidemiologie und Biometrie der FU Berlin, die mir eine sehr große Hilfe bei der Bewältigung der statistischen Hürden war.

Allen Beurteilern, namentlich Shenja Loderstedt, Holger Volk, Nadia Shihab (Neurologen), Karina Baumgart, Sinah Dohlen, Anna Dettling, Aleksandra Chirek (praktizierende Tierärzte) und Marianne Walter, Constanze Dethoff, Sophie Wasmuth und Hans Morgenroth (Studenten) gilt mein außerordentlicher Dank für die Zeit, die sie aufgewendet haben, die beiden Filme auszuwerten. Außerdem danke ich selbstverständlich allen Hundebesitzern für die Teilnahme an dieser Studie.

Mein allergrößter Dank gilt meiner Familie, allen voran meiner Freundin Kristina, die mir über die gesamte Bearbeitungszeit den Rücken freigehalten hat. Ohne ihre Unterstützung wäre diese Arbeit wohl nicht beendet worden. Ebenfalls möchte ich meinem Sohn Jakob danken, der während der Bearbeitung dieser Arbeit geboren wurde, mir sehr viel Kraft gegeben hat und unfreiwillig auf ganz viel gemeinsame Zeit mit seinem Vater verzichten musste. Ich danke meiner Mutter und meinen Großeltern, die mich stets auf meinem Weg bekräftigt haben, die mir ohne zu zögern finanziell unter die Arme gegriffen haben, wenn es nötig war und die mir beigebracht haben fleißig und hartnäckig seine Ziele zu verfolgen.

12. Selbstständigkeitserklärung

Mit meiner Unterschrift bestätige ich, dass ich die vorliegende Arbeit selbstständig angefertigt habe. Ich versichere, dass ich ausschließlich die angegebenen Quellen und Hilfen in Anspruch genommen habe.

Berlin, den 08.06.2018

Felix Giebels



mbvberlin mensch und buch verlag

49,90 Euro | ISBN: 978-3-86387-906-8