Touching Reliability

Master Thesis zur Erlangung des Grades

Master of Science in Osteopathie

an der

Donauuniversität Krems

niedergelegt von

Peter Sommerfeld

September, 2006

Betreuung: Prof. Paul Klein, PhD, DO
Université libre de Bruxelles

– Erklärung –

Hiermit versichere ich, die vorgelegte Masterthese selbständig verfasst zu haben. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten oder nicht veröffentlichten Arbeiten anderer übernommen wurden, wurden als solche gekennzeichnet. Sämtliche Quellen und Hilfsmittel, die ich für die Arbeit genützt habe, sind angegeben. Die Arbeit hat mit gleichem Inhalt noch keiner anderen Prüfungsbehörde vorgelegen.

Datum Unterschrift

– Contents –

,Abstract

Touching Reliability

Peter Sommerfeld

Objectives: The allegation of subjectivity makes reliability of manual assessment procedures a special matter of concern within the medical community. The Cranial Concept is a controversial but integral part of Osteopathy dealing with subtle phenomena that are supposed to be palpable. One of the main principles of the Cranial Concept is the Primary Respiratory Mechanism (PRM). The PRM is hypothesised to be a palpable physiological phenomenon that occurs in rhythmic cycles, called flexion-phase and extension-phase, which are independent from cardiac and respiratory rates. PRM-palpation generally opens assessment according to the Cranial Concept. A theoretical analysis unmasks the allegation of subjectivity being a scientistic argument and works out several aspects behind the concept of reliability referring to clinical and didactical aspects in manual medicine.

Design: An inter- and intraexaminer reliability study design for repeated measures has been used to test reliability in PRM-palpation. In addition possible effects of the examiners' and subjects' respiratory rates on the palpated PRM-rates were tested. The PRM rates were recorded by using silent foot switches. The respiratory rates have been recorded simultaneously by using strain gauges.

Participants: 49 healthy subjects have been palpated by two experienced examiners simultaneously twice at the head and the pelvis.

Main outcome measures: PRM-frequency (f), the mean duration of the flexion phase and the mean ratio of flexion- to extension-phase have been described as the main outcome-parameters. Inter- and intraexaminer reliability and correlations to the respiratory rates were analysed for all three parameters.

Results: Inter- as well as intraexaminer agreement could not be described beyond chance agreement, as the 95% limits of agreement showed an expected difference (e.g. for f ± 3.3 cycles/90 sec) which for all cases resembled the total range of values (e.g. for f 7 cycles/90 sec) that has been produced. A significant effect of the examiners' respiration was found for both examiners at the pelvis ($P = 0.004$ for one examiner, $P < 0.0001$ for the other examiner), for one examiner only at the head ($P = 0.0017$). No correlation could be found for the subjects' respiratory rates.

Conclusions: PRM-rates could not be palpated reliably and under certain conditions were influenced by the examiners' respiratory rates. These results do not support the hypotheses behind the PRM. The role of PRM-palpation for clinical decision making and the models explaining the PRM should therefor be thought over. What examiners perceive, when they intend to palpate the PRM, resembles autonomous nervous system regulations, known as Traube-Hering or Traube-Hering-Mayer waves.

Key words: reliability; palpation; Primary Respiratory Mechanism

Corresponding author:

Peter Sommerfeld, DO
Karl Bodingbauerstraße 23/4
A-2100 Korneuburg
AUSTRIA
peter.sommerfeld@utanet.at

Je mehr die metaphysische Entscheidung gesichert, impliziert, versteckt ist, desto mehr herrschen die Ordnung und die Ruhe über die methodologische Technizität.

[The more the metaphysical decision is assured, implied, hidden, the more order and peace are governing methodological technicity]

(Jacques Derrida, *Limited Inc.*, p 148; translation by myself P.S.)

[…] unde sicut ens dicitur unum in quantum est indivisum in se ita dicitur aliquid in quantum est ab aliis divisum. Alio modo secundum conventientiam unis entis ad aliud, et hoc quidem non potest esse nisi accipiatur aliquid quod natum sit convenire cum omni ente;

[hence being, inasmuch it is individed in itself, is called One, as much as it is divided from others. In other words this is according to the correspondence of one being with another one, which is impossible if we do not consider something that can't correspond with every being;]

(Thomas Aquinas, *De veritate*, p 7; translation by myself P.S.)

1 Aim and intention of the thesis

Addressing a common background

During the 19th century medicine moved closer to natural sciences and its methodological paradigms (von Engelhardt, 2005). At the beginning of the 21st century we may ask if medicine finally has lost (or thrown over) its reputation being an art and has become an integral part of natural science. However, we do not share this attitude. As Wieland (2004, p 24; translation by myself, P.S.) points out, „[medicine] *does not intend to gain knowledge about a part of natural or social reality, in fact it intends to act consciously and deliberately within this reality*". Following Pöltner (2002, p 22; translation by myself, P.S.) we state that „*knowledge having been collected for the case of acting (practice) is practical knowledge. Hence considering its essence medicine has to be specified as practical science.*" Despite these definitions different approaches towards a characterisation of the medical profession can be observed nowadays. Amongst these the scientific perspective could be regarded as the most popular. This might be one of the reasons why manual methods in assessment and treatment have lost a considerable amount of reputation. An additional cause can be the development of technical appliances in medicine during the last century (Lewit and Liebenson, 1993). In spite of that, a few groups amongst physicians and therapists have made the development of palpatory possibilities and abilities in diagnostics and therapy a main matter of concern. They can be looked at as the stake holders of the artistic part within the medical professions. Among these professions besides Manual Medicine, Chiropractics, a great amount of concepts within Physiotherapy and massage techniques, we also find Osteopathy. A part of Osteopathy, initiated by W.G. Sutherland (1873–1954), is the *Cranial Concept*.

This project is placed within the mentioned tension between scientific and practical attitudes. The controversial position of palpation within the diagnostic and therapeutic tools specifically used in the *Cranial Concept* (Wirth-Patullo and Hayes, 1994; Hanten et al., 1998; Rogers et al., 1998; Hartman & Norton 2002; Moran & Gibbons 2002) motivated us for further experimental analysis. In this project we try to assess the agreement within two examiners (osteopaths) concerning the palpation of a highly controversial and well discussed core element of the Cranial Concept: the *Primary Respiratory Mechanism*. We want to provide information about the reliability of manual detected data that represent a basic part in the assessment procedures of the Cranial Concept. This might be of additional interest in respect to controversial aspects concerning the physiological basis of the *Primary Respiratory Movement* (Ferre & Barbin, 1990; Green et al., 1999; Hartman and Norton 2002; Hartman, 2006) and the fact that the observation of that phenomenon is taking part near the limits of tactile perceptible events (Upledger and Vredevoogd, 1994).

Addressing the specific osteopathic background

Manual assessment and treatment plays a dominant role in osteopathic clinical practice. This is why Osteopathy can implicitly be ascribed to manual orientated therapeutic concepts. Since the very beginning of osteopathic medicine – comprehending itself as an art – osteopaths always have tried to improve manual approaches in diagnosis and treatment. Their endeavour is heading for improvement and widening of the area of manual capabilities in a practical sense as expressed in the greek term *techné*. So they go for subtle differences in tissue-tension, changes in mobility of different structures etc. by improving their own tactile and propriozeptive senses and thus making those new abilities meaningful and useful for clinical decision making and treatment. Specific osteopathic manual assessment techniques are the expression of that development. Among those we find *Listening-techniques* for the perception of local or generalised tissue-tension, the group of *Induction-techniques* for diagnostic clarifying of subtle constellations within structures and tissues, *Manual thermo diagnostics* for finding out differences of body surface temperature as possible diagnostic signs for underlying lesions or the big group of *Multi component techniques* which allow smaller amplitudes in high velocity thrusts and in the consequence make these techniques more save and effective.

There are special questions arising in this context. They are dealing with the accuracy that can be reached in palpation as well as with the limits of tactile perception. Where do we take the „gold standard" from in respect to that validation of accuracy takes place? Validation always needs a reference-frame, a benchmark. Underlying we find the old problem of the relationship between an observing subject and its observed object. In addition we are facing the tricky discussions about perception and the description of what we perceive. So the discussion may end up in questions concerning the relation between language and the world. This can become a main issue in the pedagogical field regarding the teaching of specific techniques on the one hand and in data exchange among colleagues in clinical every day life on the other hand (Johnston, 1982, p 44).

Despite the problems it produces the asserted precision and subtlety in manual dominated therapeutic and diagnostic proceedings can also be mentioned as one of the special advantages and characteristics of the osteopathic approach. This is especially true for the increasing field of functional disturbances that in many cases cannot be adequately clarified by means of classical medical diagnostics like X-ray, MRI, CT as well as various electro-diagnostic tools (like EMG, EEG etc.) or by blood tests etc. In this regard the sensitivity of osteopathic manual techniques seems to be a major prerequisite for a diagnostic approach reflecting the patient's individual context. This affects one aspect of the osteopathic term *Dysfunction*, which „*cannot* [be] *precisely defined because mere words cannot adequately describe a whole body of observations*" (Northup G.W. in: Patterson, 1990). Palpation plays a major role within this body of observations. Lewit and Liebenson (1993) mention in this context that e.g. myofascial pain without any proven pathology can only be diagnosed by palpation. This might be true for a whole variety of disturbances concerning soft tissues.

A part of the osteopathic approach, that deals with utmost subtle phenomena being most sensitive against tactile perception, is the Cranial Concept. In order to avoid confusions, this term, introduced by Sutherland (A.S. Sutherland and A.L. Wales, Eds.

1998) and Magoun (1951), will be used throughout the whole paper. Over the years different terms like *Craniosacral Osteopathy*, *Biodynamic Cranial Osteopathy* or *Craniosacral Therapy* etc. have occurred. They are but more or less describing the same and can be regarded as identical with the original ideas of the Cranial Concept. All of them are built on the principles, hypotheses and systematics of Sutherland and Magoun.

Besides the great demands that are made on the perceptive skills of the osteopath in practical work, the basic theories of the Cranial Concept are still highly controversial (Ferre and Barbin, 1990; Rogers and Witt, 1997; Green et al., 1999; Klein 2002; Hartman and Norton 2002; Downey, 2004; Hartman, 2006). A detailed report on these fundamentals as well as the problems and discussions that arise from them will follow later on. However, the Cranial Concept seems more and more being used in clinical practice. Not only osteopaths but also other therapeutic approaches seem to recognise a system with clinical relevance in its theoretical and practical sense (Abhesera, 2001).

The following theoretical introduction is about common problems of manual assessment procedures and the possibilities as well as the reasonableness and usefulness of the scientific concepts of objectivity and comparability. These general remarks are accurate for the Cranial Concept at a very high degree. Their relevance for the Cranial Concept has rarely been subject to distinct investigation. This is why we chose tactile skills demanded by the Cranial Concept for this research-project. Further on there will be given a survey on the basic elements of the Cranial Concept with a special emphasis on the idea of the Primary Respiratory Mechanism as well as a critical review on the literature dealing with the Cranial Concept.

2 Introduction

2.1 About the acceptance of manual findings

2.1.1 Criteria for a scientific approach

Clinical findings are more or less guiding the therapeutic intervention. This is why they are of fundamental importance regarding the osteopaths' responsibility to their patients. In addition to other information they guide the modern „Ulyssean" clinician safely and efficiently between *Skylla* and *Charybdis* of ineffectiveness and negligence. Being one of the main tools of the osteopath, the investigation of manual findings under scientific conditions seems highly reasonable. But what is meant by saying „scientific"? „Scientific" means an adaptation of methodological criteria from natural sciences. Following Cartesian paradigms the mathematical method dictates the object of observation. As we are dealing with assessment-procedures we have to adapt them to the principles of measurement. Identifying palpation as a measurement procedure is in fact a kind of violation. We actually have to distort the phenomenological horizon of human perception down to the mathematical determinism of a measuring device. This is the price we have to pay for this investigation. We have to keep this in mind for the final discussion.

Measuring in experimental natural sciences is founded on two main principles: *validity* and *reliability*. Being a little bit sophisticated we may argue that these principles are platonian ideas, pure metaphysics. They can never ever be matched to their full extent in empirical reality. Validity deals with an ontological problem: the relation between the measurement (and its interpretation!) and what we call the „real world". The measurement should measure what it is said to measure, i.e. it should mirror some empirical fact, something that „is" (gr.: *on*). But who tells us – beyond any doubt – what the empirical fact is like? We are entering an infinite regressus running up – following Descartes – to a dogmatic break off: God. Following clinical medical practice we install a „gold standard" (the adequate name would be „god standard") that seems to give certitude *on what there is* (see the brilliant homonymous essay by Quine, 2003). The problem is, gold standards change. So there is no reason for any belief, that the one we are referring to now is the final (divine) one.

But this study is not about validity. Adhere we face reliability. Reliability – we are going back to our well known Platonian world – is built on the heritage of repetition. In fact it has to be a special kind of repetition: repetition of the same. Reinforcing the argument we could say: repetition of the identical. Having to deal with repetition we have to deal with time and space in addition. As time seems to pass by, a sentence like „the repetition of the identical" seems rather paradoxical. Being reliable in the end does come up to anticipating the same ever and ever again. This foreshadows the permanence of Platonian ideas resting the same till the end of time. But Plato already stated that our empirical world, the world as we perceive it, does in fact change all the time. In a chronometrical sense it is never ever the same every two discrete moments. This motivated him to propose that the world we seem to live in is a „non-being" world (gr.: *me on*). Quine (2003, pp 1–2) calls this doctrine *„Plato's beard"*: *„This is the old Platonic riddle of nonbeing. Nonbeing must in some sense be, otherwise what is it that there is not?"* Now we really are in a mess. We are heading for the repetition of the same in a constantly changing world that seems to be a non-being world. Maybe Plato would laugh at us at the moment. Anyway, we are motivated to try to „impossible the possible" (pardon – this is irrationalism).

So what we are trying for seems rather impossible. In order to find a way through we have to meet certain concessions and we always have to bear in mind the metaphysical basis that is grounding the scientific concept of reliability or validity. So let's have a look at the usefulness of the concept of reliability in the practical world of the clinician. Portney and Watkins (2000, p 61) point out that

> »the usefulness of measurement in clinical research and decision making depends on the extent to which clinicians can rely on data as accurate and meaningful indicators of a behaviour or attribute. The first prerequisite, at the heart of measurement, is *reliability*, or the extent to

> which a measurement is consistent and free of error. Reliability can be conceptualized as reproducibility or dependability. If a patient's behaviour is reliable, we can expect consistent responses under given conditions.«

Regarding this description we may find three major parts within this concept: *reliability* itself, *consistency* and *reproducibility*. The terms seem to be synonyms but they are not. They are more or less focusing on the same thing under different perspectives.

Consistency

Consistency comes from a logical background. Being consistent means being free of contradictions. This is one of the main pillars of Aristotelian logic known as the *principle of contradiction*: A cannot be Non-A. Following this principle the meaningfulness of clinical findings or the usefulness of therapeutic techniques should be free of contradictions. Different examiners should e.g. on the average agree on the existence or absence of positive findings. Considering two examiners *A* and *B* we may say: if examiner *A* identifies a dysfunction in joint *X*, examiner *B* should also find a dysfunction in the same joint without knowing about the findings of *A*. Due to methodological implications it is important to have a certain variability within possible findings (values). They have to show at least two possibilities: positive or negative (e.g. „Yes, there is a dysfunction" or „No, there is no dysfunction"). If the collected data don't show any or just too little variance (e.g. all examiners find no dysfunctions in all subjects) one seems to infer that there is perfect consistency and agreement among the observers because they do not contradict each other. There should also be agreement about the spectrum of indications and effects for techniques that are commonly used. If examiner *A* judges technique *X* in the case of dysfunction *Y* of a subject *Z* as adequate, examiner *B* should not suggest that technique *X* is contra indicated in the case of dysfunction *Y* of subject *Z*. In simple words we may say that consistency has to do with reliability on a logical basis. Regarding different examiners consistency guarantees smooth communication. Regarding the patient it prevents confusion and provides good collaboration for treatment.

Reproducibility

Reproducibility is right in the centre of what we described as „repetition of the same" before. Multiple assessments or ratings should be repeatable under the „same conditions". This is a metaphysical demand and we have to diminish its strength to „at least similar conditions". This is true for repeated measurements of one examiner as *intraexaminer reproducibility* as for measurements that are done by different examiners in the sense of *interexaminer reproducibility*.

One major problem with reproducibility is the similarity of the conditions under which measurements are taken. It addresses the concept of time and space as mentioned before. If we want to talk about reproducibility we have to act on the assumption that at least two measurements have taken place. We may focus the whole problem in one statement: it is impossible to do two measurements at the same time at the same place. Good old Leibniz (1996, p 36) might shed light on our problem by giving the following sharp definitions:

> »Zeit ist die Ordnung des nicht zugleich Exitierenden […]
> Raum ist die Ordnung des Koexistierenden, oder die Ordnung der Existenz für alles, was zugleich ist.«
>
> [Translation by myself P.S.:
> »Time is the order of non-synchronous existence […]
> Space is the order coexistence or the order of existence for everything that is at the same time.«]

Either we get two measurements at the same place at different times or we get two measurements at different places at the same time. Considering the first case the demand for „repetition of the same" in biodynamic systems is a pure contradiction as life itself is an ongoing process of becoming. Considering the second issue we might have two examiners assessing simultaneously. Again they can't find the same as they are assessing different regions. We will meet both problems again in the methodological part of the thesis as *unadjustment* of time or location (see point 6.5.3, p 81ff).

For intending the best possible similarity of conditions for all measures, which means de facto a certain constancy in the conditions of the rated subjects, all examiners should rate the same subjects as simultaneously as possible. Simultaneous assessment may be important in order to rule out possible therapeutic effects that can occur as soon as the examiner puts her/his hands upon the subject. In many cases simultaneous ratings are impossible, because intraexaminer blindness cannot be guaranteed. In order to blind an examiner against himself, repeated ratings have to be done anyway.

Thus specific findings *X* of examiner *A* on subject *Z* at a certain moment t_1 should also be found by examiner *B* at nearly the same moment t_2 (in which $t_1 \neq t_2$), while *B* shouldn't know anything about the findings of *A*. Following the cited demands, specific findings coming from one subject should be comprehensible (in the sense of consistency) for a homogenous group of observers. The influence of the period of time between the ratings on the reproducibility of the rated data is related to the severity of the rated substrata's deviation from normal physiological conditions. For instance, in the case of a severe arthritis with great loss of mobility and definite symptomatic signs, we can assume similar results from different examiners for repeated ratings within big time-intervals (maybe over months). Considering the realm of functional disorders, where loss of mobility or pain may not be so distinct and the body's own regulatory mechanisms might influence the situation, it becomes much more difficult to find similar conditions. Connall et al. (1980) report that interexaminer agreement was slightly moderate for the affected segments of patients with cervical or lumbar pain. But there was low or no agreement for asymptomatic subjects.

Reliability

Consistency and reproducibility are important preconditions for reliability. In the sense of trustworthiness of data, reliability refers beyond the first two criteria. It brings in a momentum of safety that is needed in clinical practice. Reliability refers to the clinician's responsibility. It prevents clinical acting from becoming arbitrary. Reliability spans the gap from unattainable perfect metaphysical options via attainable imperfect empirical circumstances to practical ethical implications. As clinicians we always have to end up with a decision for therapeutic acting facing our patients. Sometimes there is little time and a small or confusing amount of information. So we are looking for information we can rely on as good as possible at the moment.

Thus the analysis of techniques for assessment and treatment following the scientific concept of reliability can provide helpful information for clinical decision-making. It may offer an advice on the strength of the meaning of certain findings as well as on the clinical usefulness of techniques a clinician is going to apply. Values based on personal experience, being imparted by teachers and colleagues, can be rounded up by means of adequate research. The definition of criteria for reliability (and validity) should support the possibility for differentiation between efficient and inefficient diagnostic and therapeutic tools and define a certain spectrum of meaningful indications. This can be seen as one of the main applications of the aspect of reliability. Reliability is therefore determinable by assessing repeated measures under the criteria of consistency and reproducibility.

The demand for consistency, reproducibility and reliability is not true for manual findings alone. Any kind of observation (in this context referring to clinical acting) like case history, the whole range of instrumental medical diagnostic methods, inspection, neurological and orthopaedic tests etc. are – as criteria for clinical decision-making – to be considered critically under these aspects too.

2.1.2 Facing some fundamental problems

In order to question their clinical relevance, manual techniques are repeatedly criticised (Lewit and Liebenson, 1993; Fitzgerald et al. 1994). Therefore the following chapter brings a short discussion of some of the classical problems in connection with palpation and manual techniques in the clinical field.

Perception and knowledge – a short story of epistemology

One main target for criticism against manual assessment procedures is the fact, that to a great extent the examiner relies on – what is called – *subjective* phenomena of perception. This is a neo-positivist point of view. We may ask at that moment: is there anything like objective perception or objective experience? In fact, speaking about perception or experience, there always must be a subject around. At least someone who is able to understand and express the data produced by a measuring device that again must have been planned and constructed by a subject. So finally we have come to the point where subject and object have to meet. What now? One solution has been tried by Ludwig Wittgenstein's logicistical *representation theory* as developed in his *Tractatus logico-philosophicus* and the program of *logical-empiricism* represented by the members of the *Viennese Circle* (ger.: *Wiener Kreis*). This movements were heading for the excision of the subject, a clean world, free of metaphysics. In order to get rid of any psychological and other subjective ballast the young Wittgenstein (1984, p 38) ended up with *elementary propositions* (ger.: *Elementarsätze*):

> »4.21 The simplest kind of proposition, an elementary proposition, asserts the existence of a state of affairs.«

Wittgenstein's elementary propositions brought about the *protocol-statement discussion* (ger.: *Protokollsatz-Diskussion*) of the *Viennese Circle*. *Protocol statements* should just contain the pure protocol of the observer. But statements always need elements of some kind of language. So we would need a language that is the object's language in the sense of producing true propositions about the object. But how can we be sure about that? Otto Neurath, one of the members of the *Viennese Circle*, thought about a scientific cleaning machine. You can insert statements into the machine and the machine will indicate any contradictions. As soon as a contradiction occurs, you have to change the inserted statement or … – the entire machine (cp. Zeidler, 2000, p 26ff). Following Gödel's theorem, truth never can be proofed „within" an axiomatic system. So you have to construct a cleaning machine for the cleaning machine of the cleaning machine and so on. Looking back we may state that the neo-positivists program failed. But positivism is a good example for what we get out of an „objective" experience.

So let's go back in history and put it the Kantian way now as he suggests in his *Critique of Pure Reason* (1929, B 62):

> »What we have meant to say is that all our intuition is nothing but the representation of appearance; that the things which we intuit are not in themselves what we intuit them as being, nor their relations so constituted in themselves as they appear to us, and that if the subject, or even only the subjective constitution of the senses in general, be removed, the whole constitution and all the relations of objects in space and time, nay space and time themselves, would vanish. As appearances, they cannot exist in themselves, but only in us. What objects may be in themselves, and apart from all this receptivity of our sensibility, remains completely unknown to us. We know nothing but our mode of perceiving them […]«

What ever the *object-in-itself* (ger.: *Ding an sich*) might be, we never ever will know because experience always depends on the *object-for-us* (ger.: *Ding für uns*), what Kant called *appearance* (ger.: *Erscheinung*). Appearances are the results of what the acting *intellect* (ger.: *Verstand*) did with the chaotic material coming from the senses. Appearances are something like the object within the subject. By reasoning about appearances the intellect puts objectivity into these appearances. This makes us consider in a way „as if" there were objects outside. But Kant was not an idealist in the sense of proposing objects would only exist in our heads. He just stated that objects outside cannot be recognised as they are (as objects-in-themselves). They only can be recognised as appearances that get treated by the *principles of pure reason* (ger.: *Prinzipien der reinen Vernunft*) – categories and schemata – and the acting intellect.

Besides positivism and Kantian epistemology there is another interesting attempt towards the problem of perception: phenomenology. In a phenomenological manner we can state that experience is always someone's experience, perception is always someone's perception. A machine cannot experience or perceive anything. In order to be able to even speak about experience or perception one has to have an understanding about what experience or perception does mean. Martin Heidegger calls this *pre-under-*

standing (dt.: *Vorverständnis*). The problem is that in most cases we are not aware of that implicit kind of knowledge. This is highly true for scientific and positivist approaches. By overseeing that their principles and axioms are built on *lifeworld* (ger.: *lebensweltlich*) experience, they are falling back into the same metaphysical realm they wanted to overcome. The positivist may build an apparatus for measuring what she/he calls „electromagnetic waves“. As soon as she/he is metering her/his apparatus and making the protocol statement „this is red light“ she/he has to know the meaning of „red“. This meaning comes from what Edmund Husserls called *Lebenswelt* or Heidegger conceptualised as *Vorverständnis*. Without knowing what „red“ is, the poor positivist never will be able to describe what the apparatus measures, whereas the term „electromagnetic waves“ reflects a metaphysical entity. So there is always a subject and an object together.

Like Kant the phenomenologists do not deny the being of external objects. There are objects. But we do not know what objects are „as objects“ (*objectum qua objectum*). Anyway, phenomenologically spoken this is a meaningless statement as there are only objects „for us“. So talking about objects is always a subject talking about an object. We can‘t tear them apart. Every attempt to dissolve this connection to one (*objectivity*) or another (*subjectivity*) direction is prepared to fail. Being inside (*subjectivity*) is always being outside (*objectivity*) at the same time as Heidegger explains (cp. Heidegger 1993, p 62): in a primary sense, a sense of *Being-open-to-the-world* (ger.: *Weltoffenheit*) we are ever dwelling „outside“ within a world that has already been discovered by us, that has met us; in perceiving we are not in first instance leaving an inner sphere where we might have been encapsulated; in the right sense our being – Heidegger‘s *Dasein* – has to be understood as *Being-in-the-world* (ger.: *In-der-Welt-sein*).

So if subjectivity should be an allegation, as proposed in the beginning, it is true for any statement. But there is another weapon, now coming from the scepticists‘ realm. It is called illusion. Every perception, some scepticists tend to say, contains illusion. In consequence we have to doubt any perceived data. There was one philosopher who made doubting a core principle of his methodology: René Descartes. Towards the end of his *1st Meditation* he carries scepticism to extremes (Descartes, 1986, p 72):

> »Supponam igitur non ‚optimum' Deum, fontem veritatis sed genium aliquem malignum, eundemque summe potentatem et callidum, omnem suam industriam in eo posuisse, ut me fallerert«
>
> [translation by myself, P.S.:
> So I suppose that not an ‚utmost gracious' God, the source of truth, but rather some evil as well as powerful and crafty spirit concentrates all his attention on misleading me]

After the darkness of the *1st Meditation* Descartes starts his well known argumentation. He is finding his *fundamentum inconcussum* in what he called the *res cogitans* which is within himself (ibid., p 78):

> »Adeo ut, omnibus satis superque pensitatis, denique statuendum sit hoc pronuntiatum, *Ego sum*, *ego existo*, quoties a me proferetus, vel mente concipitur, necessario esse verum.«
>
> [translation by myself, P.S.:
> After considering the whole thing more than sufficiently, I have to state, that the sentence *I am, I exist*, is necessarily true, regardless how often I pronounce or deliber it.]

Despite Descartes' fundamentum, illusion is always a possibility. For him this is due to the defectiveness of our senses. This is why he is put into the rationalists' drawer. So we know that we might fail and end up with wrong statements. But this is not because of subjecitve pollution of a clean world of true objects. This is because statements on empirical facts are always hypothetical. The certitude and evidence Descartes was heading for, is beyond our cognitive faculty. So the statement that *„the objection to palpation being 'subjective' is further enhanced by the discovery of palpatory illusion"*, mentioned Lewit and Liebenson (1993, p 586), can now be answered questioning: Yes, but what is the problem?

Observation and description

> »Die Schematisierung des Gegenstandsbereiches einer empirischen Wissenschaft erfolgt durch *Beobachtung*. Durch Beobachtung wird ein empirischer Gegenstandsbereich erschlossen [...] indem die Beobachtung immer wieder Neues entdeckt. Das Neue muss freilich nicht nur entdeckt, es muss auch festgehalten werden. Würden Beobachtungen nicht festgehalten, dann würde sich die Beobachtung in der nächsten Beobachtung verlieren. Die Beobachtung drängt darum zur *Beschreibung*. Die Beobachtung ist ihrer Beschreibung bedürftig. Sie ist derart ergänzungsbedürftig, dass man sogar gemeinhin den Unterschied zwischen Beobachtung und Beschreibung nivelliert.«
> (Zeidler, 2000, p 126)
>
> [translation by myself, P.S.:
> »In empirical sciences the schematisation of a domain of objects results from *observation*. The domain of objects becomes accessible by observation […] as observation discovers news over and over again. However, just discovering these news does not suffice, they have to be recorded. By not doing so, one observation would get lost into the other. Therefore observation pushes for *description*. Observation is in need for description. This need for endorsement is so strong that in general the difference between observation and description is being levelled.]

Empiricists use to think that there is something like pure observation. As cited above, this is impossible. Observation only makes sense as a conscious act of perception. Otherwise we would talk about pure gathering of data without even knowing that we are gathering them. So there is no observation without description. As we mentioned above, there has to be someone who is able to „read" the outcome of any measuring apparatus. „Reading" (which in addition means somehow understanding) the outcome does imply being „inside language". Being inside language means being someone who is able to address to someone else and being addressed by someone else within an endless series of meaningful contexts. This stands for being within a cultural and social tradition, having history and behaving towards ourselves and the world. Looking back to the neopositivsts' *protocol statements* mentioned above we may once more proclaim the parousia of the subject. As soon as we are using words – even while we are thinking – the object's observation has lost its virginity, a virginity it never really had.

So finally the reasons why commonly confidence in tactile ascertained data is so little, may also be of cultural, social or other origin. Nevertheless a main aspect of the problem seems to be communicative of nature. When I (sorry, but we have to break the rules and use the non-word „I" now) state that the flower is red, I never will know whether the redness I am talking about is the same for someone who is agreeing with me. But as if by a miracle we seem to understand each other in many cases. But there are in fact a lot of cases where communication seems to collapse or tick over. In relation to our topic the problem lies in the mode of perception we are dealing with which is tactile or proprioceptive. As we are used to express visual experiences at a very high degree, ordinary language seems to reach limits as soon as we want to communicate about tactile or proprioceptive experiences.

The communication problem occurs specifically in practical training or in clinical practice. Under these circumstances qualitative descriptions or paraphrases are frequently being used among osteopaths and other manual therapists. Often metaphoric forms help to characterise the ascertained data. In simple cases terms like „blocked" or „compressed" etc. are used. The term „metaphor" indicates, that we have to transfer meaning. Sometimes this transfer is doubtful and we are running into difficulties. Does the term „blocked" for a joint mean, that there is no more movement between the articulatory partners or that mobility is just reduced or kinematics have changed? Scientifically spoken we are facing an inevitable inaccuracy of language. Scientifically the situation may sometimes be handled by introducing categorical (nominal, ordinal) and numerical (ordered, discrete) scales or specific scores. But this is a Phyrric victory as scales and scores may on the one hand indeed smoothen communication to some extent but on the other hand always are accompanied with loss of meaning and the tendency of losing the subject (human being) out of sight. Johnston (1982, p 45) hits the mark by stating that „*physicians elaborated on concepts of what they were finding, rather on how they were eliciting findings*."

There is another difficult issue concerning the physiological, functional or morphological substrata of the perceived phenomena. We can question about the parameters that define the normal conditions for instance during a movement in a certain joint or within the tension of certain tissues. Further on, we can ask how much the real position of a bone or the real tension or tension-pattern of tissue (what is the meaning of „real"?), as far as we can determine or measure it, corresponds to what the examiner is talking about after she/he palpated it. The more accurate the descriptions of palpatory findings and the more subtle and controversial the substrata underlying the various tactile observations in the sense of tension, motion etc. are, the bigger the problem of communicability and comprehensibility of subjective impressions becomes. The Cranial Concept deals with a lot of subtle and controversial procedures for assessment and treatment. Therefore this is of special importance for this thesis.

Communication within the teaching process

We want to refocus on the communication problem we mentioned already above, specifically dealing with didactical problems occurring in practical training in osteopathy and manual medicine. It is mainly the aspect of reliability that dominates this field and, derived from it, the definition of common or particular (under certain conceptual circumstances) criteria for validity (Johnston, 1982, p 46). The core question is: What can a student rely on?

The uncertainty of students in respect to their findings and conclusions is a common observation. Whilst instructors might be able to rely on many years of experience, completed by sound medical knowledge, students may just have seemingly incoherent and doubtful isolated findings at their disposal (sorry, we are somewhat inflating the instructors' competence; to tell the truth we can smoothly exchange students by instructors, but the situation of authority dominating the classroom produces the described gradient of competence which in many cases is pure staging or rite). So instructors and assistants are called for help in order to proof the rightness of the findings. The ideal picture many students have in mind is to find a certain situation which has been explained theoretically (e.g. a third lumbar vertebra in extension with rotation and side-bending to the right) as an objective truth in vivo. In addition there should be no controversy about this truth among different independent examiners (perfect reliability; as we see, students are often heading for a Platonian world). On the contrary students repeatedly have to experience, that the correctness of theoretical prime examples is not beyond doubt and that there is a continuous controversy going on about whether such an example or situation does clinically exist as an „objective" actuality (pp 19ff). This implies, that events or constellations are being observed, whose pure existence or characteristics are called into question or are at least potentially contestable. On the other hand there are always situations where manual findings diverge. In this case a lot of questions may arise. Who is right? Is there more than one correct possibility? How can one judge who is right? Commonly there is thought that more experienced observers are right (in first row the godlike instructor), but even they can fail (Mior et al., 1990). In the end most students realise during their education that the ideal picture of objective undoubted criteria as a basis for any clinical acting can only be kept under very limited conditions. Thus they are obliged to constantly renew their personal objectives. The most of us are aware of the fact that this isn't an easy undertaking especially under the premise of the biggest possible responsibility we owe our patients.

Having confidence in the confidence of ones own perception – a way out?

Osteopaths seem to have found a special way out of these difficulties by giving a big emphasis on intrasubjective consistency. This is formulated in the following concise and frequently used sentence: believe what you feel, don't feel what you believe. Thus the confidence in the correctness of the perception of the clinician who dominantly must rely on palpation should be strengthened. In many cases this demand can lead into an ethical dilemma. This is especially true for the inexperienced osteopath, but for the experienced one too, if she/he wants to keep a critical distance against her/his own believes and the methods she/he applies. How far is an osteopath allowed to follow her/his subjective impressions before running the risk of possibly being ineffective or doing harm to the patient? So we are facing *Skylla* and *Charybdis* again. The responsibility towards our patients and the cultural burden that is founding the doctor (osteopath)-patient-relationship, which latently implies the danger of abuse of power – during treatment the patient literally gives himself into our hands – put a special ethical emphasis on this issue.

Manual versus apparatus methods in medicine

So what we can do is washing our hands in innocence and transmitting the task on a *deus ex machina*. Taking reliability into account, there is a certain common consent that tactile perception as a diagnostic means in medicine ranks behind visual perception and far beyond data produced by an apparatus. Lewit and Liebenson (1993, p 586) state in this context:

> »The three basic assessment methods used in clinical medicine have always been inspection, auscultation and palpation. Recently, however, they have been overshadowed by modern apparatus and laboratory methods. The method that has suffered most by this development is palpation, which at present days plays a major role in manipulation and massage, but is almost ignored by the rest of the medical profession.«

Boline et al. (1993) have found, that assessment via technical equipment need not in every case to be reliable, especially in comparison with some tactile methods. These authors investigated the interexaminer reliability in diagnostic procedures for the lumbar spine. 28 patients with low back pain were assessed by three examiners using visual, palpatory and instrumental means. As instruments a portable EMG surface scanner and a portable dermothermograph were used. Interexaminer reliability has been computed by using the kappa-coefficient (κ) and percent agreement (see Table 1 p 25). Whereas palpation for pain and visual observation produced good to excellent interexaminer agreement, the two instruments reached only poor values and were judged as being unacceptable for clinical decision making.

Such results implicate the need for a critical discussion about reliability within the meaningfulness of commonly used assessment procedures in manual orientated concepts. This will help to realise the strengths and weaknesses in the diagnostic approach with regard to a more effective use of the own methods in relation to others. During the recent years this has been a rising demand by different traditions in manual medicine (Johnston, 1982). Keating et al. (1985, p 129) think about this situation, that „*despite its lengthy commitment to research, osteopathic medicine has not realised its potential as a clinical science, nor has it succeeded in integrating osteopathic research into clinical practice*. [...] *A consequence has been an unnecessary gap between the researcher and the clinician*."

Summary

Condensing the core issues lying underneath the topic of this project we can summarise the following intensely interrelated points:

- *The subject-object-discussion*; subject and object are individably interrelated. Speaking about an object being this and that is always a subject speaking. Hence the allegation that perception is subjective does not make any sense.
- *The relation between language and world*; every time we are observing something, we have to describe this observation in order to be able to perceive (which means being conscious of) what we are observing.
- *The communication problem*; we are living and behaving inside language. All we have at our disposal are signs. Signs are indicating endless possibilities of contexts in meaning. Every attempt to pin any sign down to a fixed meaning is metaphysics and leads into idealistic aberrations. So we have to face uncertainty in communication. As, unlike for visual experiences, we are generally not used to describe tactile or proprioceptive experiences, this enhances difficulties in the field of manual medicine. The problems affect the clinical and didactical field.

- *Ethical implications*; facing our patients we have to meet and ground decisions. As we are deciding inside an empirical world, every decision may fail. But we owe responsibility to our patients.

- *Justification for a scientific approach*; Mootz et al. (1989; p 440) note in this context that *„the role of palpable spinal dysfunctions in producing clinical syndromes is still unclear"* and think that this may be due to the *„lack of standardised, valid, reliable palpatory indicators of spinal joint dysfunctions"*. Schöps et al. (2000, p 2) think, that *„the status of manual assessment of the cervical spine could be optimised by establishing criteria for documentation, assessment criteria, as well as repeated controlled phases of training"*. Reliability studies concerning diagnostic procedures can aid the clinician in optimising the choice of therapeutic and diagnostic means, which in the end might comfort the patient's safety and wellbeing. So there is nothing bad about science or a scientific approach as long as we keep in mind that our self-conception as human beings is not the sum of scientific descriptions.

2.1.3 Methods for proofing the degree of agreement between different measures

One possibility of analysing the noted aspects and problems are known as *intra- vs. interrater* (*-examiner or -observer*) *reliability studies for repeated measures*. Findings from different examiners can be investigated with respect to their degree of *intra- and/or interindividual agreement* (*concordance*). Together with studies concerning validity this belongs to *methodological research* which *„involves the development and testing of measuring instruments for use in research or clinical practice. This approach is used extensively in health care research, as clinicians work toward establishing the reliability and validity of clinical measurement tools"* (Portney and Watkins, 2000, p 351).

This approach corresponds to the procedure of comparing two measuring methods for one independent variable. This is frequently used in the clinical field in order to judge the advantages vs. disadvantages of a new measuring method against an established one (Bland and Altmann 1986). McConnell et al. (1980, p 441) think that *„an essential element in the acceptance of a diagnostic modality is the level of agreement that can be achieved* *with its use by different examiners"*. Besides interexaminer agreement, intraexaminer agreement can be investigated too. So the consistency of individual findings can be judged. Different studies have shown that inter- and intraexaminer agreement need not be corresponding. Interexaminer agreement can be low whilst intraexaminer agreement is high. The results of such research projects should provide reliable criteria for the reliability of manual assessment-procedures in clinical every day life.

Specialities concerning data analysis

Problems regarding the way of executing data analysis in reliability studies are frequently noted. Mootz et al. (1989) say that, according to Alley (1983), reliability studies of palpation prior to 1983 suffer methodological and statistical deficiencies. In this context an interesting contribution comes from McConnell et al. (1980). These authors have compared two ways of data analysis in respect to their meaningfulness for assessing interexaminer reliability.

Bland and Altmann (1986) point out, that in the assessment of agreement between different measuring methods the terms *correlation* and *agreement* frequently are not clearly separated and lead to inappropriate statistical analysis. These authors suggest an alternative approach, plotting the data in a way which shows the difference against their mean and calculating the bias, estimated by the mean difference and the standard deviation of the differences. They point out that the use of *product-moment correlation coefficient* (r) or *interclass correlation coefficient* (ICC) is inappropriate for assessing agreement.

Haas (1991) is referring to the specific statistical problems concerning reliability studies in manual medicine. In the very beginning this author accentuates that percent agreement is misleading in judging reliability, because chance agreement is not taken into account. Haas suggests different methods, as *Cohen's kappa* (κ) for measuring concordance in nominal data (e.g. Yes/No decisions), *weighted kappa* (κ_W) for ordinal data (e.g. the severity of a dysfunction measured by a scale from 0 to 5). For continuous data the author suggests *analysis of variance* (anova) with the calculation of the *interclass correlation coefficient* (ICC), which can be interpreted in a similar way as *kappa* (0 would be no correlation at all, 1 would be perfect correlation).

Values for κ	Strength of agreement
< 0.20	poor
0.21–0.40	fair
0.41–0.60	moderate
0.61–0.80	good
0.81–1.00	very good

Table 1
Guidelines for the interpretation of kappa (κ) in respect to the agreement of two measurements. A value of 0 indicates no agreement, the value 1 indicates perfect agreement (according to Altmann, 1999).

As every statistical method has its limits which are tightly connected to the hypothesis, it's conclusions and the methods being used, professional advice about the choice of data analysis and interpretation is of essential importance in reliability studies.

How blind are studies in the field of manual assessment and therapy?

For the comparison of subjective measures, blinding the two observers against each other is an important precondition. In many cases simultaneous examinations are impossible to realise. If simultaneity is possible, the fact of assessing different zones can cause a methodological problem (unadjustment). When measures are done one after the other, therapeutic effects cannot be ruled out (Hawk et al., 1999). In this case one can assume that two examiners will not find the same condition. The more subtle the phenomena are, on which the diagnostic procedure is based on (which is especially true for palpation in the Cranial Concept), the more sensitive the assessment-situation becomes for future data interpretation.

2.2 Specific research on the agreement of manual findings

The examples given in the following chapter are taken from manual assessment of the musculo-sceletal system and cardiac rhythm palpation. They show that the results of reliability studies in these fields frequently show poor or missing interexaminer agreement. Thus should be underlined, that the whole problem is not only true for the Cranial Concept but for the whole area of assessments in manual medicine.

2.2.1 The reliability of manual findings of the musculo-sceletal system

In principle manual assessment besides internal (e.g. pulse diagnostics), orthopaedic (e.g. standing flexion test) and neurological (e.g. sensibility testing, reflex testing) tests can refer to different parameters like tissue tension, pain, skin characteristics (moistness, temperature, roughness etc.), trigger points, myofascial tensions etc. All these data are primarily perceived via cutaneous receptors of the examiner's hand. The assessment of the positioning of osseous structures as well as active and passive mobility testing and the quality of the end of the range of movement need propriozeptive perception including the whole posture of the examiner. The perceptive structure of the manual assessment therefore is primarily built of two tactile modal components who can be completed by the patient-therapist interaction:

- Exteroception for judging surface texture, temperature, moisture, tension etc.
- Proprioception for judging motion and tension–patterns
- The patient's verbal and / or muscular responses on the examiner's manual intervention in the sense of reflex, pain or irritation.

Passive motion testing, which means assessment of joint-mobility which is induced and executed by the therapist, takes a great part in functional osteopathic diagnostics. In many cases it represents the deciding factor, which structure in relation to another should finally be treated. The group of other parameters as mentioned above are leading the osteopath in a focusing process to a distinct place where the passive motion test as a final criteria determines the area and part of the quality of the therapeutic intervention. The results of the studies, represented below, show that especially passive motion testing achieves very low values concerning reliability. As the palpation of the Primary respiratory Mechanism (PRM) can also be seen as passive motion testing, these results may be relevant.

McConnell et al. (1980) compared the findings of six osteopaths who assessed the spines of 21 subjects with acute pain. The assessment procedure has not been standardised, as, following the authors' opinion, in clinical every day life an individual choice of techniques can be expected. The conformity of the examiners was given as all of them were graduates from osteopathic colleges. During the examination all spinal segments from C0 to S1 had to be assessed in respect to the presence/absence of a dysfunction by a scale ranging from 0 to 3. In addition questions concerning the planned treatment, diagnosis and examination had be answered. The authors carried out two ways of data analysis in order to accentuate the relevance of using the correct statistical design: (a) expected disagreements on a single patient based on the uniform distribution hypothesis, (b) agreement in cluster areas. Interexaminer agreement was found to be low and varied with regard to certain regions and acute segments.

Mootz et al. (1989) did specific research on the intra- and interexaminer reliability of passive motion palpation of the lumbar spine. The authors refer to previous studies who in general report low values which can differ depending on the assessment technique and the region. They report no interexaminer agreement for passive motion palpation ($\kappa = 0.06 – 0.17$). Intraexaminer consistency was also poor ($\kappa = –0.01 – 0.48$).

Keating et al. (1990) investigated the interexaminer reliability for eight evaluative dimensions of lumbar segmental abnormality. The segments Th11–S1 were assessed by three chiropractic physicians on 25 asymptomatic subjects via: (a) visual observation, (b) Dermatograph, (c) osseous pain, (d) soft tissue pain, (e) muscle tension, (f) static (misalignment) palpation, (g) active motion palpation, (h) passive motion palpation. In order to determine reliability, the most reliable dimensions and the region where agreement was best, kappa (κ), intraclass correlation coefficient (ICC) and product-moment correlation (*r*) were calculated. The authors found out that methods, commonly judged as the most „subjective" ones, such as osseous and soft tissue pain palpation, were the most reliable, whereas visual observation and the Dermatograph ranged behind. Low or no reliability was found for misalignment palpation, palpation of muscle tension, active and passive motion palpation. The results for passive motion palpation coincide with former studies.

The project of Boline et al. (1993) has been already mentioned before. The results confirm the high reliability of osseous and soft tissue pain palpation. Hubka et al. (1994) assessed the reliability of palpation for cervical spine tenderness in the intervertebral joint portion of the segments C2–C7 between two examiners for 30 patients with motion-dependent neck pain. The authors report high reliability for this diagnostic method ($\kappa = 0.68$; $p < 0.001$) and coincide with the results of Keating et al. (1990), who investigated the lumbar spine.

Hawk et al. (1999) note, that a combination of different diagnostic techniques as a basis for therapeutic intervention corresponds more to the „real life" clinical setting. In their preliminary study the authors assessed intra- and interexaminer agreement with respect to the question whether a certain spinal segment should be manipulated or not. Each examiner was allowed to use her/his own combination of techniques. The spinal segments of 20 subjects were examined from Th12/L1 to L5/S1. The authors found higher intraexaminer reliability which seemed to vary with the examiner's experience. Interexaminer reliability was low. Results and methods of this project are similar to the data reported by McConnal et al. (1980).

Schöps et al. (2000) assessed interexaminer agreement between five examiners for the manual examination of the neck with 20 patients suffering from neck pain and 20 asymptomatic subjects. The patients and the subjects were assigned randomised and the examiners did not know which group they came from. Joint-facets, spinous processes, cervical- and shoulder muscles were assessed for tension and passive mobility testing for the segments C0/C1 to C7/Th1 was done. Significant interexaminer agreement could only be found for palpation of tension in joint-facets and superficial neck muscles as well as for induced kinesalgia. Passive mobility testing showed no agreement. On the whole there was low to moderate interexaminer agreement ($0.2 < \kappa < 0.6$).

2.2.2 Palpation of body rhythms

Besides the assessment of musculo-sceletal dysfunctions several rhythms of the body can be examined and used as diagnostic tools too. In principle this can be seen as a group of phenomena which occur in repeated cycles, whose frequency, regularity etc. can be judged. Here we find manual cardiac pulse diagnostics or palpation of thoraco-pulmonary respiratory excursions on the ribs, the sternum or the abdomen. As the core-point of my investigation is the examination of a rhythmical event besides cardial and thoraco-pulmonal respiratory rates, the Primary Respiratory Mechanism (PRM), which is hypothesised by the Cranial Concept, It might be useful to have a look at the common problems concerning rhythm-palpation. In principle the situation is similar to motion palpation. Besides tactile perception proprioception according to amplitude and frequency of the rhythmic event plays a major role.

Smith and Craige (1980) examined the palpation of frequencies from 1 to 40 Hz on 10 subjects using a palpation stimulator with regard to accuracy and possible enhancement in palpation of the precordium. The subjects palpated a disc, which was set sinusoid swinging by a generator whose frequency randomly varied between 1 and 40 Hz. With the other hand they could control the amplitude whose range was from several hundredths of a millimetre to approximately 1 mm and increase it until they could feel the movement. As the authors were mainly interested in the different thresholds for the perception of specific frequencies, the strain gauges measuring the amplitude where not calibrated. In a second trial the palpating hand has been restrained by putting pressure on it from above. It could be shown that with the non restrained palpation technique the threshold for perception began to decrease from 5 Hz upwards. For the lower frequencies (< 5 Hz) the amplitude for the free hand had to be 4.7 times higher than for the restrained one.

With regard to specific bimanual palpation techniques, where the contacting hand stays relaxed whilst guidance and pressure is exerted by the second hand which is put on top of the first, these results are very interesting and confirm the usefulness of such techniques. Another application of this phenomenon can be found in practical courses, when the teacher puts her/his hands on top of the student's hands. This technique is frequently used for learning to palpate the subtle phenomena of the Cranial Concept.

Myers et al. (1987) assessed the interexaminer reliability of six vascular surgeons palpating the femoral and popliteal pulses on 22 patients with occlusive peripheral arterial disease. The examiners had to record the absence or presence of the pulse and whether the amplitude was normal or reduced. For the first criteria moderate agreement ($\kappa = 0.53$ for the femoral pulse; $\kappa = 0.52$ for the popliteal pulse) has been found, for the judgement of the amplitudes agreement was found to be no better than expected by chance ($\kappa = 0.15$ for the femoral pulse; $\kappa = 0.01$ for the popliteal pulse). The authors suggest that pulse palpation alone is insufficient for diagnosing occlusive arterial diseases.

Lundin et al. (1999) investigated the reliability of distal pulse palpation (A. dorsalis pedis and A. tibialis posterior) on 25 patients with suspected lower limb arterial disease in comparison with the ankle/brachial index (ABI). The authors infer that the rate of misdiagnosis (30%) and the poor interexaminer agreement is to high in order to use pulse palpation as a single diagnostic method for arterial diseases. A further interesting aspect of this study is, that examiners in a peaceful environment achieved much better agreement ($\kappa = 0.68$) as examiners in a busy outdoor clinic ($\kappa = 0.38$). The diagnostic accuracy of the first group is tolerable.

Overseeing the presented studies it is striking, that even in fields where the physiological substratum which underlies manual assessment is beyond question, interexaminer reliability remains questioned.

2.3 Summary

The following arguments might be reasonable for founding the importance of reliability studies on manual findings for the osteopathic profession:

- Manual techniques represent the main source for osteopathic diagnostic procedures.
- Achieving definite palpatory skills for assessment takes a great part in the education of the osteopath. This is why the check of the grade of correctness and accuracy of the student's ascertained findings is of essential importance for the student.
- As manual ascertained data have a big influence on the therapeutic intervention in osteopathic treatment, the examination of the validity of the used techniques by quantitative and objective means of measurement seems necessary and useful with respect to clinical efficiency and safety.
- Quantitative assessment of the validity of manual diagnostic tools can make the dialogue with colleagues who don't know much about the principles of osteopathy easier.
- The knowledge about the reliability of the repertoire of the used manual techniques, can make explanations to patients easier for the osteopath.

Interexaminer reliability studies on manual assessment-techniques bring up specific problems for osteopaths and hold common risks due to the characteristics and the interpretation of data:

Reproducibility of passive motion palpation (testing) has repeatedly been found to be poor (Mootz et al., 1989; Keating et al., 1990; Schöps et al., 2000). Passive motion testing takes a prominent part in osteopathic assessment procedures. As Osteopathy in common, the Cranial Concept is especially dealing with the assessment of very small movements. Other assessment procedures, judging small differences like the Gillet-test (Meijne et al. 1999) or palpation of the sacroiliac joint (O'Haire and Gibbons 2000) show low interexaminer agreement. Therefore it can be assumed, that the existing problems are multiplied. The gathering of data can be expected to be very sensible against methodological aspects. It is self evident that the results' meaningfulness increases with the number of subjects and examiners.

3 Fundamentals of the Cranial Concept

3.1 Hypotheses of the Cranial Concept – Their origin and development

In order to be able to sufficiently justify the hypothesis on which the thesis and its quantitative examination is based on, it is essential to give a short review on how the Cranial Concept originated and developed. The following analysis is mainly based on two sources: The collected writings of W.G. Sutherland, edited by A.S. Sutherland and A.L. Wales (1998), which were published under the title *Contributions of Thought* (COT) and a reprint of the original version of H.I. Magouns (1997) *Osteopathy in the cranial field* (OCF).

3.1.1 The beginning – Sutherland

The idea of cranial mobility and rhythmics

The Cranial Concept goes back to the American Osteopath W.G. Sutherland (1873–1954). The starting point for his deliberations can be found in the philosophy of A.T. Still (1828–1917), the founder of Osteopathy, to whom he repeatedly refers. An important factor for the emergence of the Cranial Concept is the idea of mobility within the osseous and membranous structures of the scull. This concept of hypothetical cranial mobility originated from one of the fundamental principles of Osteopathy: the interdependency of structure and form in combination with intensive studies of anatomy and physiology. In a recorded talk at the Des Moines Still College of Osteopathy and Surgery in the year 1944 (COT, pp 145–146) Sutherland refers to his time as student at the American School of Osteopathy when he answers a colleagues question „*where he did find the bug to think out this cranial stuff*": The ideas of his teacher A.T. Still and anatomical studies on the disarticulated skull, especially the form of the sutures, „*bevelled like the gills of a fish*" (at other places often mentioned in connection with the Squama ossis temporalis), let him think of something like articular mobility and the function of a *respiratory mechanism*:

> »The thought came, like a bolt from the blue: „Bevelled like the gills of a fish; indicating articular mobility for a respiratory mechanism.« (COT, p 146)

By analysing the anatomical structures of the skull in the context of the form-function-relationship Sutherland developed, as much as one can gather from what is mentioned in the existing written sources, the following two essential hypotheses:

(1) the idea of mobility within the osseous and membranous skull, which is

(2) functionally related with „respiration" in a broader sense.

Rhythmicity which is inherent to respiration and expresses itself in repeated cycles can therefore seen as an essential part of the nature of the Cranial Concept.

The development of the basics of the Cranial Concept

One of the first public documentation of the early deliberations and techniques concerning the Cranial Concept is a paper from the district meeting of the Minnesota State Osteopathic Association from the year 1929 (COT, pp 31ff). Sutherland points out his new ideas at the end of an article about how to treat influenza with involvement of the respiratory system. His initial idea is the presumption that tissue tension, which for him is so typical in influenza, originates from tension in the Falx cerebri and the Tentorium cerebelli:

> »We are inclined to think that the tissue tensity effects so markedly present in influenza have their initiation through the falx cerebri and tentorium cerebelli, with resultant restriction of drainage channels [...]« (COT, p 41).

According to this Sutherland describes a technique for treating these tensions, which he still ascribes an experimental phase and ends his deliberations as follows:

> »A deeper understanding of the apparent functioning of the falx cerebri and the tentorium cerebelli, in coordination with cranial articular movement rhythmical with that of the diaphragm, may bring forth an effective specific treatment of such tissue tensity effects in acute diseases« (COT, p 46).

Yet in this early phase the basic characteristics of the Cranial Concept can be seen emerging:

- A dominant role of the tension of the meninges with accentuation of the Falx cerebri and the Tentorium cerebelli.
- The possibility of articular movement within the connections between the osseous elements of the skull.
- The existence of rhythmical movement, which at this stage is identified with the respiration of the thoraco-pulmonary system.

The meaning of „respiration" as metaphor and driving agent in the Cranial Concept

Following the principle of the form-function-relationship Sutherland infers, that, if there is cranial mobility, there also need be a motor which drives the system. As within the skull there are no muscles who obviously would be able to cause movements within the sutures, syndesmoses and synchondroses of the calvaria and the cranial base, Sutherland begins to search for a mechanism which is inherent to the system and, as he repeatedly emphasises, autonomous. A first hint can be found in the bulletin No. 65 of the *International Society of Sacro-Iliac Technicians* (1940):

> »It was stated that muscular agencies are unnecessary in cranial mobility as its activity is involuntary, occurring as the periods of respiration, [...]« (COT, p 131).

One of the first publications where Sutherland arises this question is an article from the year 1933 which has been printed in the *Western Osteopath* entitled *Cranial Membranous Articular Strains* (COT, pp 69–77). In the course of the article we find:

> »The question naturally arises: What actuates the falx-tentorium balance-reciprocant in its functioning? For the time being, suffice it to say: Some latent pulsatory or rhythmical agency provides the actuation« (COT, p 74).

> »Therefore one can hypothesize that, possibly, the cerebral convolutions and fissures were designed to accommodate pulsatory rhythmical activity of the brain itself, that is, that the brain functions automatically in rhythmic actuation through its various convolutions.«

Further in Sutherland (1939; Reprint 1994; p 51):

> »According to my present hypothesis, interpreted through various phenomena resulting from the application of cranial technic, the brain involuntarily and rhythmically moves within the skull.«

Sutherland obviously expresses his ideas very carefully and by doing so he leaves an unresolved question concerning the explanation of the physiological backgrounds of the Cranial Concept which up to our days could not be adequately answered. About 40 years later Kappler (1979, p 14) puts it as follows:

> »But the exact nature and mechanism of this energy or potency will have to wait for scientific explanation [...]«

In consequence this mechanism has to fulfil the following three conditions: It has to be able to

(1) initiate movements within the mentioned articulations of the cranial system in an autonomous manner and

(2) repeat them rhythmically in recurring cycles. A third condition can inevitably be derived from the fact that Sutherland exclusively used his hands for assess-

ing mobility and rhythmicity on the scull:

(3) the cranial movements initiated by the postulated mechanism have to be manually perceptible (COT, p 119).

In further publications Sutherland repeatedly mentions cyclic phases of motion in relation to respiration, using terms like *expiration* or *inspiration* (COT, pp 61, 73, 75, 76) describing the two phases of cranial movement. If one does not know about the fundamentals of the Cranial Concept, these early explanations do not clearly differentiate between the respiratory movements of the thoraco-pulmonary system and *cranial respiratory movements*, which ought to be independent from the thoraco-pulmonary system and for which he uses the terms *expiration* and *inspiration* in a metaphoric way (see the noted quotation from above: COT, p 131).

The Primary Respiratory Mechanism (PRM)

In an undated letter Sutherland makes a distinct difference between the respiratory movements coming from the thoraco-pulmonary system and the *cranial respiration* which he summarizes under the term *primary respiratory mechanism* (PRM) (COT, p 137). Using the term *primary* he even goes one step further and places the PRM within a functional hierarchy above the thoraco-pulmonary respiration. A further reference can be found in Sutherland (1939; Reprint 1994; p 24):

> »I view the cranial articular structure as a primary respiratory mechanism, and that it functions in conjunction with the brain, the ventricles, and the intracranial membranes; the diaphragmatic respiratory mechanism being secondary thereto.«

Furthermore this original terminology for the cranial rhythm will be used. The term *Cranio Rhythmic Impulse* (CRI) as introduced by Woods and Woods (1961) can be seen as identical with the PRM.

Concerning the explanatory models for the PRM the cerebrospinal fluid (CSF) represents another basic functional element. An early reference can be found in an undated manuscript entitled *The Incitation of Cranial Articular Mobility*:

> »One might say, that the cerebrospinal fluid is a compression lubricant, functioning much in the manner of the lubricant utilized in a hydraulic lift apparatus« (COT, p 98).

An extended explanation is given in the shorthand noted lectures, given at the College of Osteopathic Medicine and Surgery in Des Moines, Iowa in April 1948 under the title *Cerebrospinal Fluid Fluctuation and Central Nervous System Motility* (COT, pp 189–198). The primary position of the PRM can partly be explained by the special role of the CSF as seen in Osteopathy. In this context Sutherland cites a passage from A.T. Stills *Philosophy of Osteopathy* (p 39):

> »A thought strikes him that the cerebrospinal fluid is the highest known element in the human body, [...]« (COT, p 188).

Adding the functional dominance of the CSF within the PRM the Cranial Concept is completed by a further fundamental criteria: the physiological, clinical and therapeutical relevance of the PRM.

This syncretistic mixture of physiology and metaphysics (nearly physico-theological argumentation), which is expressed in the short citation from A.T. Still above, reflects a characteristic feature of certain osteopathic approaches in the beginning and even nowadays. It can be regarded as one cause for criticism against osteopathic models. Kappler (1979) on the one hand mentions that there is more than mere mechanics of movement behind the PRM and refers to the fuzzy term „energy", which is used a lot in alternative or border sciences for inexplicable phenomena. On the other hand this author insists in the necessity of assessing the hypotheses of the Cranial Concept under scientific criteria.

The cranio-sacral unity

A further basic element in the Cranial Concept besides the explanatory-models concerning the origin of cranial movement is the functional unity of cranium and pelvis. A first summarising publication in the bulletin No. 45 of the *International Society of Sacro-Iliac Technicians* comes from a lecture held at a meeting of the International Society of Sacro-Iliac Technicians in June 1940 under the title *The Core-link Between the Cranial Bowl and the Pelvic Bowl*:

> »The intraspinal membranes that surround the spinal cord act as a reciprocal tension tissue that links and regulates the cranial articular mechanism with the pelvic articular mechanism during periods of respiration, [...]« (COT, p 128).

Synthesis

In a talk at the *Des Moines Still College of Osteopathy and Surgery* in 1944 Sutherland introduces the basic elements of his hypothesis of the PRM and the Cranial Concept in an explicit and compressed manner:

> »In my hypothesis, I have described what we call the primary respiratory mechanism. This mechanism includes the brain, the cerebrospinal fluid, the intracranial membranes and the articular mobility of the cranial bones: also the spinal cord, the intraspinal membranes, the same cerebrospinal fluid and the involuntary mobility of the sacrum between the ilia« (COT, p 147).

The following list summarises the basic elements of the PRM representing the anatomical and physiological core of Sutherland's concept:

- The autonomous motility of the central nervous system
- The cerebrospinal fluid as a dynamising factor for cranial mobility
- The intracranial and intraspinal membranes functioning as reciprocal tension membranes

- The mobility within the osseous elements of the skull
- The involuntary movement of the sacrum, caused by the intraspinal membranes

Finally a more elaborated analytical listing of Sutherland's hypotheses and their therapeutical consequences in chronological and/or functional order as far as possible. These items are relevant for the hypothesis of the thesis:

- The idea of mobility within the area of the osseous and membranous skull.
- The connection between cranial mobility and „respiration" in the broadest sense.
- Cranial mobility is palpable by a skilled person (osteopath).
- The dominant position of the meninges, acting as reciprocal tension membranes in the cranial system.
- The tension of the intracranial membranes is externally (from extracranial) manual detectable by a skilled person (osteopath).
- The cranial movements are initiated by an autonomous mechanism and they go by in repeated cycles (phases).
- The autonomous, cyclic movements of the cranial system, called PRM, are externally manual detectable by a skilled person (osteopath).
- The CSF represents a further basic element of the Cranial Concept in a mechanical, physiological and metaphysical sense.
- The PRM is functionally of primary importance and therefore has therapeutical and clinical relevance.
- The intraspinal membranes act as transducers for the PRM from the cranium to the pelvis.

3.1.2 Magoun – Systematesising the Cranial Concept

H.I. Magouns book *Osteopathy in the Cranial Field* (OCF), first published in 1951, represents a first comprehensive summary of Sutherland's ideas in a systematic and didactical manner. Due to its clear structure Magoun's book serves as a basis for teachings in Cranial Osteopathy to this day.

Magoun gives an overview to the basic elements of the Cranial Concept in the following five points (see OCF, pp 16–18):

(1) The *cerebrospinal fluid* in relation to its (a) exceptional functional position, (b) fluctuation and (c) physiological characteristics and functions.

(2) The *meninges* or *„reciprocal tension membranes"*, which influence cranial mobility and act as transmitters for movement and tension.

(3) The *central nervous system* and its immanent motility.

(4) The *mobility within the bones of the cranium*, which is caused by the following factors: (a) the fluctuation of the CSF, (b) the action of the meninges, especially the falx cerebri and the tentorium cerebelli and (c) the expansions and contractions of the brain, which initiate movement to the surrounding structures.

(5) The *involuntary mobility of the sacrum* between the ilia, which is not influenced by posture.

Referring to the characteristics of the movements caused by the PRM, Magoun defines:

> »The movement of the primary respiratory mechanism is constant and rhythmical in nature and occurs in a cycle [...] All [cranial bones] fall into the classification of flexion or external rotation for the inspiratory phase of the cycla and extension or internal rotation for the expiratory phase« (OCF, p 19).

Magoun clearly speaks about a rhythmic and cyclic characteristic and a certain constancy of the PRM. He works out the nomenclature for the movements of the osseous elements of the cranium during the phases of the PRM in the sense of Flexion/Extension and Internal and External Rotation.

Concerning the relation between thoraco-pulmonary respiration and the PRM Magoun specifies:

> »The primary respiratory mechanism is superior to costal respiration, [...] Active articular motion of the skull, as part of the mechanism, does not necessarily coincide with costal respiration. It can, however, be made to coincide. [...«] (OCF, p 21).

A distinct differentiation between cranial rhythmical mobility and thoracic breathing action is being made while he does not exclude the possibility of simultaneity between the two rhythms. Further Magoun states that the elements of the PRM under physiological and quiet conditions do act synchronous.

> »The fluctuation of the cerebrospinal fluid and cranial articular motion coincide under normal resting conditions. The fluctuation of the fluid, changes in the morphology of the central nervous system and motion of the craniosacral mechanism are synchronous at such times, also« (OCF, p 21).

Concerning the „motor" behind the PRM, Magoun seems to support the hypothesis of an inherent motility of the central nervous system:

> »The central nervous system or the brain and spinal cord with the inherent, jellyfish-like motility which can be seen at operation and is recognized by the scientific world« (OCF, p 17).

3.2 Models for the PRM

The hypothesis of cranial mobility and the uncertainties concerning the agent behind the autonomous activity of the craniosacral system have brought about new hypotheses and various projects in order to base the theories on scientific research or find conclusive physiological explanations. While Sutherland and Magoun thought, the cause for the PRM would be a certain motility of the tissue of the central nervous system, other explanations have been worked out.

3.2.1 The „pressurestat model"

Upledger (1994, pp 22ff) developed an alternative theory to Sutherland's hypothesis. As the cerebrospinal fluid (CSF) is understood as a mechanical driving force in the Cranial Concept, he picks up the idea of a biomechanical model: a closed hydraulic system. Within this system the hydrostatic pressure underlies rhythmical changes. He further thinks that the strength of nervous tissue is not sufficient to act as a pumping mechanism who causes the changes in pressure. This is why he suggests a pressurestat model, which is based on the hypothesis, that temporarily more CSF is going to be produced as can be absorbed which results in changes of intracranial volume and pressure. Upledger works out two theories concerning the control-mechanism behind the cyclic production and absorption of CSF. One possibility might be a neuro-vegetative reflex combining proprioceptors within the sutures (preferred the sagittal suture) and the plexus choroideus where the main production of CSF is going on. Thus a stretch of the structures within the sutures in consequence leads to an increase of intracranial volume which would cause a cut down of CSF-production. Chandler et al. (1979) assessed the relation between absorption and production of CSF and changes of CSF pressure on the living and dead dog. They could observe a hysteresis-phenomenon in 78% of the living animals, but not in dead animals. The authors suggest the possibility of a pressure-sensitive control-mechanism for CFS absorption. A further explanation could, so Upledger, be given in a valve like mechanism within the straight sinus, which controls the drainage of the jugular vein. Concerning the neuro-vegetative reflex theory Upledger refers to the histological studies of E.W. Retzlaff et al. (1982).

Despite experimental results that seem to support Upledger's model, the basic approach of the theory is inconsistent with the generally accepted Monroe-Kellie-Hypothesis which says that changes of intracranial pressure are mainly regulated by the exchange of fluids between three compartments: brain, CSF and blood. Aschoff (2000) states:

> »After the fusion of the fontanels within the first months after birth the neurocranium represents a nearly completely closed space, whose vol-

> ume (ca. 1500–1800 ml) is occupied by three components: brain ca. 88%, CSF ca. 9% and blood ca. 3–5%. Each increase of one of these compartments can be compensated by the decrease of the other two (Monroe-Kellie-Doctrine, formulated by Burrows 1848).« [translation by myself, P.S.]

Aschoff further explains that in case of pathological changes of intracranial pressure the blood-compartment is the fastest changeable compartement. This mirrors in rapid changes of intracranial pressure following vasodilatation in the central nervous system or hypercapnia for instance. Especially the CSF compartment (ventricles, cisterns, furrows of the brain) can in the consequence act as a reserve space. That, even under physiological conditions, sutures might facilitate a certain compliance to the neurocranium, is improbable but has repeatedly been suggested (Adams et al. 1992; see p 32). A possible elastic function of the sutures within pathological intracranial pressure-increases is still under discussion.

3.2.2 The „muscle reaction model"

A further theory for explaining the autonomous cranial movements is suggested by F. Becker (in: Upledger 1994). Becker presumes, that cranial mobility could be a consequence of tonic reactions of postural muscles to gravity. Muscle-caused changes in tension would then cause changes in pressure within the CFS compartment. With respect to physiologically proven influences on intracranial pressure this theory seems most improbable.

3.2.3 The „tissue pressure model"

Norton (1991) developed a mathematical model. He starts from the suggestion, that the phenomenon the observer perceives as the PRM is a resultant of an overlap of heart- and respiratory rates of the observer and the subject which takes place at the contact surface between the observer's hands and the subject's skin. The tissue pressure model bases on the following three assumptions:

- Slowly adapting mechano-receptors within the palm of the hand (type SA) can react on skin displacements within a frequency range which is very similar to the frequencies that have been assigned to the PRM. Most likely the Merkel cells would correspond to this task.
- The stimulation of mechano-receptors under certain condition creates the impression of movement. This happens under the influence of pressure- and volume changes of fluids and their compartments within contacting tissues of examiner and subject (e.g. skull and palm). Norton starts from the Starling-Hypothesis which defines the flow rate of the ultrafiltration exchange in the capillary system:

$$\textit{Fluid movement} = k\,[(P_c + \pi_i) - (P_i + \pi_p)] \tag{1}$$

if:

k = Filtration constant for the capillary membrane
P_c = Capillary hydrostatic pressure
P_i = Interstitial fluid hydrostatic pressure
π_p = Plasma protein oncotic pressure
π_i = Interstitial fluid oncotic pressure
(cp. R.M. Berne and M.N. Levy, 1993, p 473)

He suggests, that changes of P_c, the parameter for the mean hydrostatic pressure in the capillaries, could most likely correspond to the PRM and creates the following equation, where P_c is understood as a resultant of arterial and venous pressure:

$$P_C = \frac{P_a\left(\frac{R_v}{R_a}\right) + P_v}{1 + \left(\frac{R_v}{R_a}\right)} \tag{2}$$

if:

P_c = mean microcirculatory pressure
P_a = pressure in the large arteries
P_v = pressure in the large veins according to Norton (1991)

- The PRM might mirror a complex function of examiner's and subject's respiratory and cardiovascular rhythms.

Standard values for examiner's and subject's P_a and P_v were inserted into the simulated computer model, regarding physiological conditions, heart rate- and blood pressure variabilities were coupled to respiratory frequencies whose changes were randomised. Finally a net pressure (P_{net}) which stimulates the receptors of the examiner's hand has been computed from the difference between the examiner's and the subject's P_cs. In addition the curve of the P_{net} has been smoothened by an algorithm, which simulated the characteristics of the slow adapting SA II mechano-receptors. Norton proves high correlation between the PRM-simulations computed by the tissue pressure model and direct measurements of the PRM that were published by Frymann (1971) or reported by other authors. The model provides a consistent physiological explanation for what different examiners seem to perceive assessing the phenomenon described as the PRM but it cannot reveal any clinical relevance.

3.2.4 The „Entrainment-Hypothesis"

McPartland and Mein (1997) discuss another explanatory theory for the PRM referring to a principle which is called „Entrainment" by the authors: Entrainment means a kind of harmonic integration of so called *biological oscillators*. Biological oscillators, so the authors, is a term subsuming all rhythmical phenomena occurring in biological systems, like cardiovascular- and respiratory rates, cyclic variabilities of heart rate and blood pressure (described as *Traube-Hering-Mayer waves*), system inherent oscillators of striated and smooth muscle fibres, cortical metabolism, endocrine glands, glia cells etc. All this be the expression of an interaction between ortho- and parasympathetic influences on the organism. Measurements of the Entrainment frequency (McCraty et al. 1995, Tiller et al. 1996) ought to produce values about 0.125 Hz which corresponds to about 7.5 cycles per minute.

The similarity to the frequency range which is reported by the Cranial Concept and the basic principle of rhythmicity suggest the conclusion that the PRM could be explained by Entrainment. Finally the authors state, that due to the complexity of the postulated phenomena an exact experimental proof of Entrainment is difficult. In consequence this is also true for the PRM. Thus two hypotheses which are both difficult to proof want to support each other.

3.3 Experimental research and discussions on cranial mobility

Concerning scientific research to support the Cranial Concept, two main approaches can be observed: projects who try to proof the existence of cranial mobility and projects who want to proof the existence of the PRM as rhythmic movements which are not caused by thoraco-pulmonary respiration and the cardiovascular pulse. In addition histological studies on the sutures have been done. In the consequence we want to give a short review over some of the frequently cited articles.

The orthodontist Baker (1971) executed a static measurement of changes of the maxilla on one patient during a treatment- and control period of six months. Osteopathic and orthodontic treatment where done parallel. Comparing 20 alginate impressions of the maxilla he could measure a change of the lateral width of the maxillary arch up to 0.0276 inch (0.01 cm).

Viola Frymann (1971) did one of the first major experimental research project that intended to proof the PRM. Until these days this study is frequently cited. The dynamic measurements were done over a period of some years. Frymann wanted to assess three main questions:

(1) Does a motion like the PRM really occur?

(2) Can it be mechanically recorded?

(3) If it exists, what is its relations to known physiological functions?

For recording cranial mobility Pick-offs (displacement transducers) were used that were mounted externally on the subject's head. The measuring device was able to record minimal movements. A movement range between 0.0005 – 0.001 inch (0.0012 –

0.0025 cm) has been recorded. Exact technical details of the apparatus' accuracy are not published). In order to be able to measure rhythmic motion which is independent from thoraco-pulmonary respiration and cardiac pulse the transducers had to be mounted with high pressure and in order to rule out respiratory movements, the subjects had to hold their breath. Norton (1991) states that a high contact pressure does not correspond to the real situation, where the PRM is palpated by utmost minimal pressure. In a first sequence measures were only done by the displacement transducers, in further sequences a pneumograph and a pletysmograph had been added simultaneously.

The author suggests that there exists a movement that is inherent to the cranium and that it can be recorded mechanically. Further on she refers to similarities to other physiological known rhythms. We are expressing our respect for the pioneer character of this research, but without any doubt there are methodological deficiencies, which don't allow such a clear conclusion, like the author did. There are no details given about the accuracy of the measurements, how many measures have been done on how many subjects and the manner how the data have been analysed. Therefore it is not possible to interpret the correlation between the recorded data or the influence of artefacts, inaccuracies etc.

The author talks about a low correlation between the data of the pneumo- and pletysmograph and the displacement transducers and interprets this with the existence of an independent rhythm. Norton (1991) computed Frymann's raw data of respiration and pulse by the tissue pressure model. The curves the computer model calculated, significantly corresponded to the curves Frymann recorded via the transducers on the cranium. In addition Norton states that in Frymann's recordings nearly every peak of the respiratory curve is near a peak of the PRM curve (Norton 1991, p 988). According to such a result there would be a high correlation between respiration and the PRM.

Michael and Retzlaff (1975) as well as Retzlaff et al. (1975) did measurements on anaesthetised squirrel monkeys. The authors assessed mobility and rhythmicity of the parietal bones in comparison with cardiac and respiratory rates. In order to measure the movements directly on the bone, screws were drilled into the parietal bone and connected with displacement transducers. The authors observe that when the animals head is free the movements of the parietal bone correspond to cardiac and respiratory motion. If the head is firmly fixed by a stereotactic frame, ruling out disturbing influences and artefacts, there can be recorded a frequency with 5–7 cycles per minute. Further on possible correlations with changes of CSF pressure are brought up to discussion by observing simultaneous changes of rhythmicity in the parietals during extension and flexion of the head. Concluding the authors ask for possible aetiological facts behind the recorded data.

There are no details given about the number of measurements that were done on how many animals nor about the accuracy of the measuring device and possible erroneous data. Retzlaff (1987) says in a later interpretation of the results that they would support two hypotheses: Sutherland's hypothesis of cranial mobility and the presumption that changes of CSF pressure initiate movements in cranial bones. In this context it might be interesting to state that these authors as well as Frymann have to use a definite fixation of the head in order to be able to record frequencies that do not correspond to cardiac and respiratory rates. This contradicts to the palpation techniques used in the Cranial Concept that work with minimal pressure.

In order to support the theory of cranial mobility Retzlaff et al. (1978) do additional histological research on the sutures of the squirrel monkey. As the preparation of the specimen is known to be difficult, the authors develop a special method where the structures within the suture are not changed or destroyed (Retzlaff, 1987). In 10 specimen no ossification at all can be found within the sutural zone. Further the authors report that ligamentous structures (Sharpey fibres), reticular and elastic connective tissue, vascular structures and nerve fibre endings can be differentiated within the suture. This whole complex, called the sutural ligament, represents the contact zone between the bones of the cranium (Retzlaff 1987). It exists of periost and a richly vascularised matrix of connective tissue. The orientation of the fibres seems to indicate mechanical functioning, differs from suture to suture and would correspond a syndesmosis.

Ferré and Barbin (1991) contradict the hypothesis of cranial mobility from the viewpoint of the anatomist and bring the usefulness of the Cranial Concept up to discussion. So the authors ask to consider that on the one hand the synchondroses of the cranial base are fused by ossification. So at the end of the age of eight there is no more possibility for movement within the anterior and middle cranial fossae. A theoretical possibility for movement rests only in the posterior cranial fossa and is untenable at the end of adolescence. On the other hand the sutures of the calvaria ossify at the inner table parallel to the disappearance of the growth cartilages of the long bones. A certain but very limited activity the authors grant the lateral sutures of the calvaria in respect to constraints occurring during mastication, emphasising that in 100 sections nearly all samples, contrary to the Cranial Concept, showed a „*compact and resistant region*" (p 168).

The authors further state that the biomechanical unity of the base of the skull can be divided into three main functions: (a) the anterior cranial fossa has to cope with constraints coming from mastication, (b) the posterior cranial fossa has to adapt to the forces of the cervical and nuchal muscles and (c) the middle cranial fossa acts as a transfer-zone between posterior and anterior constraints. The trabecular structure of the bones of the cranial base mirror these functional aspects. Summarising the authors state that the cranial base „*relies on mechanical aeronautical solutions combining strength with lightness*" (p 169). Any displacement within the osseous structures of this system, so the authors, stands in total contradiction to its biomechanical properties.

Referring to the PRM, Ferré and Barbin mean that it is obviously inadmissible that the central nervous system should execute any kind of autonomous peristaltic-like movements as it hasn't any contractile elements like the viscera. Due to known intracranial security systems like the Cisterns and the Granulationes subarachnoidales which compensate changes in CSF-pressure or -volume, the authors do not believe in an interaction between CSF-flow and sutural movements. In addition physiological CSF pressure is too small (0.4 N/m^2) in order to have any mobilising effect on the sutures. Finally Ferré and Barbin conclude that known physiological rhythms like cardiac and respiratory rates do have a proven influence on intracranial pressure.

This critical contribution shows that hypotheses like cranial mobility and autonomous rhythmicity that underlie the Cranial Concept cannot be looked at as scientifically proven or experimentally solved problems.

Adams et al. (1992) assessed the mobility of the parietal bones in relation to the sagittal suture on anaesthetised house cats. The movements were analysed in the frontal and transverse plane. For the measurements the animals' heads were fixed in a stereotactic frame after a measuring device had been screwed to the exposed parietal bones. In order to measure minimal displacements the authors chose an isotonic measuring device whose elements do not limit or hinder small displacements and state that in former studies isometric devices with a low compliance were used.

Blood pressure and cardiac rates were directly detected from the femoral artery, respiratory rates and depth by a pneumotachograph and intracranial pressure changes by a needle inserted into the lateral ventricle. The following measurements have been executed: (1) without any additional external or internal forces, (2) during extracranial pressure change by manual compression of the head, (3) during intracranial pressure change by: (a) hypercapnia, (b) Norepinephrine-injection and (c) injection of artificial CSF into the lateral ventricle. In case (1) the authors could measure lateral and rotatory movements of the parietal bones, that correspond to cardiac rates, changes in blood pressure or respiratory rates. The addition of manual compression causes displacements of the parietal bones and temporarily changes cardiac and respiratory rates, the depth of respiration and intracranial pressure. Any changes of intracranial pressure also cause changes in movement of the parietals. No changes of cranial mobility could be observed when the animal's head has been firmly fixed in the stereotactic frame during the injection of artificial CSF. The authors report a lateral range of movement for the sagittal suture with ca. 130 mm and suggest that sutural mobility could play an important role in cranial compliance besides the Monroe-Kellie-Hypothesis.

The material presented in this article is difficult to interpret as the authors don't give any detailed information on how many measures on how many animals were done. So there are no resulting mean values or standard deviations. In addition the authors state that they present mainly the data that support their hypotheses. So the data analyses cannot be seen as objective and scientific. This study does not deal with the PRM at all. But it is interesting that the authors could not find any indication for an autonomous cyclic event beyond the known physiological rhythms.

Further experimental Studies were done by Tettambel et al. (1978), Mitchell (1979) and Rommeveaux (1992). As the methods of data collection and analysis are not clearly described in the articles, any interpretation is difficult.

3.4 Possible physiological explanations of the PRM

Already Frymann (1971) assumed that the PRM could be identical with known physiological phenomena. She suggests rhythmical changes in neuronal activity of the vasomotor centre as described by T.C. Ruch and J.F. Fulton or cyclic changes in blood pressure heart rate as observed and described by Mayer around the end of the 19th century. She further cites the *Central respiratory drive potentials* (CRDP's) as described by Sears and the cyclic movements of the CFS as observed by Moskalenko.

Despite the different intentions Anatomists, Physiologists, Neurophysiologists, Neurosurgeons, Cardiologists etc. do assess changes of intracranial pressure, CSF dynamics, brain pulsations, cyclic changes in heart rate and blood pressure or similar issues. Their results can shed light on the hypotheses of the Cranial Concept. Kappler (1979) as well as Patterson (1990) mention that in future the integration of research results from outside osteopathy will be of major importance for the explanation of osteopathic theories and models. Publications concerning the issues mentioned above are numerous. We picked some in order to show the relevance for the explanation of the PRM.

Early publications on the phenomenon of CSF pulsation and resulting rhythmic movements of the brain under the influence of cardiac pressure changes were done by O'Connell (1943) and Bering (1962). Quantitative measures of rhythmic pulsations of the brain stem of ten anaesthetised house cats by a fibre-optic displacement transducer were done by Britt and Rossi (1982). They could observe two modes of displacement: (a) A movement with relatively low amplitude (110–270 mm) and short duration (330–400 ms) which corresponds to each cardiac systole (A-wave component). (b) A slower movement with an amplitude of 300 – 950 mm and a duration of 2.4 – 5.1 sec, which corresponds to each respiratory cycle (P-wave component).

Feinberg and Mark (1987) assessed brain motion and CSF circulation on 25 healthy subjects and 5 patients using MR Velocity Imaging. So the authors could show and quantify brain motion in vivo. Brain motion has been interpreted a consequence of kinetic energy of the pulse wave by momentum transfer (mainly by the CSF), changes of intracranial blood volume and expansion of the arteries.

Daley et al. (1982) assessed changes of intracranial pressure under the influence of the cardiac action on six adult and one paediatric patient under intensive care. The following parameters of cyclic changes of intracranial pressure have been measured: (a) the mean value of sampled pressure, (b) the amplitude of the fluctuation (deviation of the sampled pressure from mean) and (c) the latency interval the occurrence of the R wave of the ECG and the subsequent peak of the intracranial pressure. The authors suggest that their results support the hypothesis that changes of intracranial blood volume in connection with cardiac rates can cause fluctuations of intracranial pressure. As the authors found low correlation between the hourly recorded cardiac pulse and the latency interval they discuss changes of venous drainage besides arterial influences on brain motion.

The cited publications indicate that rhythmic brain motion as well as CSF dynamics do exist, influenced by heart action and respiration. Sutherland's hypothesis, that cranial mobility (in the sense of articular motion within the sutures and synchondroses of the cranium) is caused by brain motion and CSF-fluctuation is left open. A further divergence with the Cranial Concept and the cited data is the phenomenon of the PRM, whose frequency ought to be lower than and independent from cardiac and respiratory rates. In this context investigations of cyclic changes of blood pressure (Mayer waves) and heart rate (*heart rate variability* [HRV]) can be interesting, because they show a frequency range which is similar (< 0.1 Hz) to the frequencies reported for the PRM by the Cranial Concept.

Changes of blood pressure have already been observed 130 years ago. They are known as low frequent (LF) *Mayer waves* at a frequency of about 0.1 Hz and as high frequent (HF) *Traube-Hering waves*, coupled with respiration, at a frequency of about 0.4 Hz (Karemaker, 1999). Exact long term frequency domain analysis of HRV has shown different physiological frequency ranges. Camm et al. (1996) note four spectra:

As control centres for HRV central (vasomotor and respiratory centres) and peripheral (oscillations of arterial pressure and respiration) mechanisms are discussed (Camm et al. 1996). While vagal activity as a main contribution to the HF-spectrum is quite proven, disagreement exists for the LF-spectrum between explanations via a sympathovagal balance or mere sympathetic modulations. Physiological explanations for the VLF- and ULF-components warrant further elucidation (Camm et al. 1996). A control of the

LF-component by oscillations of blood pressure is increasingly discussed (Karemaker 1999). Regulating factors are more likely to be seen in a central nervous pace maker system than in baroreflexes (Cooke et al. 1999).

Spectrum	Abbreviation	Frequency	Cycles x min^{-1}
Ultra low frequency	ULF	≤ 0.003 Hz	≤ 0.2
Very low frequency	VLF	0.003 – 0.04 Hz	0.2 – 2.5
Low frequency	LF	0.04 – 0.15 Hz	2.5 – 9
High frequency	HF	0.15 – 0.4 Hz	9 – 24

Table 2
Frequency spectra from frequency domain long term analyses of HRV. The LF-spectrum corresponds to Mayer waves, the HF-spectrum corresponds to Traube-Hering waves (according to Camm et al., 1996).

As these publication show that physiological frequency ranges being similar to the frequencies reported for the PRM have been observed and proven. But they are not resulting from a system that is independent from the cardiac and respiratory system. They seem to originate from oscillations of the autonomous nervous system in order to control heart rate, respiratory rate and blood pressure. Recent experiments could show that these variabilities in heart rate and blood pressure can be understood as a physiological substratum of the PRM – like McPartland and Mein (1997) suggested with the *Entrainment-hypothesis* (see point 3.2.4, p 45ff) – and seem to represent the phenomenon underlying PRM-palpation (Nelson et al., 2001; Sergueef et al., 2002, Nelson et al., 2006). If this has any clinical relevance, as indicated by the Cranial Concept, remains unclear.

3.5 Summary of research on the hypotheses of the Cranial Concept

Static (Baker 1971) and dynamic measures (Frymann, 1966; Retzlaff, 1975; Tettambel, 1978; Rommeveaux, 1992; Adams et al. 1992) of the possibility of cranial mobility and rhythmicity have repeatedly been executed. The interpretation of the results of the discussed studies in the sense of proven existence of the PRM is impossible because of methodological deficiencies (Ferre and Barbin 1990, Rogers and Witt 1997, Green et al. 1999). Explanations of the PRM via physiological proven rhythms (e.g. HRV) or the *Tissue pressure model* (Norton 1991) would be possible. However, they have neither been used nor been seriously discussed by the osteopathic profession up to now. So the hypotheses of cranial mobility and the existence of the PRM cannot be regarded as scientifically proven. In the case of this research this means that the existence of the reference of the main outcome variables is unclear.

3.6 Palpation of the PRM

3.6.1 Specific nomenclature and palpation of the cycles of the PRM

Following thoraco-pulmonary respiration, in the beginning Sutherland (Sutherland, 1998) used the terms *Inspiration (Inhalation)* and *Expiration (Exhalation)* for the cycles of the PRM. Later these terms have been completed by the terms *Flexion* and *Extension* as well as *Internal Rotation* and *External Rotation*, when the movements of the cranial bones during the Cycles of the PRM are described (Magoun 1997). Like in thoracic respiration Inspiration or Flexion or External Rotation goes along with a more active widening of the whole system, while Expiration or Extension or Internal Rotation means a more passive component in the sense of contraction of the whole system.

Elements	Inspiration phase	Expiration phase
median bones	Flexion	Extension
bilateral bones	External rotation	Internal rotation
Extremities	External rotation	Internal rotation
Antero-posteriorer cranial diameter	decreases	increases
Transverse cranial diameter	increases	decreases
Calvaria	comes down	lifts up
Sacrum	moves cranial	moves caudal

Table 3
Local and global changes, induced by the PRM during its Inspiration- and Expiration phase.

The terms Flexion/Extension are used for the median positioned bones of the cranial system which rotate around transverse axes. These are: the Occipital bone, the Sphenoid, the Ethmoid, the Vomer, the Mandible, the Hyoid and the Sacrum. The terms Internal/External Rotation refer to the bilateral bones like: the Parietals, the Temporal bones, the Zygomatics, the Maxillae, the Palatines, the Nasal bones and the Lacrimal bones (Magoun 1997, pp 19–20).

3.6.2 A model for cranial movement

A *cogwheel model* is frequently used in the Cranial Concept to describe the movements of the median elements of the craniosacral system schematically (Sutherland, 1990, p 43; Magoun, 1997, p 18; Liem, 1998, p 236). The model is shown in Fig. 1. The interaction of the elements of the system and the transfer function of the fascial system are de-

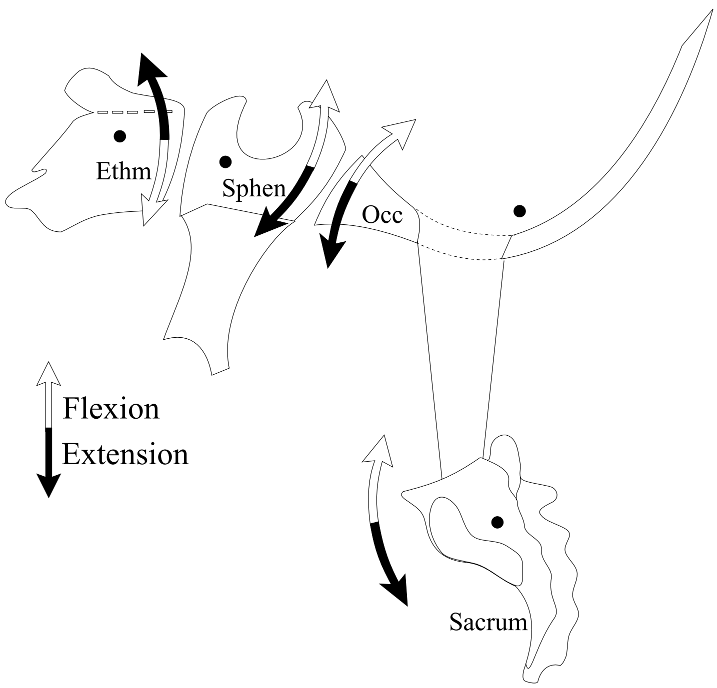


Figure 1
Schematic view of the *cogwheel model*, as used in the Cranial Concept. It shows the main osseous elements of the craniosacral system positioned in the middle and their movement-tendencies during flexion and extension (according to Sutherland, 1990).

Ethm .... Ethmoid bone
Sphen .. Sphenoid bone
Occ ...... Occipital bone

scribed to cause global rhythmical changes of the entire body during the phases of the PRM. During the inspiration-phase a widening of the transverse diameter of the cranium is been described, whilst the antero-posterior diameter decreases, the sacrum moves mainly cranial and flattens a little bit while the extremities show an external rotational tendency. During the Expiration phase the opposite movements occur (Magoun, 1997, pp 33ff; Liem 1998, pp 220–221). For that see also Table 3. Following the Cranial Concept, one can learn to palpate these changes everywhere on the body (Becker 1997, p 41). There are special regions where the PRM can be palpated better. These are frequently used for palpatory training and basic assessment. These are the parietals, the squamous part of the occipital bone, the greater wings of the sphenoid bone and the sacrum (Sutherland, 1998, pp 161–162; Upledger, 1994, pp 40–50; Liem, 1998, pp 289–293).

3.6.3 The quality of cranial palpation

Sutherland himself repeatedly emphasises the importance of tactile sense for the osteopath (Sutherland 1994, p 67; 1990, p 151ff, 170; 1998, p 16). He uses the picture of *knowing* and *thinking* fingers:

> »While your fingers are there feeling, seeing, thinking and knowing, [...]« (Sutherland 1990, p 151).

Palpation as described in the Cranial Concept requires minimum pressure, mental concentration as well as physical relaxation of the examiner (Sutherland 1990, p 151; Magoun 1997, p 55; Upledger, 1994, pp 37ff). This should bring together the osteopath's capacity of matching the tissues' tension as well as to withdraw himself and be just passive observer (Upledger, 1994, p 30; Liem 1998, pp 276–277).

4 Studies on craniosacral palpation and its reliability

4.1 A selection of studies

Upledger (1977) did a study on the reproducibility of craniosacral findings. 25 children were examined by four skilled examiners, whilst only one examiners assessed all children. Following a protocol scheme of frequent craniosacral dysfunctions, 19 parameters had to be judged by a three point ordinal scale due to the severity of the dysfunctions. Before that, pulse and respiratory rate of the child and the examiner have been measured and the examiners palpated the frequency of the child's PRM. Reliability coefficient and interexaminer agreement have statistically been analysed for the 19 parameters. The author reports a percent agreement between 65–92 %. These results seem to indicate good to nearly perfect interexaminer agreement.

As percent agreement is an invalid method for analysing agreement (see point 2.1.3, p 23ff), these results are misleading. Wirth-Pattullo and Hayes (1994) analysed Upledger's raw data with adequate methods (analysis of variance for repeated measures) and calculated an Intraclass Correlation Coefficient (ICC) of 0.57. This result indicates only moderate agreement.

The palpated PRM frequencies are reported in a range between 8 – 16 cycles per minute (0.13 – 0.26 Hz) and are, so the author, not related to heart- and respiratory rates of subjects and examiners. Upledger does not rule out the possibility, that the PRM frequency might be a modulation of other frequencies. These data are not based on statistical analysis and therefore are not merited. Wirth-Pattullo and Hayes (1994) did a regression analysis on Upledger's data and calculated $r = 0.007 – 0.164$ which means low correlation. Hanten et al. (1998) found similar results in their reanalysis. The reliability of measured PRM frequencies is not analysed in this study.

Upledger and Karni (1979) tried to show time-related correlations between manual findings of one examiner (Upledger) and different electromechanical recordings during a craniosacral treatment. Details about number and characteristics of the subjects are missing. The notes of the examiner are reported in eight categories, which are in the authors' opinion commonly used in the Cranial Concept: Normal rhythm, Stillpoint, End of Stillpoint, Release, Shifting, Pulsation, Wobbling, Torsion. The electromechanical recordings included four channels: strain gauges on the thorax for the recording of respiratory movements, one ECG and two iEMGs on the thighs. Analysing the mechanical recordings and the notes of the simultaneous executed palpation, the authors suggest to have found significant time-related relations.

The authors give no explanation for the choice of the specific measurement parameters. No details are given on how many subjects under which criteria were measured, how many measures were done on the whole and where possible sources of error might be. Further there are no reports about how the data have been analysed. Regarding the methodological deficiencies, the reported results are misleading.

Norton et al. (1992) assessed the manual examination of the PRM of 24 healthy subjects, examined by 12 examiners in order to support the *tissue pressure model*, suggested by one of the authors (Norton 1991; see p 43ff), by experimental data from clinical practice. The examiners were experienced osteopaths, specialised in cranial techniques. In order to record the palpated PRM phases the examiners had to activate a knee switch. In addition the amplitude of the PRM had to be judged using a five point scale. Besides the frequencies of the PRM the authors also analysed the lengths of the flexion- and extension phase (cycle length).

The authors did a simple and multiple linear and curve linear regression analysis as well a an analysis of variance. They found a mean frequency of 3.7 cycles/min (0.06 Hz) for the PRM which is lower than the frequencies reported in earlier studies by Frymann (1971) and Upledger (1977) who described a frequency range between 6 – 12 cycles / min (0.1 – 0.2Hz). The mean cycle length was found to be 16.5 sec with a standard deviation of 2.8 sec whilst the duration of the extension phase was longer than the flexion phase. The amplitudes judged by the 5 points scale ranged from 2 to 4 points and therefore produced a mean value of 3 points. No linear relation was found between the reported amplitudes and the palpated cycle lengths. In order to be able to analyse a possible relation between the interexaminer agreement and the experience of the examiner, one subject was assessed by all examiners within three hours. The examiners were divided into three groups of less than 5 years, 5 to 10 years and more than ten years of experience. According to the palpated frequencies and cycle lengths no significant difference between the groups were found. The mean of the standard deviation of the measured variables was found to be least within the most experienced group.

The analysis of cycle length in addition to frequency domain analysis of manual detected PRM phases constitutes an efficient means for a more adequate data analysis. The reported amplitudes, as also stated by the author, lack significance, maybe because of a lack of variance between the subjects and therefore don't produce meaningful data. The interexaminer analysis has limited meaningfulness as the data were detected from one subject only.

Wirth-Pattullo and Hayes (1994) assessed the reliability of craniosacral rhythm palpation and the relation of the PRM with subjects' and examiners' cardiac and respiratory rates. 12 adult subjects with different case histories were assessed by three examiners. The examiners were physiotherapists and had visited two to seven courses in craniosacral therapy. They used cranial techniques since four years. The PRM was palpated at the head for one minute after an orientation period of two minutes by one examiner after the other. The examiners verbalised when they felt the extension or flexion phase. This was recorded by an assistant who counted the repetitions of the so reported flexion phase during one minute. The subject's and examiner's respiratory- and cardiac rates have been measured before, the subject's rates also after each palpation session by a nurse.

The reported frequencies of the PRM range from 3 – 9 cycles / minute (0.05 – 0.15 Hz) with mean values ranging from 4.5 to 7.0 cycles / min and a standard deviation ranging from 0.8 to 1.5 cycles / min. Analysing interexaminer agreement the authors did a analysis of variance for repeated measures (anova) and calculated an ICC of –0.02 ($p = 0.0001$). The authors suggest different possibilities interpreting a negative ICC and conclude that in their case this is due to lack of agreement. On the contrary interrater agreement for the nurses who measured cardiac and respiratory rates was found to be moderate (ICC = 0.66 – 0.76). Regression analysis, which was done in order to find correlations between the palpated PRM and cardiac as well as respiratory rates showed low linear relation and therefore was not significant. The authors suggest that, following their results, craniosacral rhythm palpation does not seem to be a reliable examination tool and its use for clinical decision making should be considered. Further there seems to be no relation between the palpated PRM frequencies and the cardiac and respiratory rates of the examiners and the subjects.

The presented results can be qualified in some points. The examiners were not experienced osteopaths. During the educational period an osteopath runs through about ten to fifteen courses in cranial theory and practice. Therefore the examiners cannot be regarded representative for a group of experienced cranial clinicians. The measuring time for the PRM frequency was too short. As frequency was measured by counting the cycles per minute, the possible beginning of another cycle at the end of the measuring period could have produced significant measure errors, especially when the cycles were reported verbally. The authors refer to this deficiency in their article. The palpation has not been carried out simultaneously, so the conditions for the examiners have been unadjusted by time. The outcome of the regression analysis has to be seen with limitations as the two related variables have not been measured simultaneously and the agreement for the measurement of cardiac and respiratory rates is only moderate. In this context also see the invited commentary by Echternach (1994).

Norton (1996) did an investigation of intra- and interexaminer agreement of findings in craniosacral rhythm palpation. Six experienced osteopaths assessed nine subjects. The subjects were palpated simultaneously at the head and at the feet, the phases of the PRM were recorded by a knee switch as used in earlier studies (Norton et al. 1992). The accuracy of the activation of the knee switch has been proven by the author in preliminary ratings, where the examiners had to record respiratory rates and a regression analysis produced significant interrater correlation ($r = 0.934$; $p < 0.001$). As in earlier studies (Norton et al. 1992) besides the frequencies the duration of the cycles of the PRM were integrated into the analysis by letting the examiners activate the switch during the whole flexion period. The duration for recording was one minute, after a one to two minutes break the examiners changed positions. Four subjects have been assessed by all six examiners, five subjects by two.

Regression analysis showed a significant intraexaminer correlation for the findings (cycle length) detected at the head and at the pelvis ($r = 0.926$, $P < 0.001$), interexaminer agreement was found to range from poor to beyond chance agreement values (head: $r = -0.275$, $p = 0.115$; pelvis: $r = -0.296$, $p = 0.089$). The palpated frequencies ranged from 2.14 to 7.38 cycles/minute (0.04 – 0.12 Hz). The author suggests that the results contradict the model of motion transfer between the cranium and the sacrum. On the other hand they support the hypotheses of the *tissue pressure model* (Norton 1991) by showing significant intrasubjective agreement.

The integration of cycle duration into the analysis and the simultaneous taking of palpation (measures) optimise the methodological approach. Therefore the data can be regarded as considerably meaningful. Unfortunately the number of subjects and comparable measures is small. The choice of the examiners is perfect.

Hanten et al. (1998) assessed the reliability of craniosacral rhythm palpation and the correlations to subjects' cardiac and respiratory rates with two examiners and 30 healthy subjects. The examiners had 11 months of experience in craniosacral palpation.

Heart rates were measured by Polar® heart rate monitors, the respiratory rates by visual observation of chest movements, taken by an assistant. PRM cycles were indicated by the examiners by external and internal rotation of the foot and counted by another assistant. All measures were taken synchronously twice over a period of three minutes. The repetition was done in order to be able to analyse intraexaminer reliability.

The authors report mean palpated PRM frequencies with 4.2 (0.07 Hz) and 3.6 (0.06 Hz) cycles/minute with a standard deviation ranging from 1.0 to 1.2 cycles/minute. Analysis of variance (anova) was done to judge intra- and interrater reliability. The ICC for interrater reliability was found to be low (0.22) like in former studies (Norton, 1996; Wirth-Pattullo and Hayes, 1997). Intrarater consistency showed good values (ICC = 0.78 and 0.83). The authors suggest that one examiner can palpate PRM cycles consistently whereas interexaminer agreement is low. Multiple regression analysis produced low values showing no significant relation between PRM frequencies and heart- as well as respiratory rates. By using a measuring duration of three minutes, taking synchronous measures and recruiting a greater number of subjects the authors follow an optimised methodological design (in relation to Wirth-Pattullo and Hayes, 1994). Unfortunately the examiners were relatively inexperienced.

Rogers et al. (1998) analysed intra- and interrater reliability for craniosacral rhythm palpation simultaneously at the head and at the feet. 28 adult subjects with different past and present medical problems. The examiners were a physiotherapist with five years experience and a nurse with 17 years experience in craniosacral palpation. The simultaneously working examiners were blinded by a curtain hanging in between and the PRM was recorded by the activation of a silent foot switch at the beginning of each flexion phase. Each subject has been measured four times. For the third and fourth measurement the examiners' positions changed between head and feet. Between the measures there was a two minute break. The total spread of recorded PRM-frequencies ranged from 0 to 8.42 cycles/minute (0 – 0.14 Hz), with mean values ranging from 3.17 to 4.37 cycles/min and a standard deviation ranging from 1.04 to 1.63 cycles per minute. A frequency of zero corresponds with a shut down of the whole system as described in cranial therapy and is called *stillpoint*. The authors wanted to include the possibility of the stillpoint into their analysis.

The authors undertook an analysis of variance like in previous studies by Wirth-Pattullo and Hayes (1994) and Hanten et al. (1998). The differentiation of two assessment locations (head and foot) produces multiple correlation values. Like in previous studies the authors find low interrater reliability (ICC at the head = 0.08, ICC at the feet = 0.19). In contrast to the previous results (Norton, 1996; Hanten et al., 1998) intrarater reliability was also found to be low (ICC at the head = 0.18 and 0.26; ICC at the feet = 0.3 and 0.29). By integrating the possibility of a stillpoint the authors did another analyses, where the examiner who recorded a frequency of zero has been filtered and found the ICC to be a little bit higher (ICC at the head = 0.23; ICC at the feet = 0.6). But the pure possibility, that one examiner detects a stillpoint while the other finds normal PRM cycles, is interpreted by the authors as a sign for inconsistency.

The authors conclude that, following their results, craniosacral motion cannot be palpated reliably. They end up with the question how experts in cranial therapy can agree in complex qualitative assessment criteria when they don't agree in the simple beginning of the flexion phase. This study can be regarded as an adequate approach towards investigation of the reliability of craniosacral rhythm palpation. As to interexaminer reliability the results agree with results in previous studies (mentioned before).

4.2 Synthesis

On the whole the existing material, assessing the reliability of cranial findings, is not very extensive. In this context Echternach (1994) states, that experts in cranial therapy are obliged to work very hard in order to prove their hypotheses. So this thesis can be regarded as a further small contribution. The reasons for diverging results between elder and recent studies have already been mentioned. They are to be found in data analysis. The bare results of recent research suggest different causes and consequences.

Several methodological aspects can be criticised. But the fact, that all recent studies end up with poor results concerning interexaminer reliability, should be thought over. All authors of recent publications agree, that in relation to low interexaminer agreement, clinical decision making on the basis of craniosacral rhythm palpation has to be regarded as doubtful. As cranial techniques are used more and more, further research seems useful.

5 The hypothesis of the thesis

The definite mechanical aspects of the Cranial Concept have brought up the idea, to do a quantitative investigation of the assessment of the phases of the PRM. Referring to the sources and publications around the Cranial Concept and regarding the education in cranial theory and practice, one has to be aware that explanations to a great extent are qualitative of nature and follow metaphysical-philosophical and sometimes even religious approaches. In this sense arguments critisising a quantitative approach to the Cranial Concept are understandable. But by denying any scientific background we have to be prepeared to be confronted with insuperable barriers concerning communication in training, with patients or with colleagues and medical professionals who are not acquainted with the fundamentals of the Cranial Concept. In consequence this might lead to esoteric exclusion of the Cranial Concept or even Osteopathy in general and a growing disability for interdisciplinary dialogue. The introducing words in the information folder for the courses in *Biodynamic Cranial Osteopathy* at the *Internationale Schule für Osteopathie* in Vienna from November 1998, lead by Jim Jealous, a present leading protagonist in the field of craniosacral therapy, express this problem:

> »[...] Osteopathy has shamefully hidden it's greatest mystery as well as it's biggest resources [...]«

In addition Jeffrey et al. (1997) mention in their *Methodological Manifesto* addressing quantitative research methods in complementary and alternative medicine, that investigations of so called interparadigmatic therapy concepts with the tools of commonly used valid scientific methods is meaningful and possible. The following presumptions are the basis for the hypothesis. They have been analysed with respect to their correctness and relevance within the Cranial Concept in the previous chapters:

- The possibility of the existence of an autonomous mechanism which acts in rhythmically repeated cycles and is defined as *Primary Respiratory Mechanism* (PRM) within the Cranial Concept.
- The autonomy of this rhythm against other known physiological rhythms like thoracic respiration and heart rate as well as against the influences of voluntary motion and posture.
- The influence of the expansive and contracting cycles of the mechanism on the following structures in the sense of changes in the mentioned physical and physiological parameters:
  - (1) Hydrodynamic changes with respect to the fluctuation of the CSF.
  - (2) Changes of tension in the meningeal System, which is called *Reciprocal Tension Membranes* in the Cranial Concept.
  - (3) Movement of the osseous elements of the skull in relation to their synarthrotic articulations like sutures, synchondroses and syndesmoses.
- The presumption that this mechanism is able to induce involuntary motion to the sacrum
- The presumption that the cycles of the mechanism are externally manual detectable at the head and at the pelvis for trained persons.

Summarizing the following hypothesis that underlies the following measures can be deduced and formulated:

If the PRM as a fundamental element of the Cranial Concept

(1) represents a physical or physiological phenomenon, whose effects occur as presumed by the Cranial Concept and if further on

(2) these effects can be reliably palpated by trained persons,

then statistical significant intra- and interexaminer agreement with respect to the assessment of the cycles of the PRM should be reached when two trained examiners are palpating one subject simultaneously and repeatedly within a short time interval without showing any dependency of the PRM cycles to examiners' and subject's respiratory rates.

The planned measures should help to interpret the grade of intra- and interexaminer agreement in the palpation of the PRM and thus provide information about consistency, reproducibility and reliability of craniosacral rhythm palpation. The necessity of assessing the mentioned factors has repeatedly been clarified in the previous chapters. A summarising explanation is given now:

- The assessment of the cycles of the PRM at the head, the pelvis or other locations represents one of the first steps in the diagnostic procedure used in the Cranial Concept.
- Statements about existence or non-existence and the frequency of the PRM should therefore be reproducible, otherwise

  (1) reliable palpation of the PRM seems not possible

  (2) interexaminer communication based on the hypotheses of the Cranial Concept seems not possible

  (3) it can be suggested, clinical decision making is dominated by subjective presumptions
- As the existence of the PRM is still unclear, good interexaminer agreement in palpation of the PRM may support the clinical acceptance of the cranial approach.

6 Methods

The applied methodology followed the guidelines for inter- and intrarater reliability studies for repeated measures in manual medicine, which have been discussed under point 2.1.3 on p 23ff.

6.1 The examiners

Limitations of the time-period that was at disposal for this project made it necessary to reduce the sample size for examiners to two. Both examiners finished the six year course at the *Wiener Schule für Osteopathie* two years ago. Since two years they participate in a postgraduate project for osteopathy in the paediatric field in cooperation with the *Osteopathic Centre for Children* (London), where the main emphasis for clinical work is based on the principles of the Cranial Concept. One examiner is a doctor from her basic medical education, the other one a physiotherapist. At the time the measures were done both had about 300 hours of teachings in cranial techniques and theory as well as about seven years of clinical experience behind them. Their therapeutic approach is based to a great extent on the theoretical basis and clinical means of the Cranial Concept.

6.2 The subjects

49 healthy adult subjects (n = 49) with a mean age of 37.45 ± 7.52 (min = 19; max = 61) have been assessed. 34 were female, 15 male. The subjects were recruited from our circle of acquaintances, students from the *Wiener Schule für Osteopathie* as well as acquaintances of the students. The following exclusion criteria have been defined:

- Severe trauma, surgery and current acute pain in the area of the cranium, the spine and the pelvis.
- Current and past diseases of the central nervous system and its envelopes.
- Severe mental disturbances.

Disturbances in the area cranium-spine-pelvis in the sense of mechanical dysfunctions and slight degenerative changes as well as past medical problems, traumas and surgery from outside the cited region were not excluded in order to reach sufficient data variance for the analysis. Besides the subjects had to be older than 18 years, they had to be able to understand the necessary instructions for the measures and lie in supine position for about 15 minutes.

The subjects have been informed about the procedure during the measurements as well as the fact that no treatment does take place. They received no compensation and participated voluntarily.

6.3 Measuring devices

Measuring devices were needed to record the palpatory findings (PRM-rates) of the examiners and the respiratory rates of the subjects and the examiners during palpation. For recording the palpated PRM-rates two foot switches were used, which should be activated in relation to the phases of the PRM. So the two states ON (1) and OFF (0) resemble the flexion- and extension-phase of the PRM. As the examiners are seated during palpation, the activation of the switches should not cause any trouble.

The recording of the respiratory rates has been enabled by using *strain gauges*, that were glued to metal bows. The metal bows can be fastened around the thorax via a non-elastic belt. When the bow gets stretched or relaxed by the changes of the rib cage caused by in- and expiration, the strain gauges will cause measurable changes in voltage.

As for our investigations no suitable measuring devices were at hand, development and construction of the needed equipment had to be realised be ourselves. The choice of strain gauges as well as their specific gluing to the metal bows has been carried out by Mr. Tanzer from the Measurements Group®. The whole recording-equipment including hard- and software has been developed and constructed by Mr. Stacher, a pupil from the *Höhere Technische Bundeslehranstalt Hollabrunn* (engineering school) under the supervision of Dipl.Ing. Dum as a graduate project.

6.3.1 The hardware

The measuring equipment had to fulfil the following tasks synchronously:

- Recording of the palpatory findings of both examiners via two foot switches. The switches were put into a spring-suspended metal case, that could be used as a pedal. The excursion of the pedal is 6 mm.
- To get the examiners blinded the switches had to be mute. This could be reached by using fork-light-barriers, as common types of switches always produce a click-tone. The signals of the switches could be recorded directly as digital units "0" and "1".
- Recording of the respiratory rates of the examiners and the subjects. This was carried out by strain gauges (type: N2A-06-S153R-35B) that have been glued to metal bows which have been provided with a non-elastic belt that could be fastened around the thorax (see Fig. 2 on p 71). The used strain gauges had a resistance of 350 Ohm. Strain gauges react to very small deformations via changes of their electric resistance. In order to produce a clean signal and rule out influences due to thermal fluctuations causing measurement errors, the strain gauges were connected by a *full bridge circuit* (see the circuit diagram in Fig. 2 on p 71). So four strain gauges are needed, two gauges were glued to the front side of the bow and two to the backside.

A metal bow, equipped with four strain gauges will in future be called *respiratory rate transducer* (RRT). For further processing by the measuring module the signal (S), produced by the bridge circuit, had to be amplified via a special strain gauge amplifier. The amplification could be regulated for each input-channel separately. This allowed adaptation to individual respiration excursions in order to produce the best possible data.

For direct processing of the measured signals an *ELV measuring module* type M232 was used which was connected to a computer via the serial interface. The module has five analogue input channels (CH1–5). The integrated analogue-digital-converter (A/D-converter) makes the signals readable for the computer. The five channels of the module have been occupied by (see Fig. 3 on p 72):

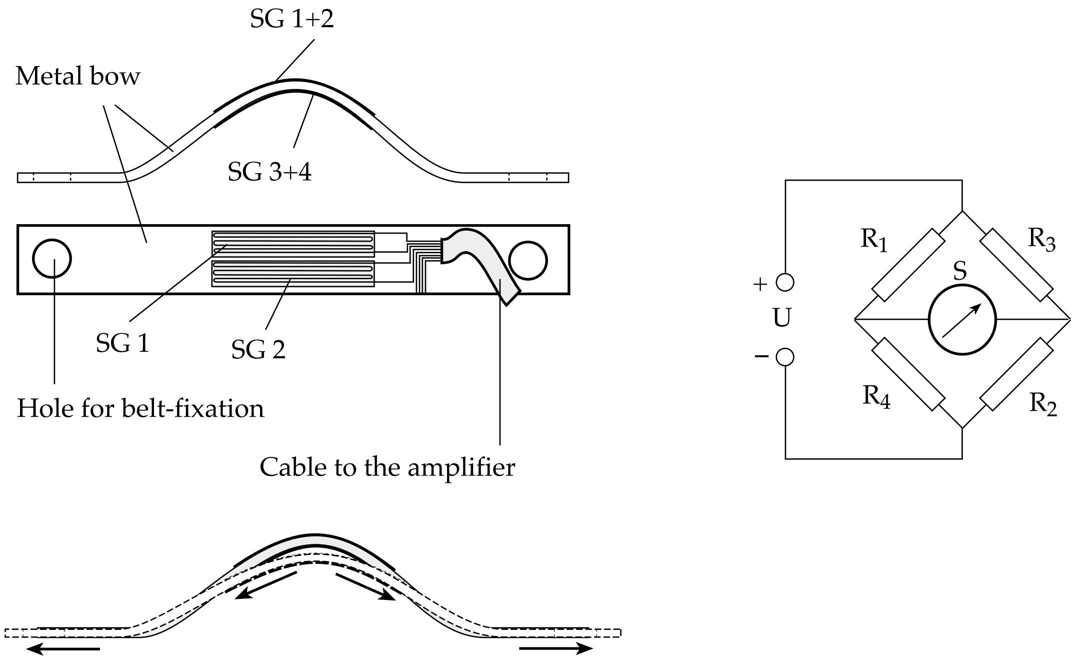


Figure 2

Depiction of a metal bow equipped with strain gauges (SG) for the recording of the respiratory rates in a view from above and from the side. Two SGs are glued to each side of the bow. All four strain gauges together represent a full bridge circuit (diagram to the right). The dimensions of the bows were about 14 cm in length and 3 – 4 cm in height. As only the frequency and not the quantity of the respiration excursions has been of interest for the measures, the dimensions were not calibrated. The lower picture shows schematically how the bow and in consequence the strain gauges are deformed during inspiration.

R_{1-4}.. 4 strain gauges
S....... Signal (measuring current)
U....... supply tension

- Foot switch 1 for examiner location (L) cranium (C)
- Foot switch 2 for examiner location (L) pelvis (P)
- RRT 1 for examiner A (RE_A)
- RRT 2 for examiner B (RE_B)
- RRT 3 for subject (RE_{SU})

To make sure that the data, produced by the RRTs are in fact mirroring the respiratory movements and not fluctuations of the supply tension, the power supply for the measuring module had to be held as constant as possible. We used an Artesyn® power supply type SSL40-7615 to guarantee that. The tension of the power supply is 15 V, its

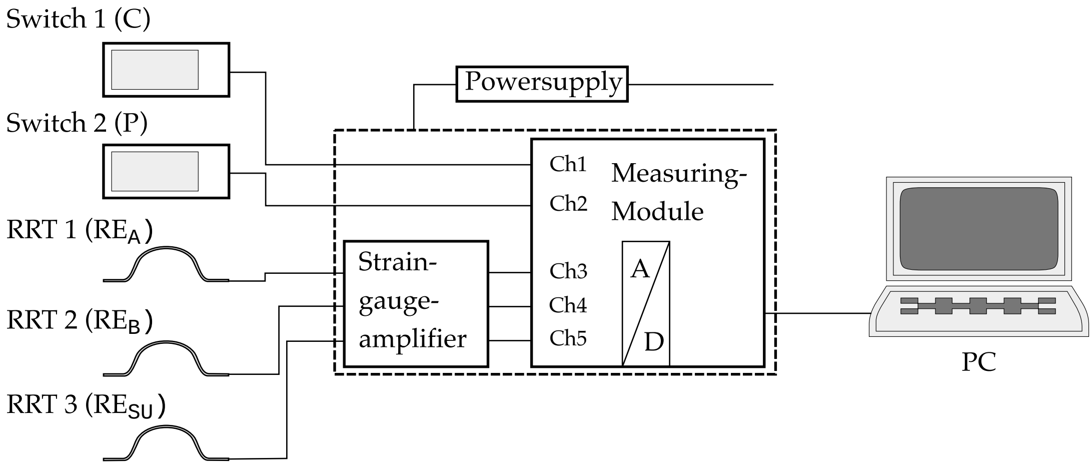


Figure 3

Module diagram of the whole measuring equipment. As the signals from the RRTs are analogue (changes of voltage) in the beginning, they have to be changed into digital signals by the A/D converter in order to be readable for the computer. The signals coming from the RRTs have to be amplified before by the strain gauge amplifier.

CH1–5 are the input channels of the measuring module. They are occupied by:
Switch 1 for examiner location cranium (C)
Switch 2 for examiner location pelvis (P)
RRT 1 for examiner A (RE_A)
RRT 2 for examiner B (RE_B)
RRT 3 for subject (RE_{SU})

strength is 2.6 A. The difference between signals coming from one RRT recording respiratory movements and the two others that did not, can be seen in Fig. 4 on p 73.

The time-related resolution of the measuring module offered two final possibilities: one signal per 300 milliseconds (ms) or one signal per 500 ms. As the future analyses of time-related data would be much easier with 500 ms, this resolution has been used. A lower resolution of 100 ms has been suggested at the beginning of the planning phase. An adaptation of the equipment to a higher resolution would have been possible by using a single-board-computer instead of the measuring module. Limited time did not allow this improvement.

6.3.2 The software

The software has been written in Lab Windows® CVI 16 Bit-Version 4.0.1. For the graphical visualisation during the measurements a user-interface has been developed, that showed each input channel (RRTs 1–3, Pedal 1, Pedal 2) as well as quantity and

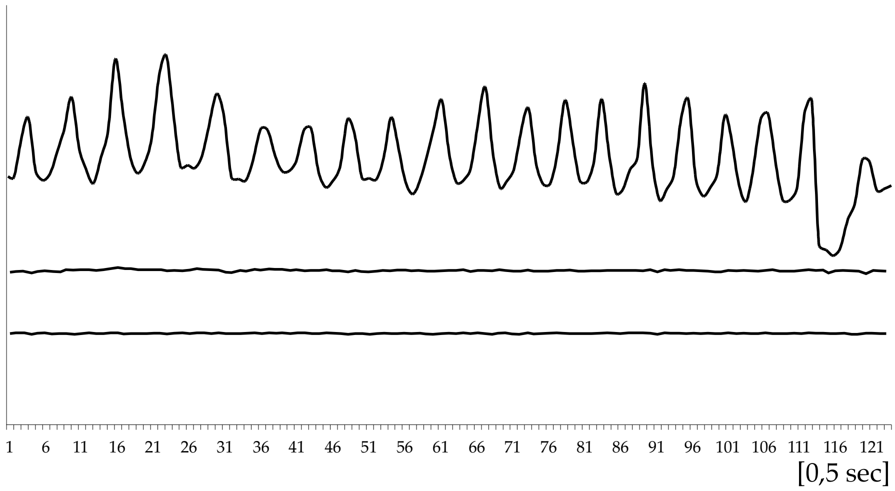


Figure 4
Recording of the data of all three RRTs over a time period of two minutes. For the control of the tension provided by the power supply, one RRT has been fastened and shows a typical curve for the respiration movements (superior graph). The other RRTs were left unused (middle and inferior graph). Comparing the superior to the middle and the inferior graph it can be seen that supply tension is constant and respiration movements can be detected.

time of each measured value via graphs. Each single measurement could be started, ended and selectively stored. Starting time and the measured time period were stored in order enable the identification of each measurement for further data analysis. After storing the data could be transferred into the spreadsheet program Excel® for further preparation and statistical analysis.

6.4 Procedure

6.4.1 Preliminary ratings

To guarantee a smooth procedure during the actual measurements, the main details have been fixed in preliminary ratings with two subjects. In the beginning the activation of the foot switch has been determined: The activation should start as soon as the examiner feels the beginning of a flexion phase of the PRM. Then the switch should be held activated over the whole flexion phase and should be let gone as soon as the beginning of the extension phase of the PRM could be perceived. Each palpated flexion phase correlates to the value „1" in the data sequence, while the extension phase correlates to the value „0". Graphically the data can be plotted as a square wave (see

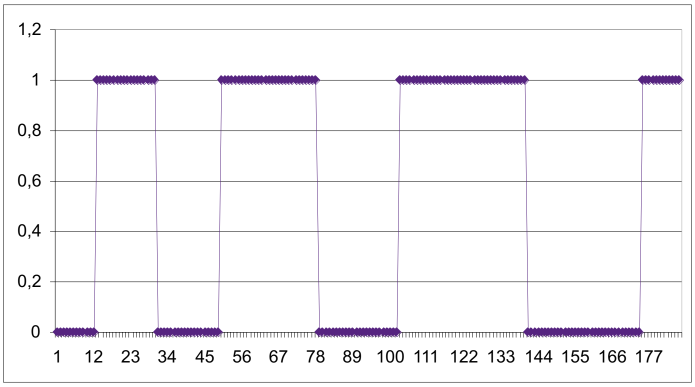


Figure 5

Graphical example of recorded signals from one foot switch over a time period of 90 seconds (sec). Every 0.5 sec a signal is detected. The value 0 means switch in position OFF and mirrors the palpation of the extension phase, the value 1 means switch in position ON and mirrors the palpation of the flexion phase.

Fig. 5). In order to develop an adequate pressure for activating the switch, the examiners could have a look at the display were the signals get plotted. This should bring an enhancement in the synchronicity of the examiner's perception and the activation of the switches. In order to avoid further disturbances and get the preconditions as good as possible for concentrated palpation, the examiners could use their preferred hand holds and had about one minute for each subject to get orientated. Both examiners used a standard vault hold with the index bilaterally on the great wings of the sphenoid and the 4^{th} and 5^{th} finger bilaterally on the lateral parts of the squama occipitalis. For the pelvis both examiners used a standard hand hold with one hand under the sacrum with the fingers pointing cranially, positioned under the basis, the thenar and hypothenar eminence under the apex of the sacrum.

During the actual preliminary ratings the final length of the measuring period has been fixed. We started with three minutes, reduced to two minutes in a second go and finally fixed a duration of 1.5 minutes (90 sec). The reduction to 90 sec was necessary in order to rule out therapeutic effects as good as possible. Using longer time periods the examiners reported difficulties to stay in a clean observing position. As in future data analysis the PRM-frequency would be one of the main parameters, regarding the problems Wirth-Pattullo and Hayes (1994) reported in connection with possible erroneous values for the number of PRM cycles when using a measuring time of just one minute, the decision for 90 sec seemed reasonable.

6.4.2 Final experimental setting and procedure

Five minutes before each set of measures (consisting of two measurements) each subject positioned her/himself supine on one of the treatment tables. So we reached a certain relaxation of the subject and a stabilisation of her/his respiratory rates. To guarantee the time period for preparation, two treatment tables were used. So there was always one subject in preparation while the other one was palpated. The examiners had to change positions within and between the tables. At the beginning of each set of measures the RRT has been fastened around the subject's thorax, then the usability of the produced data has been checked. The RRTs for the examiners have been fastened and checked at the very beginning of a whole series of measurements. So each examiner always had the same transducer connected to the same input channel. This was important for a clear identification of the respective respiration curves.

In order to be able to assess interrater agreement simultaneously, two examiner locations (L) were chosen, which were used simultaneously during each measure. One examiner sat at the head end of the treatment table palpating the cranium (C), the other one sat at the side of the treatment table next to the subject's pelvis, palpating the sacrum (P). The decision for the locations cranium and pelvis (sacrum) follow common clinical practice within the Cranial Concept.

Both foot switches were placed near the locations for palpation, while switch 1 was always used for the examination at the cranium (C) and switch 2 for the examination of the pelvis (P). Thus for the recorded data switch 1 always corresponded to location (C) whereas switch 2 always corresponded to P. To avoid any confusions during the measurements the switches were clearly marked with „HEAD" and „PELVIS". As both examiners used their right hand to palpate the sacrum, the treatment tables were positioned opposite due to their head ends. Thus switch 2 for location P could always rest between the tables at about the same place. Switch 1 had to be moved from one head end to the opposite head end of the other table before every new subject (see

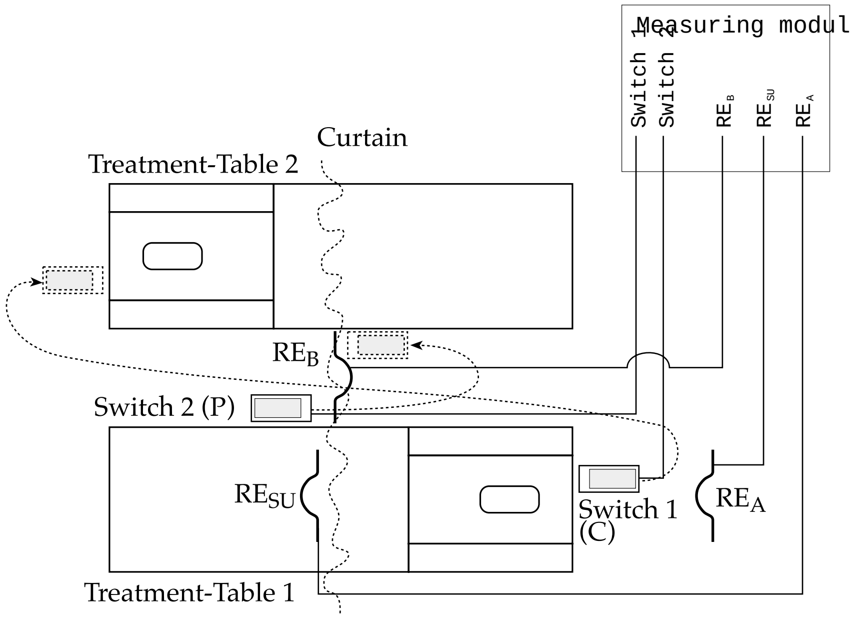


Figure 6
Schematic view of the arrangement of treatment tables, foot-switches and RRTs as well as the wiring during the measurements.While one subject has been palpated on table 1, the next subject was already lying on table 2. For the change from the first subject to the next, the RRT for the subject (RE_{SU}) and switch 2 at the head end (C) of the table had to be changed (see broken lines with arrows), while the following subject took place on table 1 for preparation time.

Fig. 6). In order to blind the examiners visually, a curtain was hung from the ceiling dividing the treatment tables into two parts. To prevent the examiners from any kind of acoustical queuing the foot switches were mute (see point 6.3.1 on p 70ff). In addition background noise from a tape recorder was used (noises from the rain forest) , especially to avoid any queuing via breathing sounds.

For future differentiation between measures taken at different locations and for the possibility of analysing intrarater reliability, each subject has been measured twice. This gives a total of 98 measures for 49 subjects. For the second measurement the examiners switched places. Besides examination location (L) we use examination time (T) as another variable. Both variables show two possibilities (levels): Cranium (C) and pelvis (P) for L, first examination time (T 1) and second examination time (T 2) for T. The location for examiner at T1 was randomised. For the final distribution see Table 4 on p 77.

N	O x L	T1	T2
23	O	B A	A B
	L	C P	C P
26	O	A B	B A
	L	C P	C P

Table 4
List of the distribution of observers (O) A and B at the locations (L) cranium (C) and pelvis (P) for the examination times T1 and T2. This list shows the randomly produced distribution at the end of all measurements. It can be seen that examiner B started 23 times, A started 26 times at the head.

18 subjects have been measured a third time (T 3) and 10 a fourth time (T 4, by changing locations again). After T 2 the examiners had to palpate the next subject and come back to the previous for T 3 and T 4. So we should get more precise information about intrarater agreement, as the same examiner assessed the same location repeatedly within a short time interval. By putting the next subject in between, a blinding-effect of the examiners against themselves should be reached. Limited time for the data analysis forced us to leave the data for the time being. Future analysis will be possible.

To avoid errors in later data-identification the data for subjects and measurements were linked via two related data bases. Each measurement has been stored separately and the name of the file had the exact information for further analysis as number of the subject, number of the measurement and the examiner that palpated the cranium. For instance file „S05_P15_A" means that subject No 5 was palpated in measurement No 15 with examiner A at the head. A sample of data collected over a time-period of 90 seconds can be seen in Fig. 7 on p 78.

6.5 Data analysis

The data analysis has kindly been prepared carried out by Mag. Kaider from the *Institut für medizinische Computerwissenschaften der Universität Wien* (*Department of medical computer sciences at the Vienna University*), who also advised us in methodological aspects concerning data analysis. The used statistics program was SAS Institute Inc. SAS/STAT® Version 6, 4th Ed., 1989.

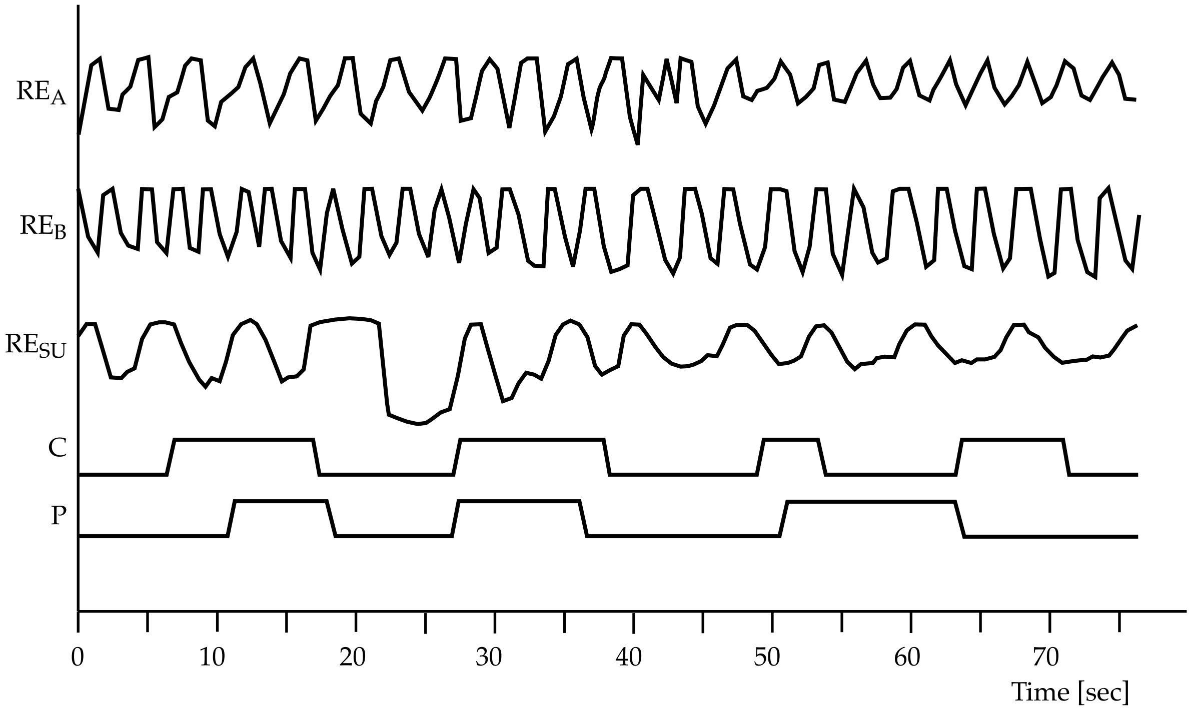


Figure 7

Sample for a typical complete measurement over a time period of 90 seconds. The topmost graph shows the respiration of examiner A (RE_A), underneath the respiration of examiner B (RE_B) can be seen. The next graph shows the respiration of the subject (RE_{SU}). The two square wave graphs at the bottom show the signals of the foot switches. The upper square wave shows the palpatory findings for the PRM at the cranium (C), the lower at the pelvis (P). For this measurement examiner A palpated the cranium.

The main outcome-variable was represented by the PRM as described in the Cranial Concept. The PRM has been palpated simultaneously by two examiners (A, B) and repeatedly by changing the measurement location (pelvis, cranium). The PRM-rates, perceived by the examiners, have been recorded as described in point 6.3 on p 69ff. The produced data have been assessed with regard to three problem-orientated parameters (see Fig. 8 on p 79):

- PRM-Frequency or -rate (f); it is represented by the number of flexion-phases within 90 seconds.
- Mean duration of the flexion-phases (MDF).
- Mean ratio of the lengths of the flexion- to extension-phase ($R_{F/E}$).

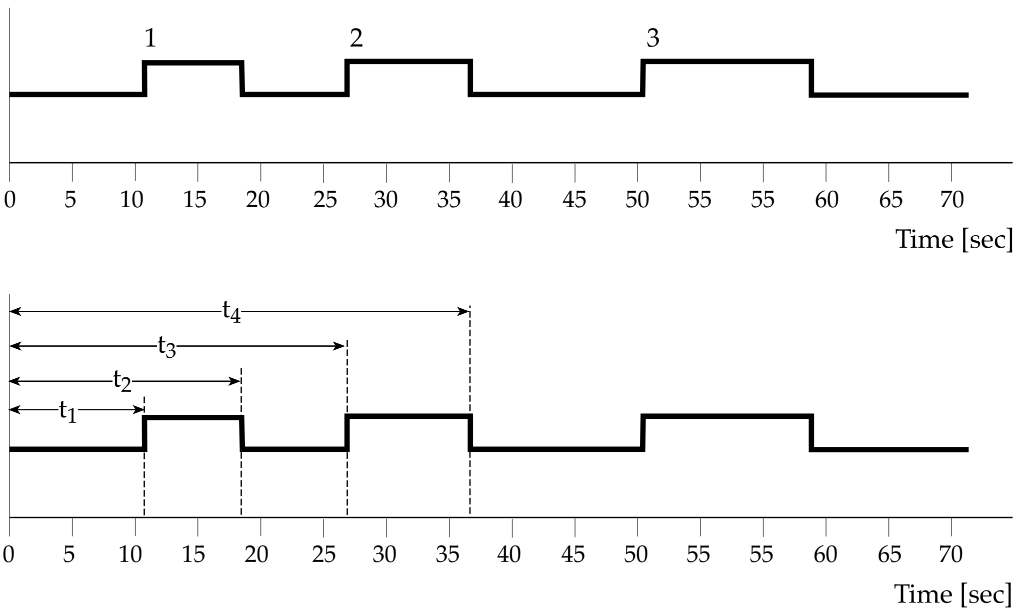


Figure 8

Schematic presentation of the determination of the three problem-orientated parameters frequency (*f*), mean duration of the flexion-phase (MDF) and mean ratio of flexion- to extension-phase ($R_{F/E}$). For *f* the number of changes from switch-position zero to one within 90 seconds had to be determined. This is shown in the topmost scheme.

For MDF the differences between t_2–t_1, t_4–t_3, ... had to be formed, from which the mean was calculated.

For the $R_{F/E}$ the ratios of the differences (t_2–t_1)/(t_3–t_2), (t_4–t_3)/(t_5–t_4), … had to be formed, from which the mean was calculated.

The decision to determine three parameters is based on the assumption that the expected distribution between flexion- and extension-phase would be irregular (Lockwood and Degenhardt, 1998). Thus the three parameters *f*, MDF and $R_{F/E}$ would enable differentiated possibilities in data analysis and future interpretation. If for instance the examiners would palpate similar lengths of the flexion-phases but different lengths of extension-phases, there would be reasonable interexaminer agreement for MDF but small or absent agreement for *f*. Another scenario cold be given by one examiners palpating about twice the length for the flexion- as for the extension-phase as the other one while the mean duration for the flexion-phases differ. In this case interexaminer agreement would be absent for *f* and MDF but reasonable for $R_{F/E}$.

6.5.1 Descriptive statistics

To start with a descriptive analysis has been carried out, considering the following three factors:

- ■ Examiner-location (L) showing two levels
  - Pelvis (P)
  - Cranium (C)
- ■ Examiner or Observer (O) showing two levels
  - Examiner A
  - Examiner B
- ■ Examination-time (T) showing two levels
  - Time 1 (T 1) which means the first measurement
  - Time 2 (T 2) which means the second measurement (= repetition for one subject)

6.5.2 Analysis of variance

In addition the three problem-orientated parameters have been tested with regard to systematic differences under various conditions. For that, including the random factor subject, a *four way analysis of variance* (anova) (Kirk, 1982) has been used. The dependent variables *f*, MDF and $R_{F/E}$ have been tested with regard to the following factors:

- ■ Differences between examiners concerning *f*, MDF and $R_{F/E}$
- ■ Differences between examiner-locations concerning *f*, MDF and $R_{F/E}$
- ■ Differences between examination-times concerning *f*, MDF and $R_{F/E}$

The random factor subject has been taken into account in the model of analysis. Further possible interactions between examiner and examiner-location have been tested by analysis of variance. If significance was present, corresponding additional analyses have been carried out by testing the differences between the examiners separately for the examiner locations cranium and pelvis. So we could find out if one examiner palpated systematically higher PRM-frequencies at one examiner-location than the other.

6.5.3 Analysis of inter- and intraexaminer agreement

Concerning the assessment of the hypothesis (see p 65) one of the essential issues in data analysis was the description of inter- and intraexaminer agreement (or reliability). The main question in this context was: can the PRM be palpated reliably and consistently?

Intraclass correlation coefficients (ICC) are frequently suggested for the description of agreement for continuous data (Haas 1991, 1995) and have repeatedly been used for the calculation of inter- and intraexaminer reliability for the palpation of the PRM (Wirth-Pattullo and Hayes, 1992, Hanten et al., 1998; Rogers et al. 1998). As Bland and Altman (1986) state, the interpretation of the comparison between two measures can be misleading, when *correlation* is mixed up with *agreement*: correlation just indicates that there exists a linear relationship between two variables. Correlation would for instance be perfect when examiner A would always measure twice the value of examiner B. But this does not mean that the examiners agree (see also p 24ff). So we decided to use an alternative method suggested by the same authors using the differences between the measurements known as the *95% limits of agreement* (Bland and Altmann, 1986).

Thus we assessed the expected variability between the palpatory findings (*f*, MDF and $R_{F/E}$ of the PRM) of the both examiners and within one examiner whilst one of the three factors (L, O, T) has been held constant. Therefore the analyses are either unadjusted by time (the comparison of measures between T 1 and T 2 that were not produced simultaneously) or by location (the comparison of measures between C and P that were not taken at the same location). In the case of the analysis of intrarater agreement both effects of unadjustment were present. The *95% limits of agreement* (Altmann 1999, p 399) describe the range of values that represents future differences between the measures (examiners) for individual measurements with a 95 % probability. As for normal distribution (95 % of the values will lie within a range of ± 2 standard-deviations from the mean) the *95% limits of agreement* can be calculated as follows:

$$\text{Mean difference} \pm 1.96 \times \text{standard deviation of the differences} \qquad (3)$$

„*The mean difference is an estimate of the average bias of one method* [which in our case is the examiner] *to the other*" (Altmann 1999, p 398). But the essential outcome for the description of agreement is the standard-deviation of the differences, which enables a prognosis for the accuracy of agreement in one individual measurement. The smaller the range within the 95% limits of agreement is in comparison to the total range of values produced on the whole, the better agreement will be. Are the 95% limits of agreement indicating about the same spread as the produced data on the whole, no agreement beyond chance can be described. At which accuracy agreement is acceptable, depends on specific clinical circumstances: „*It is not possible to use statistics to define acceptable agreement*" (Altmann 1999, p 400).

6.5.4 Analysis of covariance and correlation

As in the Cranial Concept the PRM is hypothesised to be an autonomous phenomenon, whose rhythm is independent from cardiac and respiratory rates (see p 65ff) and the examiner's own respiration might have an influence on the perception, possible interactions between palpatory findings and respiratory rates of subjects and examiners have been tested.

Therefore models for analysis of covariance (Kirk, 1982) have been used. Besides the factors subject, examiner and time, the respiratory rates have been tested as covariables in the model. These analyses were carried out at a time separate for the independent variables f, MDF and $R_{F/E}$. If interactions were present, the strength of the factor at a time has been additionally analysed separated in examiners (A, B) or locations (P, C). The strength of correlation between the three problem-orientated parameters and the respiratory rates has been additionally described by the *pearson's correlation coefficient* (r).

7 Results

7.1 Descriptive statistics

Using descriptive analysis the spread of the values for the problem-orientated parameters frequency (f), mean duration of the flexion-phase (MDF) and mean ratio flexion-to extension-phase ($R_{F/E}$), which have been ascertained at the various conditions examiner (O), examiner-location (L) and examination-time (T), have been described. The tables contain specific details about mean values, standard deviation (SD) and median values for each of the parameters at the various conditions.

Box-plots were used for the graphical representation of the data. They indicate the frequency-distribution of the ascertained values for each parameter. The ends of the vertical lines (whiskers) show the 5^{th} and 95^{th} centile. In between lie 90% of the values (known as the central range). The upper and lower limits of the box show the 75^{th} and 25^{th} centile (or 3^{rd} and 1^{st} quartile). The box stands for 50% of the measured values (Altmann 1999, pp 31–34; Bortz 1999, pp 38–46).

7.1.1 PRM-frequency (f)

The descriptive analysis for the palpated f for each examiner (A, B) at the respective examination-times T 1 and T 2 for the respective examination-locations cranium (C) and pelvis (P) produced mean values between 2.3 and 3.6 cycles/min (0.04 to 0.06 Hz) with a SD ranging from 0.6 to 1.2 cycles/min (0.01 to 0.02 Hz). More details are shown in Table 5 and Fig. 9 on p 84.

L-O-T	n	Mean			SD			Median		
		90 sec*	60 sec*	Hz	90 sec*	60 sec*	Hz	90 sec*	60 sec*	Hz
P-A-T 1	23	4.6	3.1	0.05	1.8	1.2	0.02	5.0	3.3	0.06
P-A-T 2	26	4.5	3.0	0.05	1.3	0.9	0.01	4.5	3.0	0.05
P-B-T 1	26	4.7	3.1	0.05	1.2	0.8	0.01	4.5	3.0	0.05
P-B-T 2	23	4.6	3.1	0.05	1.1	0.7	0.01	4.0	2.7	0.04
C-A-T 1	26	3.7	2.5	0.04	1.6	1.0	0.02	4.0	2.7	0.04
C-A-T 2	23	3.4	2.3	0.04	1.3	0.8	0.01	4.0	2.7	0.04
C-B-T 1	23	5.4	3.6	0.06	1.2	0.8	0.01	5.0	3.3	0.06
C-B-T 2	26	5.0	3.4	0.06	0.9	0.6	0.01	5.0	3.3	0.06

Table 5
Table of mean values, standard deviations (SD) and median values of the PRM-frequencies (*f*), palpated by the examiners (O) A and B at the examination–times (T) T 1 and T 2 at the examiner-locations (L) cranium (C) and pelvis (P).

n ... number of measurements

* As the time period for measuring was 90 seconds, the results have been transformed into the common units for cycles per minute and Hertz (Hz).

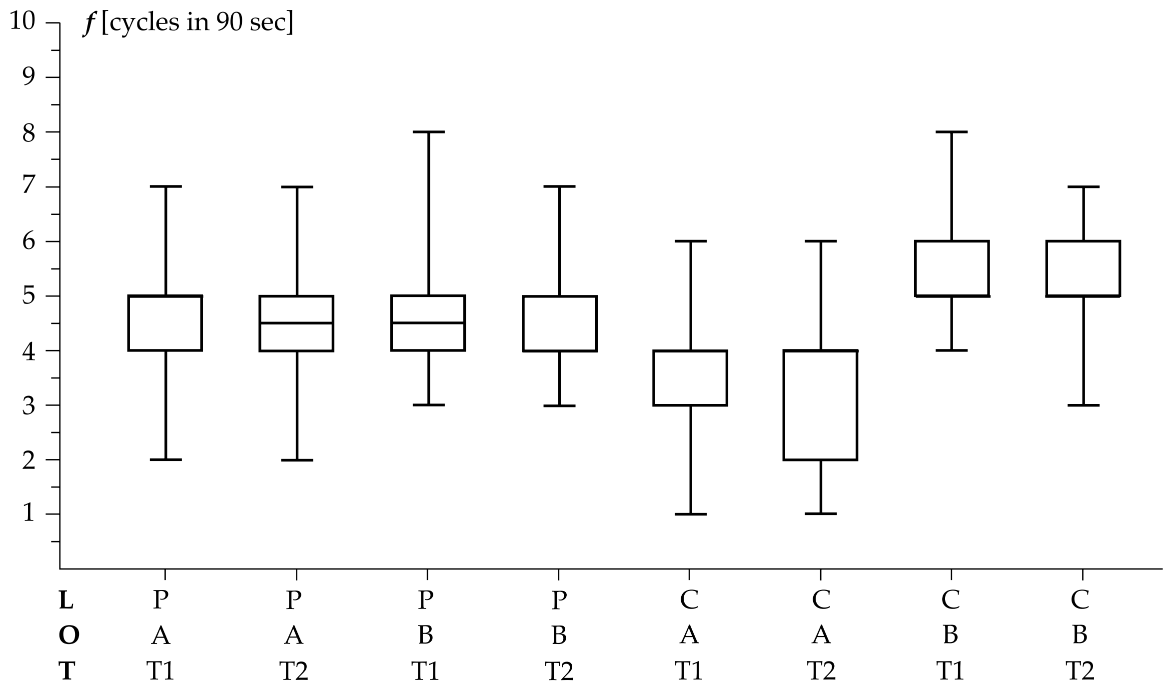


Figure 9
Boxplot of the palpated PRM-rates (*f*) under the different conditions for location (L), examiner (O) and examination-time (T). The whiskers show the central range, the box represents 50% of the measured values.

7.1.2 Mean duration of the flexion-phase (MDF)

Analysing the MDF, one of the 26 measurements had a flexion-phase palpated by A at T 1 at C which took nearly the whole measuring time of 90 seconds. As this exceptional case had relevant influence on the analysis, it was treated as a statistical anomaly and the whole analysis was repeated without the anomaly. The analysis including the anomaly was not taken into account for the final results.

As for *f*, the ascertained values under the conditions O, L and T for the respective possibilities A, B, P, C, T 1 and T 2 have been tested for MDF. MDF with mean values ranging from 8 to 10.6 seconds and a SD ranging from 2 to 5.4 seconds has been found (see Table 6 below and Fig. 10 on p 86).

L-O-T	n	Mean	SD	Median
P-A-T 1	23	8.0	4.3	7.5
P-A-T 2	26	9.2	4.0	9.3
P-B-T 1	26	9.5	2.8	9.3
P-B-T 2	23	9.6	2.4	9.1
C-A-T 1	**25***	10.2	5.4	8.6
C-A-T 2	23	10.6	4.1	10.0
C-B-T 1	23	9.0	2.3	9.1
C-B-T 2	26	9.4	2.0	9.4

Table 6
Mean values and standard deviations (SD) for the respective conditions L, O, T under which MDF has been determined. The values represent seconds.

n … number of measurements

*) One measurement of this combination of conditions has been regarded as an anomaly and was therefore excluded. So the number of measurements is 25 instead of 26.

7.1.3 Mean ratio of flexion- to extension-phase ($R_{F/E}$)

The distribution of the values for $R_{F/E}$ showed the tendency for a lot of low values and some high values, which could not be regarded as anomalies like the one measurement in the MDF analysis. Due to this skewed distribution, the values have been $\log_{10}$ transformed in order to optimise the graphical data presentation. By the $\log_{10}$ transformation the low values get stretched apart and the high values get drawn towards the

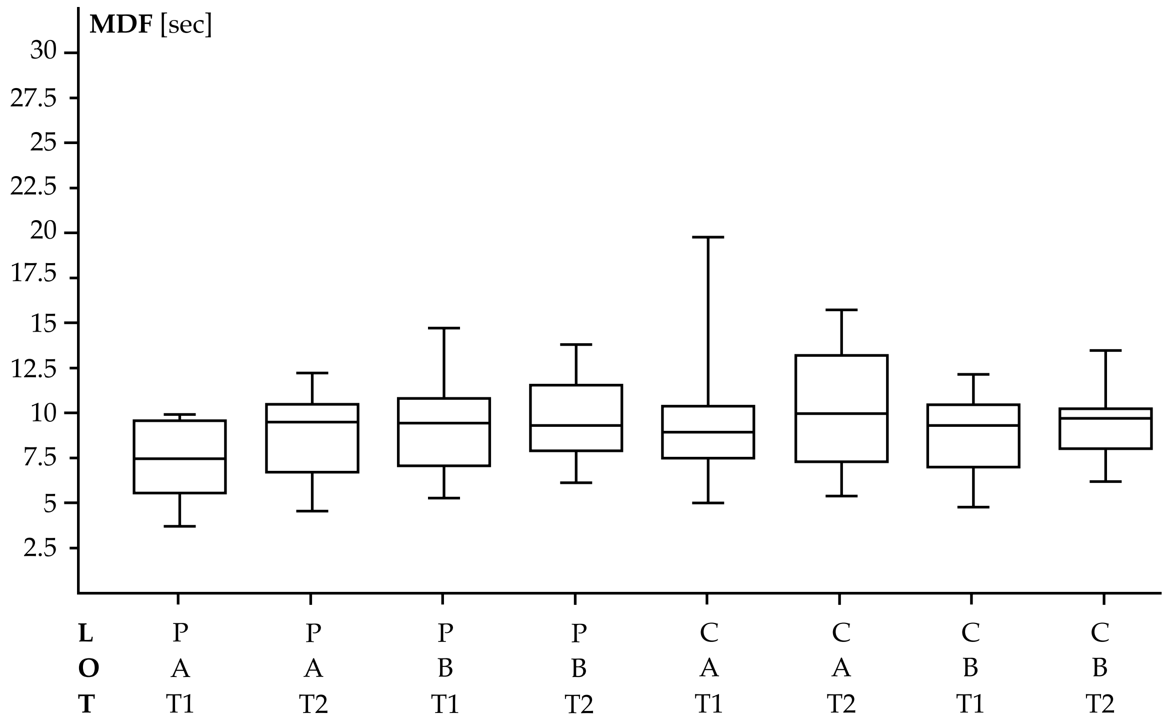


Figure 10

Boxplot-presentation of the palpated mean duration of the flexion-phase (MDF) for examiners (O) A and B at the examiner-locations (L) pelvis (P) and cranium (C) at the examination times (T) T 1 and T 2.

middle. If the transformed value (n) is put as the exponent for the basis 10, the original value can be calculated:

$$n = \log_{10} \mathrm{R}_{F/E} \quad \rightarrow = 10^{n} \tag{4}$$
$$e.g.: 0.4 = \log_{10} \mathrm{R}_{F/E} \quad \rightarrow = 10^{0.4} = 2.5$$

The mean values for $\mathrm{R}_{F/E}$ are ranging from $\log_{10}$ –0.02 to $\log_{10}$ 0.09 with a SD within the range from $\log_{10}$ 0.1 to $\log_{10}$ 0.38. For all values > 1 the flexion-phase is longer than the extension-phase and inversely. Details can be seen in Table 7 and Fig. 11 on p 87.

7.1.4 Minimal and maximal values – total range of values

For the interpretation of agreement, which will be described in point »Analysis of inter- and intraexaminer agreement (reliability)« (see p 93ff), it is interesting to know the spread of all or most (95%) of the values that have been produced under all possible conditions. Therefore the minimal and maximal values, found for each of the three

L-O-T	n	$\log_{10} R_{F/E}$			$R_{F/E}$
		Mean	SD	Median	Mean
P-A-T 1	23	–0.15	0.22	–0.13	0.71
P-A-T 2	26	0.02	0.27	0.02	1.05
P-B-T 1	26	–0.02	0.18	–0.04	0.95
P-B-T 2	23	–0.03	0.10	–0.06	0.93
C-A-T 1	26	–0.06	0.38	–0.09	0.87
C-A-T 2	23	–0.12	0.36	–0.10	0.75
C-B-T 1	23	0.06	0.16	0.04	1.15
C-B-T 2	26	0.09	0.16	0.09	1.24

Table 7
Table of mean values, standard deviations (SD) and median values of the mean ratio flexion- to extension-phase ($R_{F/E}$) palpated by examiners (O) A and B at the examination-times (T) T 1 and T 2 at the examiner-locations (L) cranium (C) and pelvis (P).

The columns on the left half show the $\log_{10}$ transformed values ($\log_{10} R_{F/E}$), the right half shows the original mean values ($R_{F/E}$). For original values bigger than one, the flexion-phase is longer than the extension-phase and inversely.

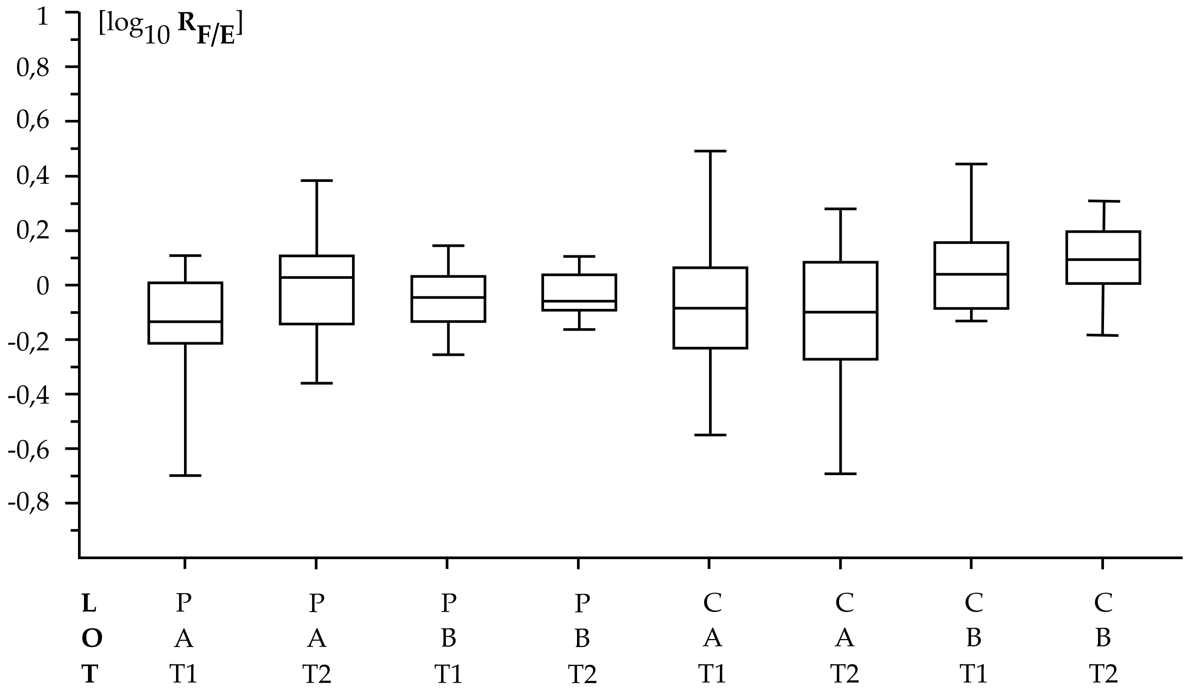


Figure 11
Boxplot-presentation of the values of palpated mean ratio of flexion- to extension-phase ($R_{F/E}$) for examiners (O) A and B at the examiner-locations (L) pelvis (P) and cranium (C) at the examination times (T) T 1 and T 2. For better representation the values are $\log_{10}$ transformed.

problem-orientated parameters, are presented. For *f* the smallest value was 1 cycle, the biggest 10 cycles per 90 seconds. By taking into account 95% of the values, 1 was the minimum, 8 the maximum. The difference therefore is 7 cycles. For MDF the smallest value was 3.1, the maximum 25 seconds. Within 95% of the values the minimum was 3.5 seconds, the maximum 19.8, which shows a range of 16.2 seconds. For $R_{F/E}$ the minimum value was $\log_{10}$ –0.9, the maximum $\log_{10}$ 1. Within 95% of the values the minimum was $\log_{10}$ –0.7, the maximum $\log_{10}$ 0.5, which produces a range of $\log_{10}$ 1.2. The values are shown in Table 8.

Parameters	within 100% of all measured values		within 95% of all measured values	
	min	max	min	max
f [cycles in 90 sec]	1.00	10.00	1.00	8.00
MDF [sec]	3.1	25	3.5	19.8
$R_{F/E}$	–0.9* (0.1)**	1.0 (10.0)	–0.7 (0.2)	0.5 (3.2)

Table 8
Summary of the minimal and maximal values that have bee found under the various conditions (examiner, location, time) for the three parameters PRM-frequency (*f*), mean duration of the flexion-phase (MDF) and mean ratio flexion- to extension-phase ($R_{F/E}$). The columns at the left take all measured values into account, the columns at the right 95%. The values for *f* represent cycles per 90 seconds, the values for MDF are seconds and the values for $R_{F/E}$ are ratio-values in
* $\log_{10}$ transformation and
** retransformed (in brackets).

7.1.5 Analysis of variance

In the model for the analysis of variance the three factors examiner (O), location (L) and time (T) have been tested for systematic differences and possible interactions between O and L for each dependent variable *f*, MDF and $R_{F/E}$.

For better understanding, a short definition of the presented values in the analysis of variance tables is given. The degrees of freedom (DF) represent the number of possibilities or levels per factor minus one. The sum of squares (SS) shows the total variability of the data. The mean square (MS) is the mean variance of the separate variances within and between groups. The F-distribution (F) shows the expected comparison for between group to within group mean variances covering the null hypothesis that all groups bear the same mean (Altmann 1999, pp 207–209; Bortz 1999, pp 237–319).

V.2.1. PRM-frequency (*f*)

The PRM-rates palpated at the pelvis produced mean values of 3.1 ± 0.9 cycles/min. At the head 2.9 ± 1 cycles/min have been found. Examiner A palpated a mean frequency of 2.7 ± 1 cycles/min, examiner B 3.3 ± 0.9 cycles/min. At examination-time T 1 the mean frequency was 3 ± 1 cycles/min, at T 2 2.9 ± 0.9 cycles/min.

The biggest influence on *f* could be found for the examiner ($P < 0.0001$), for the location it was less ($P = 0.19$). The least influence could be described for the examination-time ($P = 0.4$). Details can be seen in Table 9 and Table 10.

Factor	DF	SS	MS	F	P
L	1	2.5	2.5	1.74	0.18890
O	1	36.0	36.0	25.41	**< 0.0001**
T	1	1.0	1.0	0.71	0.40220
Sub	48	155.5	3.2	2.29	< 0.0001

Table 9

Factor	N	Mean			SD		
		90 sec	60 sec	Hz	90 sec	60 sec	Hz
L							
P	98	4.6	3.1	0.05	1.4	0.9	0.02
C	98	4.4	2.9	0.05	1.6	1.0	0.02
O							
A	98	4.1	2.7	0.05	1.6	1.0	0.02
B	98	4.9	3.3	0.05	1.4	0.9	0.02
T							
T 1	98	4.6	3.0	0.05	1.6	1.0	0.02
T 2	98	4.4	2.9	0.05	1.3	0.9	0.01

Table 10

Results of the four way anova for the factors location (L), examiner (O) and time (T) at the respective levels pelvis (P), cranium (C), examiner A, examiner B, time T 1 and time T 2 in relation to the independent variable *f*.

Table 9 shows the analysis of variance table. This table shows the PRM-frequencies separate for each group in cycles per 90 seconds, per minute and Hertz.

Sub … subjects

N … number of measurements

The test for possible interactions between examiner (O) and location (L) was significant ($P < 0.0001$). So the respective effects have been tested separately for pelvis (P) and cranium (C). The analysis showed that the means for A and B are about the same. However, this does not mean that they agree. At the head A systematically palpated lower PRM-rates with mean 2.4 ± 1 cycles/min as B with mean 3.5 ± 0.7 cycles/min. For details see Table 11 and Table 12.

Factor	DF	SS	MS	F	P
L x O	1	31.8	31.8	26.45	**< 0.0001**

Table 11
Results for the analysis of interactions (x) between the factors location (L) and examiner (O) in relation to the dependent variable *f*.

Factor L	Factor O	N	Mean 90 sec	Mean 60 sec	Mean Hz	SD 90 sec	SD 60 sec	SD Hz
P	A	49	4.6	3.0	0.05	1.5	1.0	0.023
P	B	49	4.6	3.1	0.05	1.1	0.8	0.01
C	**A**	49	**3.6**	**2.4**	**0.04**	1.4	1.0	0.02
C	**B**	49	**5.2**	**3.5**	**0.06**	1.1	0.7	0.01

Table 12
Results of the separate tests for locations pelvis (P) and cranium (C) as well as examiners A and B. Significant interactions due to Table 11 are marked in bold numbers.

7.1.6 Mean duration of the flexion-phase (MDF)

Analysis of variance for the MDF showed mean values ranging from 9.1 to 9.8 seconds with a SD ranging from 2.4 to 4.5 seconds. The MDF was found to be lower at the pelvis than at the head. Details are shown in Table 13 and Table 14.

Factor	DF	SS	MS	F	P
L	1	21.8	21.8	2.14	0.14540
O	1	0.5	0.5	0.05	0.82990
T	1	10.6	10.6	1.04	0.30980
Sub	48	1020.3	21.3	2.09	0.00040

Table 13

Factor	N	Mean	SD
L			
P	98	9.1	3.5
C	98	9.8	3.7
O			
A	98	9.5	4.5
B	98	9.4	2.4
T			
T 1	98	9.2	3.9
T 2	98	9.7	3.3

Table 14
Results of the four way anova for the factors location (L), examiner (O) and time (T) at the respective levels pelvis (P), cranium (C), examiner A, examiner B, time T 1 and time T 2 in relation to the independent variable MDF. Table 13 shows the analysis of variance table. This table shows the mean duration of the flexion-phase (MDF) separate for each group in seconds.

The tests for possible interactions between L and O was significant with regard to MDF ($P < 0.0189$). So the respective effects have been tested separately for the locations P and C. The mean values were higher for examiner B with 9.6 ± 2.6 sec in relation to A with 8.6 ± 4.1 sec at the pelvis. At the head examiner A palpated with 10.4 ± 4.8 sec longer flexion-phases as B with 9.2 ± 2.1 seconds. For details see Table 15 and Table 16.

Factor	DF	SS	MS	F	P
L x O	1	55.6	55.6	5.64	**0.0189**

Table 15
Results for the analysis of interactions (x) between the factors location (L) and examiner (O) in relation to the dependent variable MDF.

Factor L	Factor O	N	Mean	SD
P	A	49	8.6	4.1
P	B	49	9.6	2.6
C	A	**48***	10.4	4.8
C	B	49	9.2	2.1

Table 16
Results of the separate tests for locations pelvis (P) and cranium (C) as well as examiners A and B.

* One measurement has been determined as statistical anomaly and was therefore excluded (point 7.1.2 on p 85ff).

7.1.7 Mean ratio of flexion- to extension-phase ($R_{F/E}$)

The analysis of variance for $R_{F/E}$ showed a systematic tendency for significant lower mean values for examiner A with $\log_{10} - 0.07 \pm 0.32$ as for B with $\log_{10} 0.03 \pm 0.16$. Systematic differences between examination times were low. The influence of the factor examiner on $R_{F/E}$ was just not significant (P = 0.005). See Table 17 and Table 18.

Factor	DF	SS	MS	F	P
L	1	0.07	0.07	1.12	0.29070
O	1	0.49	0.49	8.02	**0.00530**
T	1	0.07	0.07	1.08	0.29990
Sub	48	3.45	0.07	1.18	0.22710

Table 17

Factor	N	Mean	SD
L			
P	98	–0.04	0.21
C	98	–0.01	0.29
O			
A	98	–0.07	0.32
B	98	0.03	0.16
T			
T 1	98	–0.04	0.26
T 2	98	–0.01	0.25

Table 18
Results of the four way anova for the factors location (L), examiner (O) and time (T) at the respective levels pelvis (P), cranium (C), examiner A, examiner B, time T 1 and time T 2 in relation to the independent variable $R_{F/E}$.

Table 17 shows the analysis of variance table. This table shows the mean ratio of flexion- to extension-phase ($R_{F/E}$) separate for each group in $\log_{10}$ transformation.

The test for possible interactions between the factors O and L was also just below the limit for significance (P = 0.058). A separate analysis for the locations pelvis and cranium showed systematic differences between A and B at the pelvis. See Table 19 and Table 20 on p 93.

Factor	DF	SS	MS	F	P
L x O	1	0.22	0.22	3.64	**0.0583**

Table 19
Results for the analysis of interactions (x) between the factors location (L) and examiner (O) in relation to the dependent variable $R_{F/E.}$

Factor L	Factor O	N	Mean	SD
P	A	49	–0.59	0.26
P	B	49	–0.03	0.15
C	A	49	–0.09	0.37
C	B	49	0.08	0.16

Table 20
Results of the separate tests for locations pelvis (P) and cranium (C) as well as examiners A and B. Ratio values are shown in $\log_{10}$ transformation.

7.2 Analysis of inter- and intraexaminer agreement (reliability)

The description of inter- and intraexaminer agreement by using the differences between the measurements enables the representation of results by the following values:

- The *mean difference* (M_{Diff}) of the compared measures. A small mean difference need not indicate good agreement for individual comparisons.
- The *standard deviation of the differences between the two measurements* (SD). Only the SD can give an expected range, wherein differences for one individual comparison can be estimated with a 95% probability.
- This range is represented by the *95% limits of agreement* bounded by a lower (low) and upper (high) limit, indicating
- the *expected difference* ($\pm E_{Diff}$) which is the relative difference being equal in both directions from mean.

7.2.1 Analysis of interexaminer agreement

Interexaminer agreement has been tested at a time separately for the locations P and C as well as the times T 1 and T 2. In the first case unadjustment is present for location, in the second case for examination-time (for further details see point »How blind are studies in the field of manual assessment and therapy?« p 25ff and point »Analysis of inter- and intraexaminer agreement« p 81ff).

PRM-frequency (f)

The description of interrater agreement between A and B for repeated measures shows 95% limits of agreement with an expected difference (E_{Diff}) of ± 3.3 cycles / 90 seconds at the cranium (C) and the pelvis (P). For simultaneous palpation at different locations (C, P) the expected difference at T 1 is ± 3.1 cycles / 90 sec, at T 2 ± 2.8 cycles / 90 sec. Details can be seen in Table 21. A graphical presentation of all results concerning agreement can be seen in Fig. 12 on p 98.

Factor	N	M_{Diff} A – B	SD A – B	95% limits of agreement Low	High	E_{Diff}
C	49	–1.65	1.70	–4.99	1.68	3.3
P	49	–0.06	1.66	–3.32	3.20	3.3
T 1	49	–0.88	1.59	–3.99	2.24	3.1
T 2	49	–0.84	1.45	3.68	2.00	2.8

Table 21
Overview of the results for interexaminer agreement for the palpated PRM-rates (*f*), described by the differences between the locations cranium (C) and pelvis (P) as well as between the examination-times T 1 and T 2. The values show the number of cycles/90 seconds. The values in the last column (E_{Diff}) go for both directions (±) from the mean.

Taking the total variance of values into account, the expected difference for interexaminer agreement is quite as large. Within 95% of all values the difference was 7 cycles (see point 7.1.4 on p 86ff and Table 8 on p 88). The expected difference is 5.6 to 6.6 cycles. Thus for the parameter *f* no agreement beyond chance agreement can be described. Comparing the individual data it might be of interest that the difference is smaller at T 2, when the examiners palpated the same subject the second time.

Mean duration of the flexion-phase (MDF)

For the MDF the expected difference between A and B within the 95% limits of agreement has been found to be ± 9.8 seconds at the head (C) and ± 8.4 seconds at the pelvis (P). At examination-time T 1 the difference was found to be ± 10.3 sec, at T 2 8.4 seconds. The results are summarised in Table 22. A graphical presentation of all 95% limits of agreement can be seen in Fig. 12 on p 98.

Factor	N	M_{Diff} A – B	SD A – B	95% limits of agreement Low	High	E_{Diff}
C	49	0.12	5.02	–8.71	10.95	9.8
P	49	–0.93	4.28	–9.33	7.46	8.4
T 1	49	–0.17	5.26	–10.48	10.13	10.3
T 2	49	0.33	4.27	–8.04	8.70	8.4

Table 22
Overview for the results describing interexaminer agreement of palpated mean duration for the flexion-phase (MDF) by using the differences between the locations head (C) and pelvis (P) as well as between the examination-times T 1 and T 2. The units of the values are seconds. The values in the last column (E_{Diff}) go for both directions (±) from the mean.

The interpretation of interexaminer agreement for the parameter MDF is similar as for *f*. Referring to the 95% of the total range of measured values, MDF shows a difference of 16.3 seconds (see point 7.1.4 on p 86ff and Table 8 on p 88). The 95% expected difference for interexaminer agreement produces values from 16.8 to 20.6 seconds. Agreement beyond chance agreement is neither predictable in this case. As for *f* a decrease of the difference from T 1 to T 2 can be observed.

Mean ratio of flexion- to extension-phase ($R_{F/E}$)

Concerning $R_{F/E}$ an expected interexaminer difference within the 95% limits of agreement showing ± $\log_{10}$ 0.8 at the head and ± $\log_{10}$ 0.5 at the pelvis has been found. The difference was the same at both examination-times with ± $\log_{10}$ 0.7. The data are shown in Table 23. A graphical presentation of all 95% limits of agreement can be seen in Fig. 12 on p 98.

The difference within 95% of the total range of ascertained values for $R_{F/E}$ was $\log_{10} 1.2$ (see point 7.1.4 on p 86ff and Table 8 on p 88). The 95% range of agreement is $\log_{10} 1$ at the head and $\log_{10} 1.6$ at the pelvis. So the situation is similar as for the other parameters (*f*, MDF). The expected range for interexaminer agreement is quite as large as the total range of values. So no interexaminer agreement can be described for $R_{F/E}$ beyond chance agreement.

Factor	N	M_{Diff} A – B	SD A – B	95% limits of agreement Low	High	E_{Diff}
C	49	–0.17	0.41	–0.97	0.63	0.8
P	49	–0.03	0.28	–0.57	0.51	0.5
T 1	49	–0.12	0.37	–0.84	0.60	0.7
T 2	49	–0.08	0.35	–0.77	0.61	0.7

Table 23
Overview for the results describing interexaminer agreement of palpated mean ratio of flexion- to extension-phase ($R_{F/E}$) by using differences between the locations head (C) and pelvis (P) as well as between the examination-times T 1 and T 2. The values are $\log_{10}$ transformed ratio values. The values in the last column (E_{Diff}) go for both directions (±) from the mean.

Summary

For all three problem-orientated parameters *f*, MDF and $R_{F/E}$ the 95% limits of agreement for interexaminer agreement are quite as large as the total range of ascertained values. So no interexaminer agreement can be described beyond chance agreement.

7.2.2 Analysis of intraexaminer agreement

To analyse intrarater agreement or consistency, data for the respective examiner were compared within the same subject between the examination-times T 1 and T 2 and the examiner-locations head (C) and pelvis (P). This analysis is therefore unadjusted by time and location. As pointed out in point »Final experimental setting and procedure« (see p 75ff), there was no time to analyse the data from the 3rd measures (T 3) that have been taken from 18 subjects and 4th measures (T 4) that have been taken from 10 subjects, where unadjustment of location would have been avoided.

The expected intrasubjective difference within the 95% limits of agreement is with regard to f ± 3.0 cycles / 90 min for examiner A and ± 2.7 cycles / 90 min for B. For MDF a difference of ± 9.7 sec has been found for examiner A and ± 4.7 sec for B. For $R_{F/E}$ the difference was ± log 0.8 for A and ± log 0.4 for B. The results are summarised in Table 24. A graphical presentation of all 95% limits of agreement can be seen in Fig. 12 on p 98.

Examiner B showed smaller differences as A for all parameters, especially for MDF. The expected difference within the 95% limits of agreement is, with regard to f 6 cycles / 90 sec for A and 5.4 cycles / 90 sec for B, with regard to MDF, 19.4 sec for A and 9.8 seconds for B and finally with regard to $R_{F/E}$ 1.6 for A and 0.8 for B. So the expected differences are quite as large as 95% of the total range of ascertained values for the respective parameters, which is 6 cycles / 90 sec for f, 16.3 seconds for MDF and a ratio of log 1.2 for $R_{F/E}$. Hence no intraexaminer agreement can be described for all three parameters.

	Factor	N	M_{Diff} A – B	SD A – B	95% limits of agreement Low	High	E_{Diff}
f	A	49	1.02	1.55	–2.01	4.05	3.0
	B	49	–0.57	1.38	–3.28	2.14	2.7
MDF	A	49	–1.79	4.94	–11.48	7.90	9.7
	B	49	0.36	2.42	–4.39	5.11	4.7
$R_{F/E}$	A	49	0.03	0.42	–0.79	0.86	0.8
	B	49	–0.11	0.20	–0.50	0.29	0.4

Table 24
Overview for the results describing intraexaminer agreement for the parameters PRM-rate (f), mean duration of the flexion-phase (MDF) and mean ratio of flexion- to extension-phase ($R_{F/E}$) by using the differences within the examiners A and B. The values are respectively in cycles/90 sec for f, seconds for MDF and $\log_{10}$ transformed ratio values for $R_{F/E}$. The values in the last column (E_{Diff}) go for both directions (±) from the mean.

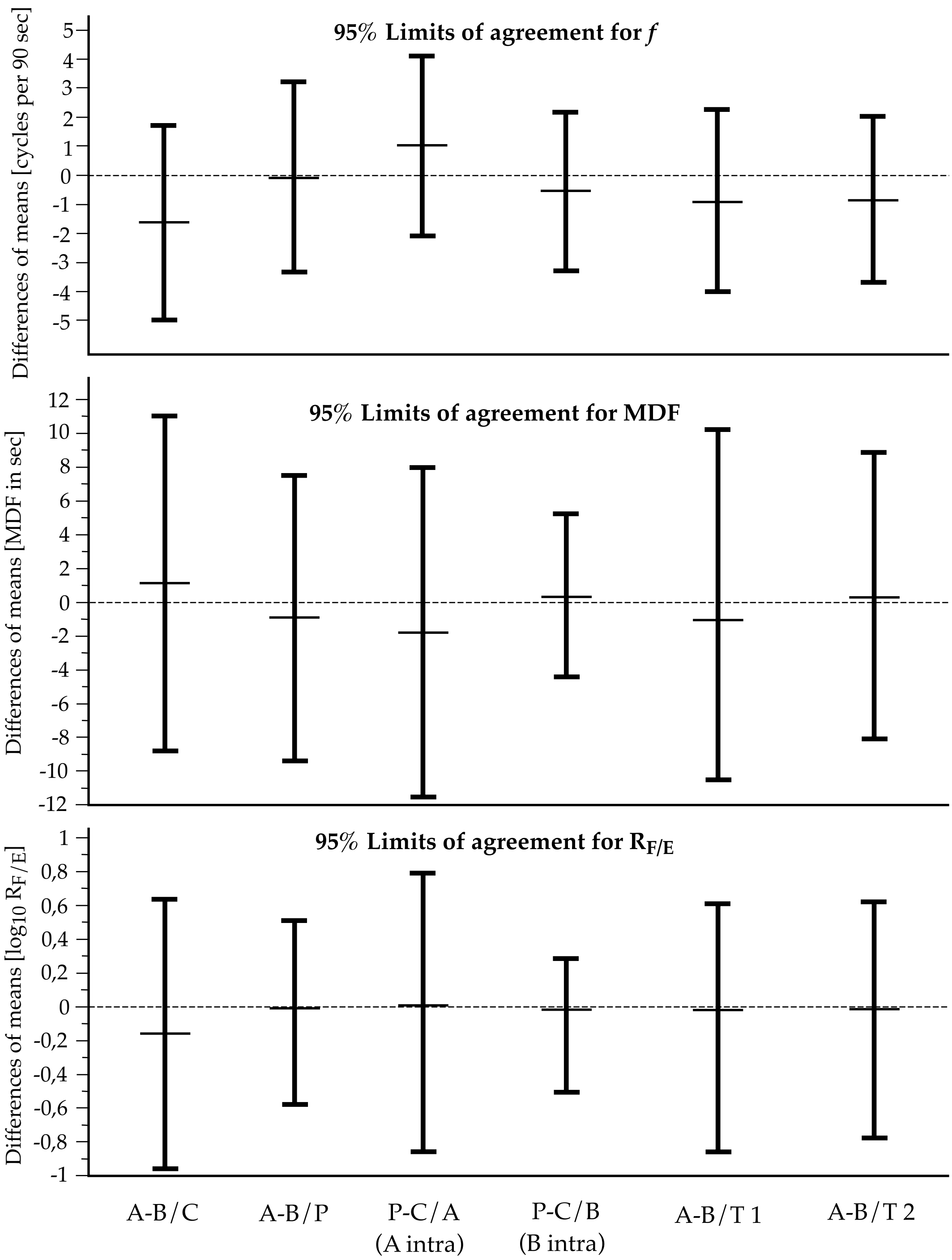


Figure 12

Plot of the 95% limits of agreement showing the *expected differences from mean* for interexaminer agreement at the locations head (A-B/C) and pelvis (A-B/P) and for the examination-times T 1 (A-B/T 1) and T 2 (A-B/T 2) as well as for intraexaminer agreement for examiner A (P-C/A) and B (P-C/B) at a time between head and pelvis and T 1 and T 2 for the same subject. The range above and underneath the mean differences (indicated by the longer lines dividing the total range into two equal halves) shows ± the inter- or intraexaminer difference that can be expected with a probability of 95% for one individual subject. The whole range represents: mean difference ± 1.96 SD of the differences.

7.3 Analysis of covariance and correlation between PRM and respiratory rates

7.3.1 Analysis of covariance

The analysis of possible effects of the respective examiner's respiratory rate (RE_O) and the subject's respiratory rate (RE_{SU}) on the palpated PRM showed the following results.

For both examiners a significant effect of the examiner's own respiratory rate could be observed at the examiner-location pelvis with $P = 0.004$ for examiner A (RE_A) and with $P < 0.0001$ for examiner B (RE_B) on the dependent variable *f*. At the head the effect was significant for examiner B only ($P = 0.0017$). No significant influence of the subject's respiratory rates (RE_{SU}) on *f* could be found. For details see Table 26 on p 100.

For the dependent variable MDF, like for *f*, a significant effect of the examiners' respiratory rates (RE_O) has been found at the examiner-location pelvis ($P = 0.0276$). For the head no significant effects could be observed. For details see Table 26 on p 100.

For the dependent variable $R_{F/E}$ no significant effects of the examiners' as well as the subjects' respiratory rates could be found. The results can be seen in Table 25.

Factor	DF	SS	MS	F	P
$RE_{SU} \times O$	1	0.006	0.006	0.10	0.7525
$RE_O \times O$	1	0.007	0.007	0.11	0.7354
$O \times L$	1	0.163	0.163	2.71	0.1017
$RE_O \times L$	1	0.008	0.008	0-13	0.7200
RE_{SU}	1	0.209	0.209	3.50	0.0634
RE_O	1	0.002	0.002	0.03	0.8710
O	1	0.426	0.426	7.12	0.0085
L	1	0.696	0.696	1.16	0.2824

Table 25
Results of the analysis of covariance for the dependent variable mean ratio of flexion- to extension-phase ($R_{F/E}$). Possible significant effects of the respective examiner's (RE_O) and subjects' (RE_{SU}) respiratory rates on $R_{F/E}$ have been tested under the influence of the factors examiner (O) and location (L).

Dependent variable	Factor	DF	SS	MS	F	P
	$RE_{SU} \times O$	1	1.9	1.9	1.84	0.1775
	$RE_O \times O$	1	0.6	0.6	0.66	0.4185
	$O \times L$	1	44.0	44.0	42.23	**< 0.0001**
	$RE_O \times L$	1	15.8	16.8	15.13	**0.0002**
	Examiner A at the pelvis					
	RE_{SU}	1	2.5	2.5	1.28	0.2642
	RE_A	1	18.2	18.2	9.23	**0.0039**
f	**Examiner B at the pelvis**					
	RE_{SU}	1	2.7	2.7	3.18	0.0812
	RE_B	1	24.0	24.0	28.11	**< 0.0001**
	Examiner A at the head					
	RE_{SU}	1	0.6	0.6	0.27	0.6036
	RE_A	1	1.5	1.5	0.74	0.3646
	Examiner B at the head					
	RE_{SU}	1	0.3	0.3	0.33	0.5706
	RE_B	1	10.4	10.4	11.16	**0.0017**
	$RE_{SU} \times O$	1	9.6	9.6	0.97	0.3260
	$RE_O \times O$	1	18.4	18.4	1.86	0.1743
	$O \times L$	1	51.9	51.9	5.27	**0.0232**
	At the pelvis					
	RE_{SU}	1	1.6	1.6	0.20	0.660
MDF	RE_O	1	42.7	42.7	5.18	**0.028**
	O	1	54.0	54.0	6.55	**0.014**
	At the head					
	RE_{SU}	1	0.2	0.2	0.01	0.907
	RE_O	1	0.2	0.2	0.02	0.892
	O	1	22.9	22.9	1.75	0.193

Table 26
Results of the analysis of covariance for the dependent variables PRM-frequency (*f*) and mean duration of flexion (MDF). Possible significant effects of the respective examiner's (RE_O respectively RE_A and RE_B) and subjects' (RE_{SU}) respiratory rates on *f* and MDF have been tested under the influence of the factors examiner (O) and location (L).

For *f* the interaction of the respective examiners A and B for the locations pelvis (P) and head (C) has been tested separately. For MDF the interaction of both examiners for the locations pelvis (P) and head (C) has been tested separately.

Significant values are marked in bold letters.

7.3.2 Analysis of correlation

To test the strength of the effects described in the analysis of covariance, the analysis of correlation was carried out. The results for the respective subgroups have been calculated separately for examiners and examiner-locations. Person's correlation coefficient (r) and P-values (P) have been used to describe the strength of correlation.

At the examiner-location pelvis a significant correlation between the examiners' respiratory rates (RE_O) and the PRM-frequency (f) could be described with $r = 0.42$ ($P = 0.0024$) for examiner A and $r = 0.58$ ($P < 0.0001$) for B. For examiner B significant correlation with regard to MDF could be observed with $r = -0.55$ ($P < 0.0001$). The negative coefficient indicates, that the examiner palpated shorter flexion-phases when her respiratory rate increased. All other correlations for the location pelvis produced no significant values. At the examiner-location head correlation could be described as significant only for examiner B with regard to f and MDF with $r = 0.45$ ($P = 0.0012$) for f and $r = -0.57$ ($P < 0.0001$) for MDF. Details can be seen in Table 27 on p 102.

7.3.3 Summary

The results indicate a significant tendency for the examiners' respiratory rates to have an effect on f and MDF especially at the pelvis. This means that the examiners tend to palpate higher PRM-frequencies and shorter lengths of the flexion-phase when their own respiratory rate increases and inversely. At the head the results for the examiners differ. Whereas for one examiner an effect can be described, for the other one it is not present. No influences of the subject's respiratory rate on the PRM could be observed at any time. For the $R_{F/E}$ no effects can be described at all.

Parameter	RE_{SU}		RE_O	
	r	P	*r*	P
		Examiner A at the pelvis		
f	0.21	0.1574	**0.24**	**0.0024**
MDF	–0.09	0.5207	–0.21	0.1497
$\log_{10} R_{F/E}$	–0.16	0.2862	–0.11	0.4714
		Examiner B at the pelvis		
f	0.03	0.8540	**0.58**	**0.0001**
MDF	–0.04	0.7969	**–0.55**	**0.0001**
$\log_{10} R_{F/E}$	–0.17	0.2517	–0.32	0.0243
		Examiner A at the head		
f	0.04	0.7689	–0.11	0.4577
MDF	–0.13	0.3681	0.18	0.2324
$\log_{10} R_{F/E}$	–0.05	0.7343	–0.04	0.7970
		Examiner B at the head		
f	0.13	0.3760	**0.45**	**0.0012**
MDF	–0.16	0.2700	**–0.57**	**0.0001**
$\log_{10} R_{F/E}$	–0.12	0.4002	–0.15	0.3171

Table 27
Results of the analysis of correlation in four subgroups between the problem-orientated parameters PRM-frequency (*f*), mean duration of the flexion-phase (MDF), mean ratio of flexion- to extension-phase ($R_{F/E}$) and the covariables respiratory rates of the respective examiner (RE_O) and subjects (RE_{SU}). Significant values are marked in bold letters.

r... Pearson's correlation coefficient

P...P-value

8 Discussion

8.1 Methodological limitations

Every conclusion that can be derived from the presented results has to include possible errors and actual deficiencies due to methodological aspects such as research design measuring devices and procedures as well as data analysis, concerning this thesis.

8.1.1 Limitations concerning examiners

The results of interexaminer agreement testing have to be interpreted under the condition that the number of examiners was minimal (2 examiners) for conducting interexaminer reliability testing. To have at least four or eight examiners or different groups with different degrees of experience would have been more appropriate to make the results more meaningful for the specific population of examiners. In order to make the measurements comparable, all examiners had to assess the same 49 subjects. However, the conditions under which this project has been realised, allowed just two examiners. But as the probability of treatment effects grows with increasing number of examiners (i.e. repetition of measurements), it was better to stick to two. Furthermore analysis of interexaminer agreement is more meaningful using a bigger sample size of rated subjects. The examiners' degree of experience can be regarded as sufficient, as they have passed an amount of training that is regarded as sufficient by the guidelines for osteopathic education in Europe.

8.1.2 Limitations concerning the subjects

The subjects were analytically considered as random factors (see point 6.5.2 on p 80ff). Prominent pathologies concerning the cranium, the spine and the pelvis were excluded, the possibility of dysfunction was included in order to reach sufficient variance among the subjects (see point 6.2 on p 68ff). The description of inter- and intraexaminer agreement for symptomatic subjects therefore was not possible in our case. Nevertheless inter- and/or intraexaminer agreement might be higher for symptomatic subjects than for normal subjects. Two groups of samples, comparing symptomatic and asymptomatic subjects in randomised appearance, like Schöps et al. (2000) did, might have brought a more specific information on interexaminer agreement under clinical conditions.

8.1.3 Limitations concerning palpation and its recording

Measuring the perception of the PRM by using a foot switch can produce erroneous values caused by difficulties in simultaneously palpating and activating the switch. We tried to minimise this effect by introducing the use of the foot switches in a preliminary test where the examiners could check the signals produced by the switch while looking at the monitor during palpation. The amount of presented measurements ($n = 98$) and the fact that the spread of the final values concerning palpated data (*f*, MDF, $R_{F/E}$) were between acceptable limits (e.g. 0.7–5.3 cycles per minute for f) could in addition support the assumption that possible errors are small. The examiners themselves reported neither problems in using the foot switches nor any subjective interference to quality of palpation. In this context additional testing of interrater reliability for thoraco-abdominal respiration palpation, like Wirth-Pattullo and Hayes (1994) and Norton (1996) did, would have enhanced internal validity of the study.

Furthermore the examiners did not work under normal and relaxed clinical conditions. The awareness, that their findings were recorded, might have produced mental stress that could have influenced the concentration needed for adequate palpation. Further circumstances that could have interfered the examiners' palpatory qualities were the fact that they should only assess the subjects even when palpatory findings seemed to call for treatment, that they had the respiration-transducers mounted around their thorax, that the patient was „divided" by the curtain or that sequences of mere testing followed one after the other. All these factors can be highly relevant as influences on the palpation of the PRM. In order to achieve the best examination-conditions possible the examination room was kept as quiet as possible, the examiners could use their own hand holds, had time to orientate themselves before the recording started and breaks of at least 15 minutes after every sixth subject. The fixations of the respiration-transducers has always been corrected as soon as they produced any unpleasant effects.

8.1.4 Limitations concerning the interpretation of inter- and intraexaminer agreement

Regarding interexaminer agreement there has always been the problem of unadjustment either by time or by location (see also point »How blind are studies in the field of manual assessment and therapy?« p 25ff and point 6.5.3 on p 81ff), which means that the compared measurements were not taken either simultaneously (1st measurement or 2nd measurement) and/or at the same location (cranium or pelvis). Any interpretation of the results has to consider, that the conditions under which the compared measures were produced, were not the same. This essential problem is inherent to reliability study designs (see point »Reproducibility« p 10ff) and the question rests whether unadjustment especially for the factor location can be ruled out at all in inter- and intraexaminer reliability studies for manual findings. In the presented case this problem comes especially true for the analysis of intraexaminer agreement, where unadjustment for both, location and time is given.

8.1.5 Limitations concerning the recording of the respiratory rates

For the recordings of the respiratory rates any effects of the respiration-transducers on the respiratory rates of the subjects and examiners cannot be ruled out. As the belt had to be tightened to a certain extent in order to obtain appropriate data, respiration was certainly affected as it had to be exerted against a certain resistance. For the examiners this is especially true, as the belts had to be fastened tight enough to prevented them from gliding downwards during the expiration phase when the thorax is narrowing.

A further problem was the accuracy of the measuring module, which could not handle a resolution of 1 signal per 100 milliseconds. As already mentioned in point 6.3.1 on p 70ff, due to technical limitations a resolution of 500 ms had to be used, which for time domain analyses of the respiratory rates is estimated to be too rough. As time- and frequency domain analysis did not go into subtle details, we hope that the measurement error, produced by this factor is not too dominant. However, these circumstances have to be taken into account interpreting the results of the analysis of covariance and correlation between the main outcome variables *f*, MDF and $R_{F/E}$ and the covariables RE_O and RE_{SU} presented in point 6.5.4 on p 82ff.

8.2 Comparing the results

8.2.1 PRM data

The PRM-frequencies measured in this study range from 0.7 to 6.7 cycles/min (0.1 to 0.11 Hz), described by mean values from 2.3 ± 0.01 to 3.6 ± 0.8 cycles/min. The results are similar to recent publications (Norton et al. 1992; Wirth-Pattullo and Hayes 1994; Norton 1996; Hanten et al. 1998; Rogers et al. 1998) and differ to elder publications (Upledger, 1977) who describe a range between 6 to 12 cycles/min (0.1 to 0.2 Hz). But even in comparison with recent results the PRM-frequency range in this study was quite low and agrees most with the values reported by Norton et al. (1992) with a mean value of 3.7 ± 0.6 cycles/min, Norton (1996) with a range from 2.14 to 7.38 cycles/min and Hanten et al. (1998) with a range of mean values from 3.5 ± 0.9 to 4.3 ± 1.2 cycles/min. For a summary of the different results see Table 28.

Publication	reported cycles/min
Upledger (1977)	8–16
Norton et al. (1992)	3.7 ± 0.6*
Wirth-Patullo and Hayes (1994)	3–9
Norton (1996)	2.14–7.38
Hanten et al. (1998)	3.5 ± 0.9* to 4.3 ± 1.2*
Rogers et al. (1998)	0–8.42
Values of this study	2.3 ± 0.01* to 3.6 ± 0.8*

Table 28
Summary of reported data on cranio sacral rhythm palpation in comparison with the presented values. The values represent minimal and maximal findings.
* These values indicate mean values ± standard deviation.

Several authors have suggested to explain the PRM as an expression of variations of known physiological rhythms like heart rate and respiratory rate (Frymann 1971, Upledger 1977, McPartland and Mein 1997). Comparing the results with the frequency spectrum of heart rate variabilities (Camm et al. 1996), with mean values ranging from 0.04 ± 0.02 to 0.06 ± 0.01 Hz, they cover the lower part of the low frequency spectrum (LF), known as Mayer waves which ranges from 0.04 to 0.15 Hz and seems to be expressed by changes in blood pressure (see also point 3.5 on p 54ff and Table 2 on p 54). Recent experiments of Nelson et al. (2001, 2006) and Sergueef et al. (2002) are supporting this hypothesis.

Similar to Norton et al. (1992) and Norton (1996) who besides the PRM-frequency (f) described the whole cycle length as well as the length of the extension and the flexion phase, in this study the mean duration of the flexion phase (MDF) and the mean ratio between the length of the flexion- and the extension phase ($R_{F/E}$) have been described. Norton et al. (1992) report a mean duration of the flexion phase with 7.7 ± 1.4 sec, the MDF found in this study is longer and ranges from 8 ± 4.3 to 10.6 ± 4.1 sec.

8.2.2 Inter- and intraexaminer agreement

Wirth-Pattullo and Hayes (1994), Norton (1996), Hanten et al. (1998) and Rogers et al. (1998) report low or missing interrater reliability for the palpation of the PRM. For analysing interexaminer reliability *pearson's correlation coefficients* (Norton, 1996) and *intraclass correlation coefficients* (Wirth-Pattullo and Hayes, 1994; Hanten et al., 1998; Rogers et al., 1998) were used (see point 4 on p 58ff and Table 29). As explained in point 6.5.3 on p 81ff, we decided to describe inter- and intraexaminer agreement by using the *95% limits of agreement* (Bland and Altman, 1986).

Nevertheless the results for interrater agreement were the same. In contrast to the other authors, we described agreement besides for the palpated PRM-frequency (f) also for the MDF and the $R_{F/E}$ and used a bigger sample size (n = 49). In neither cases interexaminer agreement could be described beyond chance agreement.

For intraexaminer agreement the results differ. Norton (1996) reports significant intrarater correlation for the cycle lengths palpated by the same examiner at the head and at the pelvis ($r = 0.926$, $P < 0.001$). Hanten et al. (1998) found intrarater reliability to be good (ICC = 0.78 and 0.83) for palpation of the PRM-rates. We found, like Rogers et al.

Publication	Interexaminer reliability	Intraexaminer reliability
Wirth-Patullo and Hayes (1994)	ICC = –0.02; P = 0.0001	
Norton (1996)	head: r = 0.275; P = 0.115	r = 0.926; P < 0.001
	pelvis: r = –0.296; P = 0.089	
Hanten et al. (1998)	ICC = 0.22	ICC = 0.78 and 0.83
Rogers et al. (1998)	head: ICC = 0.08; feet: ICC = 0.19	head: ICC = 0.18 and 0.26
	head: r = 0.12; feet: r = 0.23	head: r = 0.17 and 0.27
		feet: ICC = 0.30 and 0.29
		feet: r = 0.30 and 0.29

Table 29
Results of published data on inter- and intraexaminer reliability for the palpation of the PRM described by the intraclass correlation coefficient (ICC) and the pearson's correlation coefficient (r). The values are presented due to the methods used for two different locations and two different examiners.
r can range from 1 to –1, indicating direct or inverse correlation when reaching 1 or –1
ICC can range from 0 to 1, indicating high correlation when reaching 1

(1998) (head: ICC = 0.18 and 0.26; feet: ICC = 0.3 and 0.29), intraexaminer reliability to be low with an expected difference of ± 3.0 cycles/90 sec for examiner A and ± 2.7 cycles/90 sec for examiner B. For a summary of published data for inter- and intraexaminer reliability of the palpation of the PRM see Table 29 on p 107.

8.2.3 Correlation of the PRM to respiratory rates

Upledger (1977) measured the subjects' respiratory rates before the PRM-palpation. Wirth-Patullo and Hayes (1994) measured the subjects' and the examiners' respiratory and heart rates before, the subjects' rates also after the assessment of the PRM. Hanten et al. (1998) measured the subjects' heart- and respiratory rates simultaneously during the assessment of the PRM. All authors agree in finding low correlation between the palpated PRM-rate and the measured subjects' respiratory as well as cardiac rates. In this study the respiratory rates of subjects and examiners were recorded simultaneously during the assessment of the PRM as described above. Cardiac rates have not been measured. AS the other authors we did not find significant correlations with the subjects' respiratory rates.

Surprisingly enough there was significant correlation between the examiners' respiratory rates and the palpated PRM-rates for both examiners at the pelvis and for one examiner at the head (for detailed data see point 7.3.2 on p 101ff). These results are interesting according to the tissue pressure model (Norton, 1991) which suggests that variations in the examiner's respiratory and cardiac rates have the biggest influence on the resulting frequency calculated by the model (see also point 3.2.3 on p 43ff).

9 Conclusion

Considering that neither inter- nor intraexaminer agreement could be described for the palpation of the PRM in this project and results from other publications are similar – and taking into account the limitations discussed before – we come up to discussion suggesting the following items:

- *Palpation aiming for the patient's PRM is based on the examiner's imagination*. If this is so, the use of palpatory findings, concerning the PRM, should be thought over as means for clinical decision making. The presumption seems possible as certain influences of the examiners' own respiratory rates on the palpatory findings could be shown. Besides, as we set out earlier, the physiological existence of the PRM cannot be regarded as proven (Green et al. 1999; Downey, 2004; Hartman, 2006). Echternach (1994) states in this context that under normal conditions clinicians do not try to measure a phenomenon whose existence is unclear.
- *The PRM is a phenomenon which is too subtle in order to be palpated reliably*. This contradicts the fundamentals and the development of the Cranial Concept. In the beginning the PRM has solely been manually detected (see point 3.2 on p 41ff). To prove this assumption, the changes induced by the PRM should be measured by valid means (which is still controversial; see point 3.3 on p 46ff) and tests on manual perceptive possibilities should suggest that the threshold for such a perception is above the measured changes for the PRM. Roppel et al. (1978) report a threshold from about 0.5–0.25 mm, whereas early mechanical recordings (Frymann, 1971) showed amplitudes from 0.012 to 0.025 mm.
- *The PRM is a metaphysical, not a physiological concept*. This discussion could be carried on under a philosophical perspective which could be an interesting contribution. The frequent use of metaphoric terms like *breath of life* or the *tide* instead of the PRM which were already introduced by Sutherland himself (Sutherland 1990, 1998) suggest such an interpretation. So the Cranial Concept can be seen as a vitalistic concept. If this is so, the anatomical and physiological models explaining the PRM should be dropped. Otherwise we are facing the problems of physicotheological argumentations: the explanation of natural causality (physics) by theological arguments (supernatural forces as the *vis mediatrix naturae*).
- *The PRM is the result of the interaction between known physiological rhythms of the examiner and the subject*. In this context Norton (1991) found that within the *tissue pressure model* (see point 3.2.3 on p 43ff) simulated variations especially of the examiner's

heart- and respiratory frequencies had the strongest influence on changes of the resulting frequency. This, so Norton, could explain low interexaminer agreement. The frequency range of the presented results could also suggest possible explanations via heart rate variabilities (point 3.4 on p 51ff). Recent experiments by Nelson et al. (2001) and Sergueef et al. (2002) are supporting this hypothesis in the way that the PRM is correlated to Traube-Hering-Mayer-waves.

- *The examiners are not experienced enough*. For the present project and the publications by Norton (1996) and Rogers et al. (1998), this argument is not justified.
- *Other methodological limitations and deficiencies* (as mentioned in point 4 on p 58ff and point 8.1.4 on p 104ff) *were too dominant*. It seems quite improbable that definite results (especially concerning interexaminer reliability), presented by different research projects, that are dealing with a similar question and are following various methodological strategies as well as introducing different possibilities of describing inter- and intraexaminer reliability and agreement of the same phenomenon, should all be erroneous due to methodological deficiencies.

However, the presented results do not support theories behind the PRM, that call for an autonomous rhythmical event which is manually detectable. The results imply that the PRM cannot be palpated consistently among different examiners as well as within one examiner and under certain conditions the examiner's respiratory rates seem to have a distinct influence on what the examiner thinks to perceive as the PRM. In the consequence the produced data do not support the hypothesis proposed in point 5 on p 65ff. What the examiner actually perceives, when she/he palpates the phenomenon, described as the PRM, remains unclear.

The design used in this project can formally be seen as an observation of observers who observe a specific phenomenon, which in this case is the PRM. We can define an observer of 1st order who is the examiner and an observer of 2nd order who is the researcher (Von Foerster 1997, p123–126). Being an observation of 2nd order the results of this project do not at all refer to the question whether the PRM does exist or not. They only refer on the reliability of a specific observation, where observers try to identify a phenomenon called the PRM. So the phenomenon of identification could also be a unicorn. This would not influence the general methodic approach of this study. We regard this worth mentioning as the existence of the PRM is subjected to controversial debates, as pointed out in point 3.3 on p 46ff. In general it can be assumed that for the observer of 1st order (examiner) the objective existence of the observed phenomenon is beyond doubt. For the observer of 2nd order (researcher) this need not be the case. Criticism against interexaminer reliability studies on cranio sacral rhythm palpation (as Quaid, 1995) tends to misinterpret the problem by levelling the difference between 1st and 2nd order observation. By the way the dubiety of phenomena underlying outcome measures seems to be a characteristic element in quantitative assessments of complementary or alternative medicine (Jeffrey et al., 1997).

If the PRM should further be supported by a physiological model, new paths of explanation should seriously be considered. We suggest to keep to known physiological rhythms like heart and respiratory rates and phenomena that are related to them like heart rate variabilities (see point 3.5 on p 54ff). What now is identified as the PRM, might be a certain combination of these rhythmic phenomena, coming from the examiner and the patient, which is eventually helpful to perceive certain processes of release that occur during treatment. In this case hardly any interexaminer agreement can be achieved. The palpated frequency-spectrum resembled the LF-spectrum of heart rate variabilities, which expresses itself in cyclic changes of blood pressure. A combination of respiratory and cardiac rates and continuous blood pressure measurements from subjects and examiners seems useful in order to gain more specific information about what could be the physiological substratum of the examiners' perception. This assumption is supported by the experiments of Nelson et al. (2001, 2006) and Sergueef (2002). Mathematically extended time- and frequency-domain analyses might be helpful in this context.

Another aspect can be seen in the complex interaction between subject (patient) and examiner. Phenomena like transmission and counter-transmission, which are known from psychology, can be regarded as relevant in our case. As human beings are observing human beings, the subject's influence on the examiner and vice versa has to be taken into account. So changes of methodological aspects, regarding the system subject-examiner as a constantly changing network of complex interaction should be considered. Commonly used inter- and intrarater reliability study designs may not fit the complex situation in manual medicine, as they originate from an experimental background that is based on the principles of natural sciences (see point 2.1.1 on p 7ff). In this context we must admit that most reliability studies are only focusing on single tests. This does not reflect clinical reality. In general examiners gather a lot of information which cannot be brought down to a single test. So testing reliability should do justice to the whole clinical situation. For the future, methodological advice from psychologists or sociologists may be helpful. Qualitative and transdisciplinary approaches have to be considered as they are focusing more on the subject. Topics like responsibility and ethical implications that go along with the problem of reliability (see point »Reliability« p 11ff) might be addressed by these study types. Regarding the relationship between osteopath and patient within the clinical intervention these perspectives might be more relevant for reliability than scores indicating a sort of mathematical reliability.

Furthermore we want to emphasise that lack of inter- and intrarater agreement concerning the palpation of the PRM does not imply that the Cranial Concept has no clinical relevance. This has to be proven with standardised clinical trials. However, as far as we know this has still to come. Elder clinical projects (Frymann 1966, Woods and Woods 1961) do not match nowadays standards. One recent study by Mills et al. (2003) does show promising results but there is still a lot to do.

Reliability rests an important issue for osteopaths regarding clinical practice and training. What we tried by experiment was solving a scientific problem of mathematical measurement-reliability that has been imposed on a genuine osteopathic manual assessment procedure. But in addition we wanted to show by analytical (philosophical) means that we are also facing an epistemological problem which finally turns out to be an ethical question. As being clinicians we are always facing a fellow human being who is more or less trusting in what we are doing. What we are actually going to do should therefore be the result of reasonable deliberation and decision. So we are understandably thirsting for certitude and doubtlessness, claiming for objectivity and determinability. But all that has been unmasked as the old Platonian dream of metaphysics. So reliability does not end up as pure countable certitude. The results of scientific reliability-testing can give certain support for clinical osteopathic acting. But, doing justice to the clinical situation, the concept of reliability should be developed further as an approach that is pointing beyond measurability. So reliability might mean relying on the fundamental difference that will show up with every new patient. It might indicate repetition, but in the sense of *iterability*: repetition of the different. This might be the only way we can cope with the uniqueness of the whole person that meets our forever limited capabilities and knowledge.

10 Summary

Interexaminer reliability can be regarded as an essential problem for manual methods of assessment and treatment. The concept of reliability encompasses (1) a logical perspective (consistency), (2) a perspective of nature-philosophy concerning time and space (repetition) and (3) practical-ethical implications. These refer to (a) the examiner's clinical responsibility (gathering information for adequate treatment) as well as to (b) the instructor's responsibility (guiding the student towards deliberate behaviour in assessment and treatment). Finally reliability refers to (4) the world-language-relationship and interferes with communication (meaning of clinical findings, usefulness of manual techniques).

Studies on interexaminer reliability for manual assessment techniques within the musculo-sceletal system show moderate to poor agreement, especially for passive mobility testing which represents an essential part of the osteopathic assessment.

Part of the assessment procedure of the Cranial Concept, a part of Osteopathy, is the palpation of the rhythmic cycles of the Primary Respiratory Mechanism (PRM). The existence of the PRM could not consistently be proven up to these days. One of the hypotheses of the Cranial Concept assume, that the PRM is a physiological phenomenon, which can be palpated by trained persons. Another assumption is, that the PRM is an autonomous mechanism which is independent from cardiac and respiratory rates.

For this thesis we hypothesised that, if this is true, the PRM should be palpated consistently between two and within one examiner. In addition there should be no correlations between the examiner's and subject's respiratory rates and the palpated PRM-rates.

49 asymptomatic healthy subjects were palpated by two trained examiners. The respiratory rates of the examiners and the subjects have been recorded simultaneously. Each subject has been palpated at least twice by both examiners simultaneously at the head and at the pelvis. PRM-frequency, mean duration of the flexion phase and mean ratio of flexion- to extension-phase have been described as the main outcome-variables. One the whole the palpated frequencies were lower than reported in older sources. Inter- and intrarater agreement has been analysed by the 95% limits of agreement. Neither inter- nor intrarater agreement could be described beyond chance for all three parameters. Correlation with the examiners' respiratory rates has been significant for both examiners at the pelvis, for one examiner at the head. No correlation could be found with the respiratory rates of the subjects.

The results indicate that the PRM cannot be palpated reliably. Therefore the role of PRM-palpation as a diagnostic tool should be thought over. Furthermore the results support new physiological models for the PRM. What experienced examiners actually perceive when they assess a phenomenon which is interpreted as the PRM could be autonomous regulation processes like Traube-Hering-Mayer-waves.

11 Acknowledgements

First of all I want to thank Prof. Paul Klein, my supervisor for his inspirations and guidance throughout the whole period of the project, from the very first ideas over the actual measurements to the presentation of the final results. Without his influence, his insisting in accuracy and correctness, the whole project would not have been carried out the way it was done.

Next I want to thank Ing. Rudolf Tanzer from the Measurements Group® who kindly prepared the respiratory rate transducers. Special thanks to Matthias Stacher and his supervising teacher Dipl.Ing. Hermann Dum from the *Höhere Technische Bundeslehranstalt Hollabrunn – Abteilung Elektrotechnik-Ausbildungsschwerpunkt Regelungstechnik*, who developed and built the measuring devices and developed the software.

Another special thank you goes to Mag. Alexandra Kaider from the *Institut für medizinische Computerwissenschaften der Universität Wien* (Department of medical computer sciences at the Vienna University), who advised me in methodological aspects and carried out the statistical analysis.

A further and warmhearted thank you to my colleagues Dr. Gudrun Wagner DO and Ulrike Weber DO who kindly took over the exhausting and demanding task to be the examiners.

And finally thousand thanks to my family, for the hours, nights and days they had to do without me as well as for the support they gave, when circumstances seemed to be against the project.

Looking back, I realise how many people have been enabling this project. I would have failed, if one of them would not have participated.

Korneuburg, August 2006

12 References

Abehsera A. 2001. Craniosacrale Osteopathie unter der Lupe. Osteopathische Medizin. 2(4):4-9.

Adams T, Heisey RS, Smith MC, Briner BJ. 1992. Parietal bone mobility in the anesthetized cat. J Am Osteopath Assoc. 92(5):599-622.

Alley JR. 1983. The clinical value of motion palpation as a diagnostic tool. J Can Chiro Assoc. 27:97-100. In: Mootz et al. 1989. Intra- and interobserver reliability of passive motion palpation of the lumbar spine. 12(6):440-445.

Altman DG. 1999. 8th Ed. Practical statistics for medical research. Chapman & Hall/CRC. London, New York, Washington D.C.

Aschoff A. 2000. Intrakranieller Druck und Hirndurchblutung: Eine Einführung in die Pathophysiologie, Klinik und Behandlung des gesteigerten ICP und erniedrigten ICP. http://www.ukl.uni-heidelberg.de/nch/aufsatz/aschoff01.htm. 5.4.2000.

Baker EG. 1971. Alteration in width of maxillary arch and its relation to sutural movement of cranial bones. J Am Osteopath Assoc. 70;6:559-564.

Becker, RE, Brooks, RE [Ed.]. 1997. Life in motion, the osteopathic vision of Rollin E. Becker, D.O. Rudra Press. Portland.

Bering RA. 1962. Circulation of the cerebro-spinal fluid: demonstration of the choroid plexuses as the generator of the force for flow of fluid and ventricular enlargement. J Neurosurg. 19:405-413. In: Feinberg DA, Mark AS. 1987. Human brain motion and cerebrospinal fluid circulation demonstrated with MR velocity imaging. Radiology. 163:793-799.

Berne RM, Levy MN [Eds.]. 1993. 3rd Ed. Mosby Yearbook; Physiology. Mosby-Yearbook, Inc. USA.

Bland JM, Altman DG. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. The Lancet. 8:307-310.

Boline PD, Haas M, Meyer JJ, Kassak K, Nelson C, Keating JCJr. 1993. Interexaminer reliability of eight evaluative dimensions of lumbar segmental abnormality: Part II. J Manipulative Physiol Ther. 16(6):363-74.

Bortz J. 1999. 5. Aufl. Statistik für Sozialwissenschaftler. Springer. Berlin, Heidelberg, New York.

Britt RH, Rossi GT. 1982. Quantitative analysis of methods for reproducing physiological brain pulsations. J of Neuroscience Methods. 6:219-29.

Camm AJ, Malik M, Bigger JT, Breithardt G, Cerutti S, Cohen RJ, Coumel P, et al. 1996. Heart Rate Variability. Standards of Measurement, Physiological Interpretation, and Clinical Use. Circulation. 93(5):1043-1065.

Cooke WH, Hoag JB, Crossman AA, Kuusela TA, Tahvanainen KUO, Eckberg DL. 1999. Human responses to upright tilt: a window on central autonomic integration. J Physiol. 517(2):617-628.

Descartes R. 1986. Meditationes de Prima Philosophia. Meditationen über die Erste Philosophie. Reclam. Stuttgart.

Derrida J. 2001. Limited Inc. W Rappl [Transl.]. Passagen. Wien.

Dickhaus H, Ganser KA, Staubert A, Bonsanto MM, Wirtz CR, Tronnier VM, Kunze S. 1997. Quantification of brain shift effects by MR-imaging. http://www.ukl.uni-heidelberg.de/nch/publikat/here.htm. 15.7.2000.

Downey PA. 2004. Craniosacral therapy: is there biology behind the theory? Dissertation. Dir. MI Siegel. University of Pitsburgh.

Echternach JL. 1994. Invited commentary on "Interrater reliability of craniosacral rate measurements and their relationship with subjects' and examiners' heart and respiratory rate measurements". Phys Ther. 74:917-18.

Feely RA [Hg]. 1988. Clinical cranial osteopathy. Selected readings. Cranial Academy. USA.

Feinberg DA, Mark AS. 1987. Human brain motion and cerebrospinal fluid circulation demonstrated with MR velocity imaging. Radiology. 163:793-799.

Ferré JC, Barbin JY. 1990. The osteopathic cranial concept: Fact or fiction?. Surgical Radiol Anat. 13:165-170.

Fitzgerald GK, McClure PW, Beattie P, Riddle DL. 1994. Issues in determining treatment effectiveness of manual therapy. Phys Ther. 74(3):227-233.

Frymann VM. 1966. Relation of disturbances of craniosacral mechanisms to symptomatology of the newborn: Study of 1250 infants. J Am Osteopath Assoc. 65:1059-75.

Frymann VM. 1971. A study of the rhythmic motions of the living cranium. J Am Osteopath Assoc. 70:928-45.

Green C, Martin CW, Bassett K, Kazanjian A. 1999. A systematic review of craniosacral therapy: biological plausibility, assessment reliability and clinical effectiveness. Complement Ther Med. 7(4):201-207.

Haas M. 1991. Statistical methodology for reliability studies. J Manipulative Physiol Ther. 14:119-32.

Haas M. 1995. How to evaluate interexaminer reliability using an interexaminer reliability study design. J Manipulative Physiol Ther. 18(1):10-15.

Handy ChL. 1948. A history of cranial osteopathy. J Am Osteopath Assoc. 47:269-272.

Hanten WP, Dawson DD, Iwata M, Seiden M, Whitten, FG, Zink T. 1998. Craniosacral rhythm: reliability and relationships with cardiac and respiratory rates. J Orthop Sports Phys Ther. Mar; 27(3):213-8.

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

Hartman SE. 2006. Cranial osteopathy: its fate seems clear. Chiropractic & Osteopathy. 10:4.

Hawk C, Phongphua C, Bleecker J, Swank L, Lopez D, Rubley T. 1999. Preliminary study of the reliability of assessment procedures for indications for chiropractic adjustments of the lumbar spine. J Manipulative Physiol Ther. 22(6):382-89.

Heidegger, Martin. 1993. Sein und Zeit. 17. Aufl. Max Niemeyer Verlag. Tübingen.

Heisey RS, Adams T. 1992. Measurement of cranial mobility. David Kopf Instruments Newsletter. 31:1-5.

Hsieh CY, Hong CZ, Adams AH, Platt KJ, Danielson CD, Hoehler FK, Tobis JS. 2000. Interexaminer reliability of the palpation of trigger points in the trunk and lower limb muscles. Arch Phys Med Rehabil. 81(3):258-264.

Hubka MJ, Phelan SP. 1994. Interexaminer reliability of palpation for cervical spine tenderness. J Manipulative Physiol Ther. 17(9):591-5.

Johnston WL. 1982. Interexaminer reliability studies: Spanning a gap in research - Luisa Burns Memorial Lecture. J Am Osteopath Assoc. 81(12):43-53.

Kant I. 1929 (e-book 2003). Critique of Pure Reason. Transl.: N Kemp Smith. Palgrave Macmillan. Oxford. http://www.hkbu.edu.hk/~ppp/cpr/toc.html. 14.7.2006.

Kappler RE. 1979. Osteopathy in the cranial field: Its history, scientific basis and current status. Osteopath Physician. 46:13-18.

Karemaker JM. 1999. Autonomic integration: the physiological basis of cardiovascular variability. J Physiol. 517(2):316.

Keating JC, Seville J, Meeker WC, Lonczak RS, Quitoriano LA, Dydo M, Leibel DP. 1985. Intrasubject experimental designs in osteopathic medicine: Applications in clinical practice. J Am Osteopath Assoc. 85(3):192-203.

Keating JCJr, Bergmann TF, Jacobs GE, Finer BA, Larson K. 1990. Interexaminer reliability of eight evaluative dimensions of lumbar segmental abnormality. J Manipulative Physiol Ther. 13(8):463-70.

Kirk RE. 1982. Experimental Design. CA: Brooks/Cole Publishing Company. Belmont.

Klein P, Burnotte J. 1985. Contribution à l'ètude biomécanique de la moelle épinière et de ses enveloppes. Annales de Médecine ostéopathique. 1(3):99-105.

Klein P. 2002. Zum Mythos der Schädelknochenmobilität. Einige Überlegungen eines freien Osteopathen. Osteopathische Medizin. 3(2):17-20.

Lantz CA. 1995. Letter to the editor of „How to evaluate intraexaminer reliability using an interexaminer reliability study design". J Manipulative Physiol Ther. 18(7):486-487.

Lawson A, Calderon L. 1998. Interexaminer agreement for applied kinesiology from an interexaminer reliability study design. Percept Mot Skills. Apr;84(2):539-546.

Lay EM, Cicorda RA, Tettambel M. 1978. Recording of the cranial rhythmic impulse. J Am Osteopath Assoc. 78:149.

Leibniz GW. 1996. Schriften zur Mathematik. In: Hauptschriften zur Grundlegung der Philosophie. Teil 1. E Cassirer (Hg.). Meiner. Hamburg. S 35–81.

Levin JS, Glass TA, Kushi LH, Schuck JR, Steele L, Jonas WB. 1997. Quantitative methods in research on complementary and alternative medicine. A methodological manifesto. Medical Care. 35(11):1079-1094.

Lewit K, Liebenson C. 1993. Palpation - Problems and Implications. J Manipulative Physiol Ther. 16(9):586-590.

Liem T. 1998. 2. Aufl. Kraniosakrale Osteopathie: Ein praktisches Lehrbuch. Hippokrates. Stuttgart.

Lockwood MD, Degenhardt BF. 1998. Cycle-to-cycle variability attributed to the primary respiratory mechanism. J Am Osteopath Assoc. 98(1):35-42.

Lundin M, Wiksten JP, Perakyla T, Lindfors O, Savolainen H, Skytta J, Lepantalo M. 1999. Distal pulse palpation: is it reliable?. World J Surg. Mar;23(3):252-5.

Magoun HI, Magoun HI[Hg]. 1997. 2. Aufl. Osteopathy in the cranial field; Original Edition; The application to the cranium of the principles of osteopathy, based on the arduous study and keen clinical observation of William Garner Sutherland. SCTF, Inc. Texas.

Martins AN, Wiley JK, Myers PW. 1972. Dynamics of the cerebrospinal fluid and the spinal dura mater. J Neurol Neurosurg Psychiat. 35:468-73.

Mc Connell EA. 1943. Vascular factor in intracranial pressure and maintenance of cerebrospinal fluid circulation Brain. 66:204-228.. In: Feinberg DA, Mark AS. 1987. Human brain motion and cerebrospinal fluid circulation demonstrated with MR velocity imaging. Radiology. 163:793-799.

McConnell DG, Beal MC, Dinnar U, Goodridge JP, Johnston WL, Karni Z, Upledger JE, Blum G. 1980. Low agreement of findings in neuromusculoskeletal examinations by a group of osteopathic physicians using their own procedures. J Am Osteopath Assoc. 79(7):441-450.

McCraty R, Atkinson M, Tiller WA. 1995. New electrophysiological correlates associated with intentional heart focus Subtle Energies. 4:251-268.. In: McPartland JM. 1997. Entrainment and the cranial rhythmic impulse. Alternative Therapies. 3(1):40-45.

McGregor M. 1991. Letter to the Editor of: „Statistical Methodology for Reliability Studies". J Manipulative Physiol Ther. 14(9):544-45.

McPartland JM. 1997. Entrainment and the cranial rhythmic impulse. Alternative Therapies. 3(1):40-45.

Meijne W, van Neerbos K, Aufdemkampe G, van der Wurff P. 1999. Intraexaminer and interexaminer reliability of the Gillet test. J Manipulative Physiol Ther. 22(1):4-9.

Michael DK, Retzlaff EW. 1975. A preliminary study of cranial bone movement in the squirrel monkey. J Am Osteopath Assoc. 74:866-9.

Michael DK. 1975. The cerebrospinal fluid: Values for compliance and resistance to absorption. J Am Osteopath Assoc. 74:873-76.

Michael LD, Gallo AE, Gehling GF, Isom JB, Mauch W, Kingsley PD. 1982. Fluctuation of intracranial pressure associated with cardiac cycle. J Neurosurg. 11(5):617-621.

Mills MV, Henley CE, Barnes LL, Carreiro JE, Degenhardt BF. 2003. The use of osteopathic manipulative treatment as adjuvant therapy in children with recurrent acute otitis media. Arch Pediatr Adolesc Med. 157(9):861-866.

Mior SA, McGregor M, Schut B. 1990. The role of experience in clinical accuracy. J Manipulative Physiol Ther. 13(2):68-71.

Mitchell FL, Pruzzo NA. 1971. Investigation of Voluntary and Primary Respiratory Mechanisms. J Am Osteopath Assoc. 70:1109-1113.

Mootz RD, Keating JC, Kontz HP, Milus TB, Jacobs GE. 1989. Intra- and interobserver reliability of passive motion palpation of the lumbar spine. J Manipulative Physiol Ther. 12(6):440-45.

Moran RW, Gibbons P. 2002. Intraexaminer and Interexaminer Reliability for Palpation of the Cranial Rhythmic Impulse at the Head and Sacrum. J Manipulative Physiol Therap. 24(3):183-190.

Myers KA, Scott DF, Devine TJ, Johnston AH, Denton MJ, Gilfillan IS. 1987. Palpation of the femoral and popliteal pulses: a study of the accuracy as assessed by agreement between multiple observers. Eur J Vasc Surg. Aug;1(4):245-249.

Nelson KE, Sergueef N, Lipinski CM, Chapman AR, Glonek T. 2001. Cranial rhythmic impulse related to the Traube-Hering-Mayer oscillation: comparing laser-Doppler flowmetry and palpation. J Am Osteopath Assoc. 101(3):163–173.

Nelson KE, Sergueef N, Glonek T. 2006. Recording the rate of the cranial rhythmic impulse. J Am Osteopath Assoc. 106(6):337–341.

Nielsson N. 1996. Letter to the editor of "How to evaluate intraexaminer reliability using an interexaminer reliability study design". J Manipulative Physiol Ther. 19(2):149-150.

Norman GR. 1995. Letter to the editor of "How to evaluate intraexaminer reliability using an interexaminer reliability study design". J Manipulative Physiol Ther. 18(7):484-485.

Norton JM. 1991. A tissue pressure model for palpatory perception of the cranial rhythmic impulse. J Am Osteopath Assoc. 91:975-94.

Norton JM, Sibely G, Broder-Oldach R. 1992. Characterization of the cranial rhythmic impulse in healthy human adults. Am Acad Osteopath J. 2(3):9-12,26.

Norton JM. 1996. A challenge to the concept of craniosacral interaction. Am Acad Osteopath J. 15-21.

O'Haire C, Gibbons P. 2000. Inter-examiner and intra-examiner agreement for assessing sacroiliac anatomical landmarks using palpation and observation: pilot study. Manual Therapy. 5(1):13-20.

Ottenbacher KJ. 1994. Invited Commentary on "Interrater reliability of craniosacral rate measurements and their relationship with subjects' and examiners' heart and respiratory rate measurements". Phys Ther. 74:918-19.

Patterson MM. 1990. Calling for theories that test the underlying osteopathic concepts. J Am Osteopath Assoc. 90(9):785, 789.

Pöltner G. 2002. Grundkurs Medizin-Ethik. 1. Aufl. Facultas. Wien.

Popevec JP, Biggert TP, Retzlaff EW. 1976. Histological techniques for cranial bone studies. J Am Osteopath Assoc. 75:606-607.

Portney LG, Watkins MP. 2000. Foundations of clinical research; applications to practice. 2nd Ed. Prentice Hall Health. New Jersey.

Poupot Ch, Rommeveaux L. 1992. La mobilité des os du crâne démontrée à Grenoble. Les affiches de Grenoble et du Dauphiné. 3 Juillet:3-5.

Quaid A. 1995. Cranial controversy. Response to: Interrater reliability of craniosacral rate measurements and theis relationship with subjects' and examiners' heart and respiratory rate measurements. Phys Ther. 75(3):240.

Quine WVO. 2003. On what there is. In (idem): From a logical point of view. Nine logico-philosophical essays. 2nd Ed. revised (1st Ed. 1953). Harvard University press. Cambridge a.o.

Retzlaff EW, Michael DK, Roppel RM. 1975. Cranial bone mobility. J Am Osteopath Assoc. 74:869-73.

Retzlaff EW, Roppel RM, Becker F, Mitchell FL, Upledger JE. 1976. Cranial mechanisms. J Am Osteopath Assoc. 76:288-89.

Retzlaff EW, Michael D, Roppel R, Mitchell FL. 1976. The structures of cranial bone sutures. J Am Osteopath Assoc. 75:607-8.

Retzlaff EW, Mitchell FL, Upledger JE, Biggert T. 1978. Nerve fibers and endings in cranial sutures. J Am Osteopath Assoc. 77:474-75.

Retzlaff EW, Mitchell FL, [Eds.]. 1987. The Cranium and Its Sutures; Anatomy, Physiology, Clinical Applications and Annotated Bibliography of Research in the Cranial Field. Springer-Verlag. Berlin, Heidelberg, New York, London, Paris, Tokyo.

Rogers JS, Witt PL. 1997. The controversy of cranial bone motion. J Orthop Sports Phys Ther. 26(2):95-103.

Rogers JS, Witt PL, Gross MT, Hacke JD, Genova PA. 1998. Simultaneous palpation of the craniosacral rate at the head and feet: intrarater and interrater reliability and rate comparisons. Phys Ther. 78(11):1175-85.

Rommeveaux L. 1992. La mobilité des os du crâne: une vérité scientifiquement démontrée. unpublished data.

Roppel RM, Pierre NS, Mitchell FL. 1978. Measurement of accuracy in bimanual perception of motion. J Am Osteopath Assoc. 77:47.

Schöps P, Pfingsten M, Siebert U. 2000. Reliabilität manualmedizinischer Untersuchungstechniken an der Halswirbelsäule. Studie zur Qualitätssicherung in der manuellen Diagnostik. Z Orthop Ihre Grenzgeb. Jan-Feb;138(1):2-7.

Sergueef N, Nelson KE, Glonek T. 2002. The effect of cranial manipulation on the Traube-Hering-Mayer oscillation as measured by laser-Doppler flowmetry. Altern Ther Health Med. 8(6):74–76.

Sibley G, Broder-Oldbach RE, Norton JM. 1992. Interexaminer agreement in the characterization of the cranial rhythmic impulse. J Am Osteopath Assoc. 92:1285.

Smith D, Craige E. 1980. Enhancement of Tactile Perception in Palpation. Circulation. 62(5):1114-1118.

St.Pierre N, Roppel R, Retzlaff E. 1976. The detection of relative movements of cranial bones. J Am Osteopath Assoc. 76:289.

Sullivan Ch. 1997. Introducing the cranial approach in osteopathy and the treatment of infants and mothers. Complement Ther Nurs Midwifery. 3(3):72-76.

Sutherland WG. 1990. Teachings in the science of Osteopathy. Wales AL [Ed.]. Rudra Press. Portland, Oregon.

Sutherland WG. 1994. 2nd Ed. The Cranial Bowl: A Treatise Relating to: Cranial Articular Mobility, Cranial Articular Lesions and Cranial Technique. Free Press Company. USA.

Sutherland WG. 1998. 2nd Ed. Contributions of thought: The collected writings of William Garner Sutherland, D.O. A Strand-Sutherland & AL Wales [Eds.]. Rudra Press. Portland, Oregon.

Taylor AJ, Eckberg DL. 1996. Fundamental Realtions Between Short-term RR Interval and Arterial Pressure Oscillations in Humans. Circulation. 93(8):1527-1532.

Tettambel M, Cicora RA, Lay EM. 1978. Recording of the Cranial Rhythmic Impulse. J Am Osteopath Assoc. 78:149.

Thomas von Aquin. 1986. Von der Wahrheit. De Veritate. Meiner. Hamburg.

Tiller WA, McCraty R, Atkinson M. 1996. Cardiac coherence: a new, non invasive measure of autonomic nervous system order. Altern Ther Health Med. 2:52-65. In: McPartland JM. 1997. Entrainment and the cranial rhythmic impulse. Alternative Therapies. 3(1):40-45.

Upledger JE. 1977. The reproducibility of craniosacral examination findings: a statistical analysis J Am Osteopath Assoc. 76(12):890-9. In: Upledger JE, Vredevoogd JD. 1994. 2nd Ed. Lehrbuch der Kraniosakral-Therapie. Übers: J Weingart, HA Moll. Haug. Heidelberg.

Upledger JE, Karni Z. 1979. Mechanio-electric patterns during craniosacral osteopathic diagnosis and treatment J Am Osteopath Assoc. 78(11):782-91.. In: Upledger JE, Vredevoogd JD. 1994. 2nd Ed. Lehrbuch der Kraniosakral-Therapie. Übers: J Weingart, HA Moll. Haug. Heidelberg.

Upledger JE, Karni Z. 1979. Mechanio-electric patterns during craniosacral osteopathic diagnosis and treatment. J Am Osteopath Assoc. 78(11):782-91.

Upledger JE, Vredevoogd JD. 1994. 2nd Ed. Lehrbuch der Kraniosakral-Therapie. Übers: J Weingart, HA Moll. Karl F. Haug. Heidelberg.

von Engelhardt D. 2005. Krankheit, Krankheitsbegriff (Neuzeit). Historischer Verlauf, Grunddimensionen. In: Enzyklopädie Medizingeschichte. Gerabek WE, Haage BD, Keil G, Wegner W [Hgg.]. De Gruyter. Berlin. S 797–803.

Von Foerster H, SJ Schmidt [Hg]. 1997. 4th Ed. Wissen und Gewissen. Versuch einer Brücke. Suhrkamp. Frankfurt a. Main.

Wieland W. 2004. Diagnose. Überlegungen zur Medizintheorie. Bibliothek des skeptischen Denkens. Warendorf.

Wirth-Pattullo V, Hayes KW. 1994. Interrater reliability of craniosacral rate measurements and their relationship with subjects' and examiners' heart and respiratory rate measurements. Phys Ther. 67(10):1526-32.

Wittgenstein L. 1984. Tractatus logico-philosophicus. Werkausgabe Bd 1. Suhrkamp. Frankfurt a.M.

Woods JM, Woods RH. 1961. A physical finding related to psychiatric disorders. J Am Osteopath Assoc. 60:988-993.

Zeidler KW. 2000. Prolegomena zur Wissenschaftstheorie. Könighausen & Neumann. Würzburg.

13 Appendix

List of abbreviations

A	Ampere; Examiner A
A/D	Analogue/Digital
B	Examiner B
C	Cranium (head)
Ch	Channel
cm	Centimetre
COT	Contributions of Thought
CRDP	Central respiratory drive potential
CSF	Cerebrospinal fluid
CT	Computer tomography
DF	Degrees of freedom
ECG	Electrocardiography
EMG	Electro myography
E_{Diff}	Expected difference
F	F-value (distribution)
f	PRM-frequency (-rate)
ger.	German
HF	High frequency
HRV	Heart rate variability
Hz	Hertz
ICC	Interclass correlation coefficient
κ	Kappa-Index
k	Filtration constant for the capillary membrane
L	Examiner location
LF	Low frequency
log	Logarithm
m^2	Square metre
MDF	Mean duration of the flex-ion-phase
min	Minute(s)
mm	Micrometre
M_{Diff}	Mean difference
MR	Magnet resonance
MRI	Magnet resonance imaging
MS	Mean squares
ms	Millisecond(s)
N	Newton, Number
n	Number
O	Observer (examiner)
OCF	Osteopathy in the Cranial Field
π_i	Interstitial fluid oncotic pressure
π_p	Plasma protein oncotic pressure
p	Page
P	Pelvis
P	P-Value
P_A	Pressure in the large arteries
P_C	Capillary hydrostatic pressure
P_i	Interstitial fluid hydrostatic pressure
P_{net}	Net pressure

PRM	Primary Respiratory Mechanism
PV	Pressure in the large veins
r	Pearson's correlation coefficient
R	Resistance
RE_A	Respiratory rate of examiner A
RE_B	Respiratory rate of examiner B
RE_O	Respiratory rate of the examiners
REP	Repetition of measurements
RE_{SU}	Respiratory rate of the subject
$R_{F/E}$	Mean ration of flexion- to extension-phase
RRT	Respiratory rate transducer
S	Signal
SA	slowly adapting
SD	Standard deviation
sec	Second(s)
SG	Strain gauge
SS	Sum of squares
Sub	Subject(s)
T	Examination-time
t	Time-period
T 1	1st Examination
T 2	2nd Examination
U	Current tension
ULF	Ultra low frequency
V	Volt
VLF	Very low frequency

– List of figures –

– List of tables –

– Index –

– Raw data –

The complete raw data are available on CD-ROM. The data given below served as basis for statistical analysis. They represent the calculated main outcome measures (PRM-frequency – *f*, mean duration of the flexion-phase – MDF, mean ratio of flexion- to extension-phase –$R_{F/E}$) and the covariables for the analysis of correlation (Respiratory rate of examiner A – R_{EA}, Respiratory rate of examiner B – R_{EB}, Respiratory rate of the subject – RE_{SU}) taken from the raw data of each single measurement. Calculations of the main outcome measures are explained under point »Data analysis« (see p 77ff).

SU .... Subject
REP .... Repetition of measurements (I = T 1 & T 2, II = T 3 & T 4)
O / C .... Examiner at the location cranium
f-C .... PRM-Frequency, palpated at the cranium
f-P .... PRM-Frequency, palpated at the pelvis
MDF-C .... Mean duration of the flexion-phase, palpated at the cranium
MDF-P .... Mean duration of the flexion-phase, palpated at the pelvis
$R_{F/E}$-C ..... Mean ratio of flexion- to extension-phase, palpated at the cranium
$R_{F/E}$-P ..... Mean ratio of flexion- to extension-phase, palpated at the pelvis
RE_A .... Respiratory rate of examiner A
RE_B .... Respiratory rate of examiner B
RE_{SU} .... Respiratory rate of the subject
O/T 1-C ... Examiner at examination-time T 1 at examiner-location cranium

SU	REP	O / C	*f*-C	*f*-P	MDF-C	MDF-P	$R_{F/E}$-C	$R_{F/E}$-P	RE_A	RE_B	RE_{SU}	O/T1-C
1	I	G	4	5	7,25	7,30	0,62	1,08	28	34	26	U
1	I	U	8	10	4,69	3,50	0,76	0,65	45	51	33	U
1	II	G	2	5	19,50	7,80	1,67	3,97	31	32	29	U
1	II	U	5	4	8,50	4,63	1,15	0,38	29	34	32	U
2	I	G	6	8	3,08	4,19	0,43	0,56	45	50	37	G
2	I	U	5	5	8,60	8,90	0,88	1,40	29	33	28	G
2	II	G	5	4	7,00	13,38	0,55	1,45	26	32	28	G
2	II	U	4	3	11,88	14,00	1,22	1,15	29	35	31	G
3	I	G	3	5	6,17	8,00	0,46	0,82	28	33	21	U
3	I	U	5	5	10,20	6,00	1,16	0,75	29	34	16	U
3	II	G	2	4	9,25	10,13	0,28	0,90	30	33	11	U
3	II	U	8	8	5,75	4,06	1,38	0,74	45	51	19	U
4	I	G	4	3	7,50	12,17	0,72	1,13	28	31	10	G
4	I	U	7	7	6,43	4,79	1,20	0,67	44	50	17	G

SU	REP	O / C	*f*-C	*f*-P	MDF-C	MDF-P	$R_{F/E}$-C	$R_{F/E}$-P	RE_A	RE_B	RE_{SU}	O/T1-C
4	II	G	5	6	7,50	6,25	0,80	0,73	45	52	15	G
4	II	U	4	6	10,50	8,17	1,16	1,28	28	36	11	G
5	I	G	2	7	15,25	6,00	0,55	1,15	42	48	29	U
5	I	U	5	4	9,00	5,38	2,75	0,58	31	36	17	U
5	II	G	5	4	7,50	11,00	0,52	0,71	28	35	18	U
5	II	U	4	3	12,13	24,00	1,22	0,31	27	31	17	U
6	I	G	6	6	6,83	6,00	1,06	0,70	28	34	18	G
6	I	U	4	5	10,25	4,80	2,05	0,39	30	36	25	G
6	II	G	3	4	6,33	9,63	0,43	0,76	29	35	16	G
6	II	U	4	6	11,38	6,58	1,23	2,77	30	36	13	G
7	I	G	5	3	5,50	14,50	0,55	1,30	29	31	24	U
7	I	U	5	5	9,10	6,40	0,90	0,66	23	35	25	U
7	II	G	7	4	3,64	10,88	0,58	0,98	28	35	26	U
7	II	U	7	11	7,29	4,82	1,29	1,50	40	42	41	U
8	I	G	5	4	8,60	10,88	1,28	0,93	27	31	16	G
8	I	U	4	5	10,63	7,00	1,49	0,92	26	34	17	G
8	II	G	6	5	5,67	9,30	0,67	1,58	29	33	24	G
8	II	U	5	5	9,50	6,00	1,34	0,65	28	31	23	G
9	I	G	4	6	14,13	7,17	0,67	0,85	31	34	11	U
9	I	U	6	4	7,08	9,63	1,01	1,43	26	32	12	U
9	II	G	1	6	90,50	7,83		1,06	29	38	12	U
9	II	U	6	7	7,42	5,71	0,98	0,92	32	35	10	U
10	I	G	2	4	19,75	10,00	0,35	1,01	30	35	24	G
10	I	U	5	4	9,40	10,00	1,02	1,06	32	31	17	G
10	II	G	2	5	26,50	8,40	2,33	0,78	32	36	26	G
10	II	U	5	4	9,20	10,63	1,65	0,94	30	37	25	G
11	I	G	8	5	5,00	7,30	0,91	0,83	26	34	29	G
11	I	U	6	6	8,08	4,83	1,02	0,56	26	35	34	G
11	II	G	6	5	7,92	5,00	1,74	0,53	46	50	43	G
12	I	G	5	5	7,60	10,10	0,95	1,09	30	34	19	U
12	I	U	6	6	9,08	7,50	1,38	0,91	29	33	20	U
12	II	U	5	5	9,90	5,50	1,45	0,54	30	44	12	U
13	I	G	6	6	6,75	7,83	0,80	1,13	27	35	29	U
13	I	U	4	6	9,25	5,67	1,10	0,88	27	39	32	U
13	II	U	6	6	8,08	4,92	1,26	0,67	32	36	33	U
14	I	G	3	5	7,50	5,70	0,28	0,44	30	35	19	G
14	I	U	6	2	9,92	7,50	1,85	0,44	27	36	20	G
14	II	G	4	5	6,63	8,40	0,54	0,84	33	38	27	G
15	I	G	4	5	10,25	8,80	0,82	0,86	32	37	16	G
15	I	U	5	3	8,00	8,83	0,90	0,60	31	37	15	G
15	II	U	7	7	6,93	7,07	1,36	1,38	25	34	9	G
16	I	G	5	6	8,30	9,08	2,77	1,38	29	33	26	G
16	I	U	6	6	9,42	6,58	1,60	1,13	27	36	30	G
16	II	U	6	6	9,08	5,33	1,55	0,64	25	38	31	G
17	I	G	4	5	9,63	7,80	1,13	0,86	31	37	27	G
17	I	U	6	7	10,00	6,79	1,90	1,05	27	35	22	G
17	II	G	5	6	8,50	8,25	0,97	1,46	29	36	27	G
18	I	G	4	6	12,63	8,58	3,13	1,11	32	37	18	G
18	I	U	5	6	10,40	7,42	1,51	1,53	27	35	19	G
18	II	U	6	6	8,92	9,17	1,42	1,71	30	39	15	G
19	I	G	4	6	10,88	6,50	1,00	0,71	33	33	22	U
19	I	U	8	5	5,31	8,90	0,81	0,87	31	35	27	U
20	I	G	4	5	8,00	7,20	0,59	0,76	30	37	26	G
20	I	U	6	4	6,92	10,75	1,22	1,00	30	34	24	G

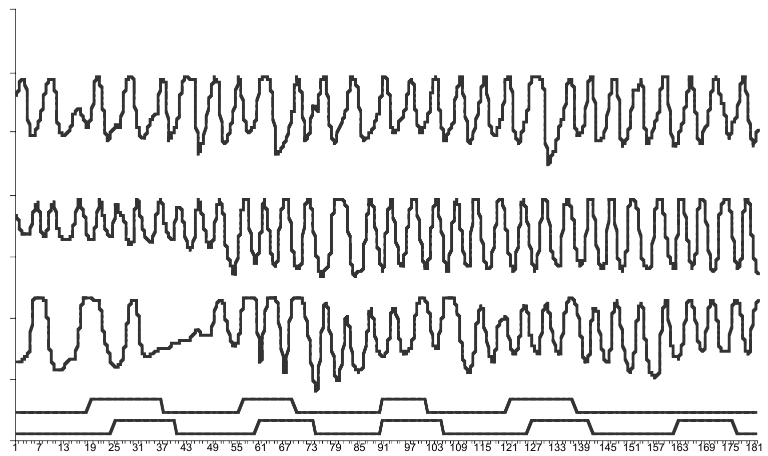
SU	REP	O / C	*f*-C	*f*-P	MDF-C	MDF-P	$R_{F/E}$-C	$R_{F/E}$-P	RE_A	RE_B	RE_{SU}	O/T1-C
21	I	G	3	6	11,83	6,42	1,26	0,84	28	30	4	U
21	I	U	6	5	5,75	9,10	0,71	0,97	29	37	5	U
22	I	G	1	7	74,00	6,50	11,38	1,01	33	33	19	G
22	I	U	5	5	6,40	10,80	0,65	1,40	30	40	18	G
23	I	G	2	5	5,50	8,40	0,29	0,88	31	32	35	U
23	I	U	8	7	4,94	4,36	0,73	0,60	41	34	28	U
24	I	G	3	4	7,83	10,38	0,50	0,72	31	34	26	G
24	I	U	6	6	5,42	6,17	1,24	0,72	38	46	29	G
25	I	G	4	4	11,25	8,38	1,30	0,73	36	34	24	U
25	I	U	6	4	6,83	6,38	0,82	0,53	30	32	20	U
26	I	G	3	5	19,50	8,60	1,97	0,73	32	31	12	G
26	I	U	5	4	9,30	9,38	1,20	0,94	34	34	16	G
27	I	G	2	4	9,50	9,13	0,13	0,69	33	35	31	U
27	I	U	5	2	8,70	3,50	0,81	0,20	33	36	33	U
28	I	G	4	5	13,00	6,80	0,81	0,71	34	35	23	G
28	I	U	5	5	8,90	12,20	2,40	1,30	32	36	19	G
29	I	G	4	5	14,00	8,90	1,21	0,85	33	35	20	U
29	I	U	5	4	9,80	10,00	1,05	1,03	34	34	19	U
30	I	G	2	4	28,50	8,50	1,05	0,68	34	34	21	G
30	I	U	5	4	9,10	10,38	0,80	2,48	32	37	23	G
31	I	G	5	4	8,10	9,88	1,97	0,94	32	34	25	U
31	I	U	6	5	7,75	7,50	0,99	1,01	29	35	26	U
32	I	G	1	3	7,50	14,67	0,13	1,41	31	34	17	G
32	I	U	4	2	14,13	25,00	1,24	10,63	30	29	17	G
33	I	G	2	3	16,25	12,83	1,00	0,82	31	34	33	U
33	I	U	4	3	11,50	9,83	1,09	0,61	31	33	27	U
34	I	G	3	5	10,33	9,50	0,63	1,00	32	36	23	G
34	I	U	5	4	9,70	10,63	1,07	1,29	30	33	19	G
35	I	G	2	3	7,25	13,83	0,31	0,89	31	33	15	U
35	I	U	5	2	10,80	4,50	1,20	0,16	31	28	19	U
36	I	G	3	3	5,83	13,17	1,79	1,00	28	30	31	G
36	I	U	4	4	11,50	10,38	1,20	0,92	30	34	32	G
37	I	G	4	4	10,00	9,38	0,98	0,68	32	33	25	U
37	I	U	5	5	10,50	7,10	2,76	0,73	28	36	26	U
38	I	G	4	4	10,00	10,63	1,19	1,04	31	34	23	G
38	I	U	3	5	8,00	9,80	0,48	1,12	32	32	26	G
39	I	G	3	5	13,33	8,90	0,82	0,87	30	27	12	U
39	I	U	4	3	12,38	9,83	1,66	0,69	26	34	12	U
40	I	G	2	4	10,00	12,00	0,59	1,09	30	34	32	G
40	I	U	5	3	9,20	10,00	1,24	1,22	28	23	31	G
41	I	G	1	4	6,50	9,75	0,20	0,80	30	31	27	U
41	I	U	5	4	10,30	9,88	1,51	0,72	24	31	21	U
42	I	G	3	4	11,17	10,00	0,60	0,88	29	33	12	G
42	I	U	5	3	11,00	10,50	1,29	0,73	25	21	11	G
43	I	G	2	4	22,50	11,50	1,96	1,16	30	34	29	U
43	I	U	5	2	11,70	25,00	1,49	1,30	29	29	31	U
44	I	G	4	4	11,63	11,50	0,69	1,97	34	29	23	U
44	I	U	4	6	12,25	7,75	1,35	1,23	33	27	19	U
45	I	G	3	4	6,00	13,38	0,32	4,36	31	28	11	G
45	I	U	4	4	13,13	3,75	1,94	0,57	32	29	13	G
46	I	G	4	4	10,00	11,50	6,72	0,88	32	29	31	U
46	I	U	5	4	10,60	7,88	1,45	0,62	33	25	32	U
47	I	G	5	4	7,90	10,50	0,93	0,82	33	32	18	G
47	I	U	5	5	9,90	9,20	1,24	1,29	34	31	11	G

SU	REP	O / C	*f*-C	*f*-P	MDF-C	MDF-P	$R_{F/E}$-C	$R_{F/E}$-P	RE_A	RE_B	RE_{SU}	O/T1-C
48	I	G	3	4	12,33	13,13	0,53	0,91	32	29	25	U
48	I	U	4	5	8,75	9,10	1,03	0,93	32	26	27	U
49	I	G	4	3	10,00	15,50	0,80	1,39	33	27	22	G
49	I	U	5	4	10,90	11,88	1,58	1,25	29	27	24	G

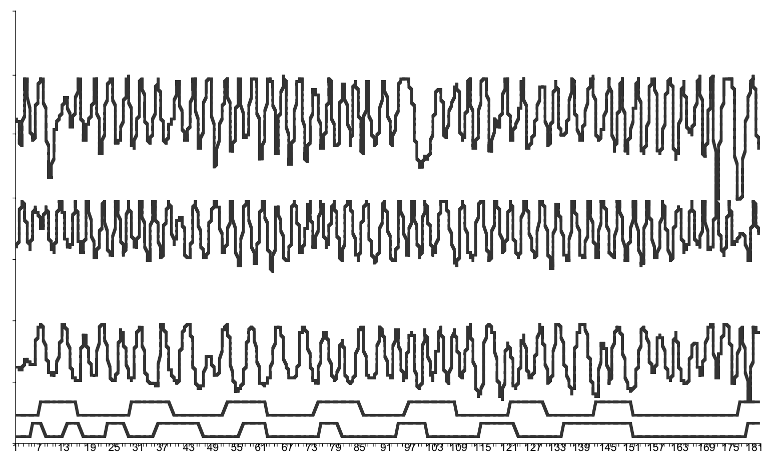
– Graphical samples –

Graphical samples of the first four measurements. The graphs represent (listed from above):

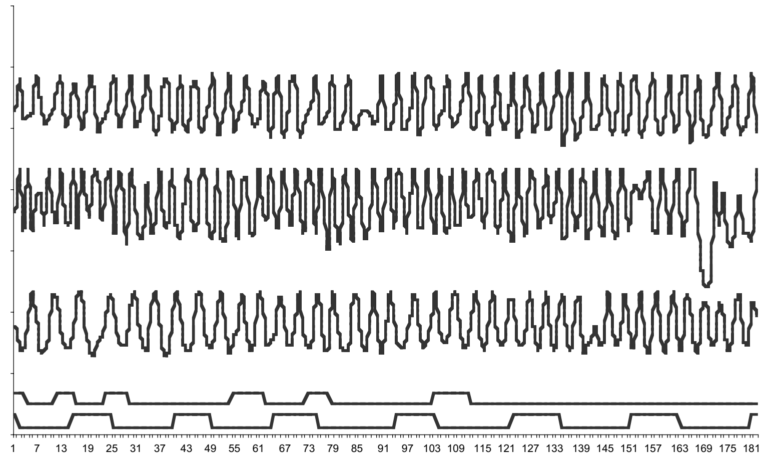
Respiratory rate of examiner A
Respiratory rate of examiner B
Respiratory rate of subject
PRM-palpation of examiner A
PRM-palpation of examiner B



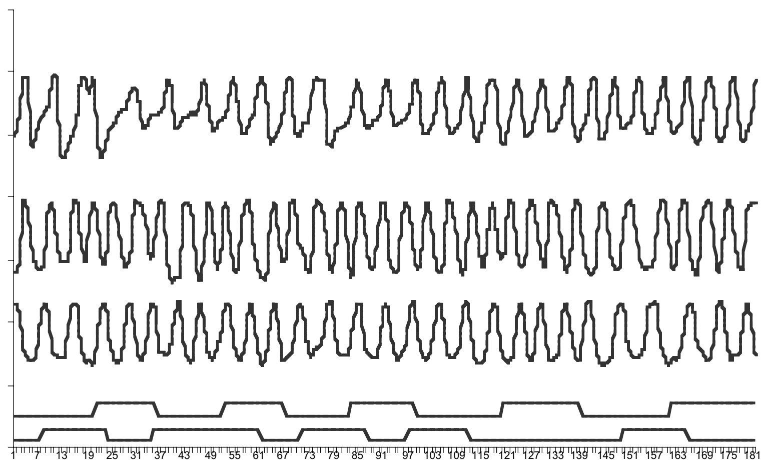
Subject 1 – Examiner A at the head



Subject 1 – Examiner B at the head



Subject 2 – Examiner A at the head



Subject 2 – Examiner B at the head