Antti Rantanen

DEVELOPMENT OF METHODOLOGY FOR ASSESSING COUNSELING INTERACTIONS

DEVELOPING THE COUNSELOR RESPONSE OBSERVATION SYSTEM AND ASSESSING APPLICABILITY OF HEART RATE VARIABILITY TO THE MEASUREMENT OF CLIENT EMOTIONS DURING VERBAL REPORTING
ANTTI RANTANEN

DEVELOPMENT OF METHODOLOGY FOR ASSESSING COUNSELING INTERACTIONS
Developing the Counselor Response Observation System and assessing applicability of heart rate variability to the measurement of client emotions during verbal reporting

Academic dissertation to be presented with the assent of the Doctoral Training Committee of Human Sciences of the University of Oulu for public defence in Kaljusensali (KTK112), Linnanmäki, on 13 December 2014, at 12 noon
Rantanen, Antti, Development of methodology for assessing counseling interactions. Developing the Counselor Response Observation System and assessing applicability of heart rate variability to the measurement of client emotions during verbal reporting
University of Oulu Graduate School; University of Oulu, Faculty of Education
Acta Univ. Oul. E 150, 2014
University of Oulu, P.O. Box 8000, FI-90014 University of Oulu, Finland

Abstract
The general purpose of this study was to develop methodology for assessing counseling interactions. The first main objective of the research was to develop measures for assessing counselors’ responses and client-centeredness (Study I). Specific aims connected to the first objective were to assess the content validity and inter-rater reliability of the Counselor Response Coding System (CRCS) and the content validity, construct validity and inter-rater agreement of the Skilled Verbal Responding Scale (SVRS), which are measures included in the Counselor Response Observation System (CROS). The second main objective was to assess the applicability of heart-rate variability (HRV) to the measurement of emotions during clients’ verbal reports of unpleasant experiences (Studies II and III). The setting was analog with client-centered counseling aims in that the subjects produced verbal reports of their unpleasant experiences. Specific aims connected to the second objective were to examine the differences in subjects’ emotions between viewing and reporting unpleasant experiences and to examine the differences in their emotions between reporting pleasant and unpleasant experiences by measuring HRV. The results supported the content validity and inter-rater agreement of the CRCS and the content validity, construct validity, and inter-rater reliability of the SVRS. Overall, the results from the CROS testing can be considered a good starting point toward its further validation. The results from the HRV experiments were consistent with earlier findings and indicated that the subjects’ sympathetic activity was lower and that their parasympathetic activity higher when they reported unpleasant pictures than it was when they viewed them. This dynamic was not observed with the pleasant pictures. The results suggest that HRV and emotions during reporting should be interpreted in relation to the initial experiences and emotions in the viewing phase.

Keywords: counselor responses, heart-rate variability, measurement
Rantanen, Antti, Ohjausvuorovaikutuksen arvioinnin metodologinen kehittäminen. Ohjausresponssien mittaussärjestelmän kehittäminen ja sykevariaatiomittauksen soveltuvuuden arviointi asiakkaan emoottioiden mittaamiseen puheen aikana
Oulun yliopiston tutkijakoulu; Oulun yliopisto, Kasvatustieteiden tiedekunta
Acta Univ. Oul. E 150, 2014
Oulun yliopisto, PL 8000, 90014 Oulun yliopisto

Tiivistelmä

Asiasanat: mittaaminen, ohjausresponssit, sykevariaatio
This thesis is dedicated to my family for their love and support
Acknowledgements

First, I would like to thank my wife Marja Rantanen for supporting me throughout the research process. Second, I would like to thank Professor Hannu Soini (Clinic of educational psychology) and Dr. Seppo Laukka (LearnLab) for giving me an opportunity to work in their research projects and for supporting me in my work. Third, I would like to thank all co-authors that have made my research possible including Laboratory technician Antti Siipo, Professor Tapio Seppänen, Professor Matti Lehtihalmes, Dr. Eero Väyrynen, Dr. Suvi Tiinanen and Dr. Jukka Kortelainen. Fourth, I would like to thank pre-examiners Dr. Puukari and professor Peräkylä for valuable feedback on my dissertation manuscript. Finally, I would like to thank researchers Teemu Suorsa and Matti Rantanen for our in-depth discussions pertained to my research topic that have helped me to organize my thoughts and to achieve my goals. This thesis was funded by Academy of Finland (Project 273517).

11.9.2014 Oulu

Antti Rantanen
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRCS</td>
<td>Counselor Response Coding System</td>
</tr>
<tr>
<td>HF</td>
<td>High-frequency</td>
</tr>
<tr>
<td>HRV</td>
<td>Heart rate variability</td>
</tr>
<tr>
<td>LF</td>
<td>Low-frequency</td>
</tr>
<tr>
<td>LF/HF</td>
<td>Low-frequency high-frequency ratio</td>
</tr>
<tr>
<td>SVRS</td>
<td>Skilled Verbal Responding Scale</td>
</tr>
</tbody>
</table>
List of original publications

This thesis is based on the following publications, which are referred throughout the text by their Roman numerals:


# Contents

Abstract

Tiivistelmä

Acknowledgements 9

Abbreviations 11

List of original publications 13

Contents 15

1 Introduction 17

1.1 Counseling interaction ................................................................. 17

1.2 Ingredients of successful counseling .......................................... 18

1.2.1 Common factors ................................................................. 18

1.2.2 Client-centered counselor behaviors ..................................... 19

1.2.3 Measurement of counselor behaviors and client-centeredness ......................................................... 20

1.3 The significance of emotions in counseling ............................... 24

1.3.1 Emotions and client outcome .............................................. 24

1.3.2 Reinforcement of adaptive emotion regulation .................. 26

1.3.3 Measurement of client emotions ........................................... 27

1.4 Critical measurement components for assessing counseling interaction ........................................................................ 31

2 Aims of the study 33

3 Methods 35

3.1 Development of the CROS ....................................................... 35

3.1.1 Participants ...................................................................... 35

3.1.2 Procedure ........................................................................ 35

3.1.3 Measures ........................................................................ 37

3.2 Assessment of applicability of HRV to the measurement of emotions during verbal reports ......................................................... 38

3.2.1 Participants ....................................................................... 38

3.2.2 Apparatus and materials ................................................... 39

3.2.3 Procedure ........................................................................ 39

4 Summary of the results 43

4.1 Development of the Counselor Response Observation System 43

4.1.1 Content validity and inter-rater reliability of the CRCS .......... 43

4.1.2 Content validity and inter-rater reliability of the SVRS .......... 44

4.1.3 Construct validity of the SVRS ........................................... 45
4.2 HRV during the oral reporting of unpleasant experiences ....................... 49

5 Discussion 51

5.1 The Counselor Response Observation System ........................................ 52
5.2 HRV and emotions during verbal reporting on unpleasant experiences .......................................................... 54
5.3 General conclusions .............................................................................. 56
5.4 Limitations and future developments ...................................................... 58

References 61
Appendices 71
List of original publications 73
1 Introduction

In the introduction, I will review the essential literature related to two key components in successful counseling interactions: counselor responses and client emotions. In chapter 1.1, I will describe counseling and counseling interactions. In chapters 1.2 and 1.3, I will review two key components in successful counseling interactions and the types of pertinent measures that have been used in the past during counseling sessions. The components of focus are counselor verbal behaviors and client emotions during sessions. In chapter 1.4, I will summarize the results from earlier studies on these components and present a conclusion regarding the measurements and measures needed in research on counseling interactions.

1.1 Counseling interaction

Generally, counseling refers to the practice of a counselor helping a client to examine and resolve psychological worries and problems (Ivey & Ivey 2007, Okun 2002). In many cases, the terms counseling and psychotherapy are used interchangeably (Okun 2002). They share the same variety of theoretical orientations and methods for helping, and similar to psychotherapy, counseling is also a profession that is implemented in a variety of settings (Ivey & Ivey 2007, Okun 2002). However, there are some differences between the concepts of counseling and psychotherapy. Counseling can be focused on either nonclinical or clinical clients, but psychotherapy typically addresses only clinical clients and disorders (Okun 2002). Furthermore, in some contexts, counseling is focused on more specific problem areas than is psychotherapy, which usually addresses larger areas of an individual’s psychological functioning (Ivey & Ivey 2007). Counseling interaction refers to the relationship between counselor and client behaviors that can be examined with a variety of counselor and client variables that are focused, for example, on participants’ verbal behaviors, the quality of the interactions or their emotions (Beutler et al. 2004, Clarkin & Levy 2004). Examination of this relationship is crucial for obtaining information on the types of counselor actions that are associated with different levels of the desired changes, which can only be extracted by using valid measures that can capture aspects of counselor actions and client responses (Hill 2004, Sloan & Kring 2007). In chapters 1.2 and 1.3, I will focus on two key components of successful counseling interactions and how they have been measured. The components I focused on were counselor behaviors and client emotions that have been connected to positive client outcomes in earlier studies.
1.2 Ingredients of successful counseling

In this chapter, I will review the essential literature pertained to successful counseling. I will describe the critical components that are associated with successful counseling and research on specific counselor behaviors that have been connected to good client outcomes. Finally, I will review the types of measures that have been used in earlier studies to measure counselor behaviors.

1.2.1 Common factors

The overall effectiveness of counseling has been well established by the research that has been conducted on the topic (Grisom 1996, Lambert & Bergin 1994, Lipsey & Wilcon 1993, Smith & Glass 1977, Smith et al. 1980). In one of the most extensive meta-analyses of counseling and psychotherapy effectiveness, Smith et al. (1980) analyzed 475 studies and found that counseling and psychotherapy explained 15.2% of the variance in client outcomes. Despite the demonstrated benefits of psychotherapy, there has been continuous debate among researchers regarding the effectiveness of various counseling and psychotherapy approaches. Many meta-analyses, including analyses of randomized clinical trials, have not found significant differences between the approaches (Ahn & Wampold 2001, Cujpers et al., Wampold et al. 1997).

The most critical influences on client outcomes have been found among common factors across different counseling orientations that are related to client, counselor and client-counselor relationship variables (Wampold et al. 1997). Especially strong evidence has been gathered in recent decades on the effect of individual counselor and therapeutic alliance on client outcomes (Crits-Cristoph et al. 1991, Dinger et al. 2008, Okiishi et al. 2003, Horvarth & Symonds 1991, Horvarth et al. 2011, Martin et al. 2000). For instance, in a meta-analysis of 15 studies Crits-Cristoph et al. (1991) found that the therapist effect explained 7.9% of the variance in client outcomes. Additionally, the effect of the client-counselor relationship on client outcomes has gained empirical support (Horvarth & Symonds 1991, Horvarth et al. 2011, Martin et al. 2000). Martin et al. (2000) analyzed 79 studies that focused on the effect of the client-therapist relationship on client outcomes and found that this relationship explained 5% of the variance in client outcomes. In a more recent meta-analysis of 211 studies, Horvarth et al. (2011) found that alliance explained 7.5% of the client outcomes.
1.2.2 Client-centered counselor behaviors

Although individual counselors and their ability to facilitate counseling relationships have been shown to be key factors in counseling outcomes, the question remains regarding the specific counselor actions that contribute to successful counseling outcomes. This question can perhaps best be answered with the results from large reviews (Orlinsky et al. 1994, Orlinsky et al. 2004) of process-outcome studies that have shown that client-centered counselor behaviors are critical in good outcomes. Orlinsky et al. (1994) reviewed 2,354 process-outcome research findings and found variables that had consistent positive relationships with successful therapy outcomes. The criterion for assessing consistent positive relationships with outcomes was that more than 50% of the findings in a process variable category supported the connection between the process variable and a positive outcome (Orlinsky et al. 1994). Therapist variables that had consistent positive relationships with good client outcomes were associated with the therapist’s focus on client problems, affect, cognition and behavioral processes (Orlinsky et al. 1994). Furthermore, therapist contributions to the therapeutic bond, therapist engagement, therapist credibility, therapist empathy, reciprocal attunement, and therapist affirmation of the patient were found to be important for successful outcomes (Orlinsky et al. 1994, Orlinsky et al. 2004). Many of these aspects form the basis of client-centered counseling strategies that emphasize attunement to the client and facilitation of clients’ descriptions of his thoughts, feelings and actions (Carkhuff 1969, Okun 2002, Rogers 1961, Truax & Carkhuff 1967).

Although being client-centered is currently emphasized as an important factor in a variety of traditions, its roots are in humanistic counseling and especially in the work of Carl Rogers (1961). Later humanistic counseling models were constructed by Truax and Carkhuff (1967). Both Rogers (1961) and Truax and Carkhuff (1967) emphasized accurate verbal responses to clients’ descriptions of his experiences and emotions. More recent developments have been presented by Okun (2002), whose ideas are similar to Truax and Carkhuff’s (1967) conceptions of facilitative conditions in counseling and psychotherapy. The essential idea in the client-centered communication strategy is to facilitate the clients’ descriptions of their experiences in ways in which they accurately report observations, thoughts and events associated with his problem (Okun 2002, Rogers 1961, Truax & Carkhuff 1967). Client-centered counseling strategy aims to help clients to connect with their feelings, thoughts and actions and to help clients construct better understandings of their experiences (Okun 2002). These aims are achieved by guiding clients to verbally describe all aspects, but
especially the troubling aspects, of their experience by focusing on main themes and emotions and synthesizing the different parts of their messages (Okun 2002).

There are two main communication techniques that are characteristic of client-centered counseling: reflection and specifying questions (Okun 2002, Rogers 1961, Truax & Carkhuff 1967). Reflection means that the client's words or phrases are repeated or the essence of the client's message is verbally summarized (Okun 2002, Rogers 1961, Truax & Carkhuff 1967). Specifying questions focus on things the client has previously said and aim to deepen the description of the client's experience (Okun 2002, Rogers 1961, Truax & Carkhuff 1967). These two techniques capture the essence of client-centered counseling strategy that aims to enable the client to produce accurate verbalized descriptions and observations of troubling and unpleasant aspects of his experience (Okun 2002, Rogers 1961, Truax & Carkhuff 1967). In this context, emotional expressions are important in recognizing what is important to the client, but client-centered counseling is not focused solely on this. Even though verbal emotional expressions are important, the key idea in client-centered counseling is that specifying the client's message will lead the client to approach and describe the emotional aspects of his experience that are meaningful to him (Okun 2002, Rogers 1961, Truax & Carkhuff 1967).

When focusing on troubling, emotional aspects in his reporting, the client approaches the aspects of his experience that were difficult and that he initially avoided (Okun 2002, Rogers 1961, Truax & Carkhuff 1967). This helps the client to perceive important aspects of his experience and to better direct his future actions (Okun 2002, Rogers 1961, Truax & Carkhuff 1967). With this strategy, the client is exposed to his unpleasant or avoided experiences and is engaged in the non-avoidant task of describing and making observations (Barlow et al. 2004, Okun 2002, Lieberman et al. 2011).

1.2.3 Measurement of counselor behaviors and client-centeredness

Despite the importance of research on counselor behaviors, and especially client-centered responses, earlier studies that used empirical methods to study them have shown several problems with their measurements. For instance, Naugle and Hall (1999) found that in many cases, counselor assessment was performed inadequately in field settings and tended to focus on whether clients liked or approved of their counselors. Further problems have been found with the quality of the applied interactional measures used in counseling and psychotherapy research. A thorough review of candidate therapist-patient interaction measures by Cahill et al. (2008)
showed that most of the measures reported in the literature did not meet their criteria for sufficient validity and reliability. Their review found 83 measures that included basic information relating to reliability and validity (Cahill et al. 2008). In their review, a measure was required to demonstrate at least one aspect of validity (face, content, concurrent, predictive, construct, discriminant) with a value higher than .50 and one aspect of reliability (internal consistency, test-retest, inter-rater) with a value higher than .70 (Cahill et al. 2008).

In our review (Rantanen & Soini 2013), we focused on obtaining information on the number and validity of the observational measures of counselor behaviors that have been used during counseling and psychotherapy sessions. Furthermore, we extracted the number of measures that focused on client-centered aspects of counselor behaviors. We extracted those measures that established basic information on their validity and reliability. The criterion for a measure that had basic validity and reliability information was that at least one aspect of validity and reliability was examined and was concluded to be sufficient by the authors of the original article. The initial search resulted in 8,234 articles, of which, 1,036 articles used a quantitative measure. Of these studies, 186 used quantitative counselor measures and made observations of authentic counseling sessions. Eight hundred and fifty studies were excluded because they did not measure counselors (239) or did not take measurements during authentic sessions (611). Of the 186 included studies, 31 used observational counselor response measures. Of these 31 measures, 25 provided established validity and reliability data. Table 2 summarizes the included measures, listing the measure, its author, and the number of studies that used the measure in our data. The number of studies that used these 25 measures is summarized in table 2. Overall, the included measures were used in 54 studies in the source data. The most frequently used measures were the Counselor Evaluation Rating Scale (Myrick & Kelly 1971) and the Counselor Verbal Response Category System (Hill 1978). Of the 25 measures, 11 measures were constructed for use by an independent observer. Many of the measures were meant to be used by the counselor or the client, making the measurements fundamentally biased.

Coding of specific counselor response units was the focus of 3 measures: the Hill Counselor Verbal Response Category System (HCVRCS) (Hill 1978), the HCVRCS-revised (Friedlander 1982) and the Motivational Interviewing Skills Code (MISC) (Miller 2000). The review shows that the HCVRCS and HCVRCS-revised are the most used coding systems but that they have certain problems in their construct. Overlapping categories and the undifferentiated use of response and response focus are found in both systems, and these scales do not allow the free combining of these...
two variables. The MISC is not intended for general coding of counselor behaviors but instead for giving scores on motivational interviewing via coding (Miller 2000). The MISC is very time-consuming, and this fact led to the development of the Motivational Interviewing Integrity Scale (Moyers et al. 2005), which does not apply the time-consuming coding protocol.

Three rating scales were found that focused on client-centered aspects of counselor behaviors: Carkhuff’s empathy scale (Carkhuff 1969), the Accurate Empathy Scale (Truax & Carkhuff 1967) and the Skilled Counseling Scale (Crews et al. 2005). The first two are different versions of the same measure and represent the ideas of the two pioneers in client-centered, humanistic counseling. In spite of the accurate, client-centered focus of their empathy measures, the items lack specificity. For example, the description of level 1 of the Carkhuff empathy scale is: “The verbal and behavioral expressions of the first person either do not attend to or detract significantly from the verbal and behavioral expressions of the second person in that they communicate significantly less of the second person’s feelings than the second person has communicated himself” (Carkhuff 1969). The items are long descriptions of actions and include parts that may be interpretive. The Skilled Counseling Scale (SCS) is a more recent and more specific 18-item scale that uses the occurrence of defined skills as scoring grounds (Crews et al. 2005). The SCS is a promising development, but it has the same problems as many of the coding systems. It provides clear-cut, identifiable descriptions of counselor actions (e.g., paraphrasing is the act of engaging in brief, accurate, and clear rephrasing of what the client has expressed), but it also has many items that are more ambiguous (e.g., immediacy entails recognition of the immediate feelings (verbal/nonverbal) expressed between the client and the counselor when discussing problem)(Crews et al. 2005). Overall, the number and use of measures for assessing counselor responses and their client-centeredness have been low, and the existing measures have substantial problems with structure and validity. These findings suggest that better-developed measures for coding counselor responses and assessing the client-centeredness of counselor actions are needed.
Table 1. The included measures and the number of candidate studies that used each measure.

<table>
<thead>
<tr>
<th>Measures with established validity and reliability (n = 25)</th>
<th>Author</th>
<th>Times used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Counselor Evaluation Rating Scale</td>
<td>Myrick &amp; Kelly 1971</td>
<td>5</td>
</tr>
<tr>
<td>Hill Counselor Verbal Response Category system</td>
<td>Hill 1978</td>
<td>5</td>
</tr>
<tr>
<td>Carkhuff Empathy Scale</td>
<td>Carkhuff 1969</td>
<td>3</td>
</tr>
<tr>
<td>Counselor Skill and Personal Development Form</td>
<td>Wilbur 1991</td>
<td>3</td>
</tr>
<tr>
<td>Helping Skills Measure</td>
<td>Hill &amp; Kellem 2002</td>
<td>3</td>
</tr>
<tr>
<td>Hill Interaction Matrix</td>
<td>Hill 1966</td>
<td>3</td>
</tr>
<tr>
<td>Motivational Interviewing Skills Code (MiSC)</td>
<td>Miller 2000</td>
<td>3</td>
</tr>
<tr>
<td>Skilled Counseling Scale</td>
<td>Crews et al. 2005</td>
<td>3</td>
</tr>
<tr>
<td>Vanderbilt Psychotherapy Process Scale</td>
<td>O'Malley et al. 1983</td>
<td>3</td>
</tr>
<tr>
<td>Accurate Empathy Rating Scale</td>
<td>Truax &amp; Carkhuff 1967</td>
<td>2</td>
</tr>
<tr>
<td>Counseling Skills Scale</td>
<td>Eriksen &amp; McAuliffe 2003</td>
<td>2</td>
</tr>
<tr>
<td>Cross-Cultural Counseling Inventory - Revised</td>
<td>Lafromboise et al. 1991</td>
<td>2</td>
</tr>
<tr>
<td>Hill Counselor Verbal Response Category System-Rev.</td>
<td>Friedlander 1982</td>
<td>2</td>
</tr>
<tr>
<td>Motivational Interviewing Treatment Integrity Scale</td>
<td>Moyers et al. 2005</td>
<td>2</td>
</tr>
<tr>
<td>Roter Interaction Analysis System</td>
<td>Roter &amp; Hall 1991</td>
<td>2</td>
</tr>
<tr>
<td>Vanderbilt Psychotherapy Process Scale -Revised</td>
<td>Smith et al. 2003</td>
<td>2</td>
</tr>
<tr>
<td>Checklist of Interpersonal Transactions -Revised</td>
<td>Kiesler 1987</td>
<td>1</td>
</tr>
<tr>
<td>Classification System for Counseling Responses</td>
<td>Highlen et al. 1982</td>
<td>1</td>
</tr>
<tr>
<td>Dietitians Interviewing Rating Scale (DIRS)</td>
<td>Gregory et al. 1995</td>
<td>1</td>
</tr>
<tr>
<td>Impact Message Inventory - Circumplex</td>
<td>Kiesler &amp; Schmidt 2006</td>
<td>1</td>
</tr>
<tr>
<td>Interpersonal Communication Rating Scale</td>
<td>Strong et al. 1988</td>
<td>1</td>
</tr>
<tr>
<td>Inventory of Countertransference Behavior</td>
<td>Friedman &amp; Gelso 2000</td>
<td>1</td>
</tr>
<tr>
<td>Collaborative Study Psychotherapy Rating Scale</td>
<td>McIntosh et al. 2005</td>
<td>1</td>
</tr>
<tr>
<td>Relational Communication Coding System</td>
<td>Rogers &amp; Farace 1975</td>
<td>1</td>
</tr>
<tr>
<td>Structural Analysis of Social Behavior</td>
<td>Benjamin 1974</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>54</td>
</tr>
</tbody>
</table>
1.3 The significance of emotions in counseling

In this chapter, I will review the essential literature pertained to emotions and their significance in successful counseling. I will describe the functional and psychophysiological basis of emotions and their connection to client outcomes. Finally, I will review the types of measures and measurement that have been used in earlier studies to measure emotions in counseling and verbalization settings.

1.3.1 Emotions and client outcome

From the motivational point of view, emotions can be perhaps best defined as experiences that are founded on activity of nervous system that helps an individual to respond to threats and life-sustaining events (Lang & Bradley 2010, Lang & Bradley 2013). Emotions have a critical role in goal-directed behavior, where the goal is determined by the things that are important for the survival or the well-being of a person (Lang & Bradley 2010, Lang & Bradley 2013). Approach and avoidance behavior are connected to motivational psychophysiological systems that help organism to respond to events that threaten him or approach goal supporting events (Lang & Bradley 2010). Pleasant emotions are associated with organism’s approach behavior (achieving goals) and unpleasant emotions with avoidance behavior (preventing goal-destructive events) (Lang & Bradley 2010).

Emotions have been usually modeled with the discrete or the dimensional model. Discrete models have been suggested by e.g. Plutchik (1962), Tomkins (1962), Izard (1972), Ekman (1971), and Panksepp (1982). In the discrete approach, emotions are seen adaptive universal processes with distinct characteristics that include appraisal, antecedent events, probable behavioral responses and emotion specific physiology (Ekman 1999). One of the most known basic emotion model is Ekman’s basic 6 including emotions of anger, disgust, fear, happiness, sadness and surprise (Ekman 1999). Dimensional view of emotions has been represented by e.g. Davitz (1969), Russell (1980), Lazarus (1984, 1991) and Lang and Bradley (2013). Lang and Bradley (2013) describe dimensional model, comprised of two motivational systems (appetitive and defensive) that are connected to approach and withdrawal behavior. These motivational systems can vary in terms of arousal, describing the intensity of activation.

Affective responding includes changes in the central nervous system, autonomic nervous system, and changes in internal and external behavior that are always connected to behavioral context of the situation (Lang & Bradley 2010). When
considering research evidence, the best physiological differentiability can be found between pleasant and unpleasant emotions supporting the dimensional approach (Lindquist et al. 2012). Although brain activity measurements have provided some differentiability connected to fear and anxiety (amygdale, orbitofrontal cortex), no clear-cut emotional subsystems supporting the basic emotions hypothesis have been found in the nervous system (Lindquist et al. 2012). Recent developments suggest psychophysiological system associated with emotions are found among distributed neural networks (Hermunstad et al. 2013, Yeo et al. 2011). In a large study by Yeo et al. (2011) where functional connectivity measurements of the brain of 1000 subjects with blood-oxygen-level dependent functional magnetic resonance imaging (BOLD fMRI) were examined and 7 functional networks were extracted: visual, somatomotor, dorsal attention, ventral attention, limbic, frontoparietal and default network. Of the interconnected neural networks, limbic network is essential for properly functioning emotions and core affect of pleasure and displeasure (Lindquist et al. 2012, Oosterwijk et al. 2012, Yeo et al. 2009).

Limbic network is consisted of bilateral anterior temporal lobe, medial temporal lobe, subgenual anterior cingulate cortex, medial and lateral orbitofrontal cortex, and the basal ganglia (caudate, putamen, globus pallidus and central nucleus of the amygdalae) and regulates the autonomic nervous system (ANS) mainly through hypothalamic pathways (Oosterwijk et al. 2012, Yeo et al. 2011). ANS is consisted of sympathetic and parasympathetic divisions (Yeo et al. 2011) and in behavioral sense it can be best comprehended as a functional extension of the limbic network (Lindquist 2012, Oosterwijk et al. 2011). In many studies, unpleasant emotions have been found to be connected to increase in sympathetic activity and pleasant emotions to increase in parasympathetic activity (Thayer & Lane 2009, Elliot et al. 2011, Bosch et al. 2009, Thayer et al. 2012). For instance, in a study by Elliot et al. (2011), 33 subjects participated in intelligence test where they were exposed to a threat cue. When compared to control group, parasympathetic activity was decreased and the change was associated with poor cognitive performance (Elliot et al. 2011).

Several studies have shown that emotions have a critical role in successful counseling and therapy (Boritz et al. 2011, Orlinsky et al. 1994, Orlinsky et al. 2004, Carryer & Greenberg 2010, Missirlian et al. 2005, Gassman et al. 2006). For instance, emotional arousal has been connected to successful client outcomes in a study by Carryer and Greenberg (2010) who compared the level of emotional arousal in 38 depressed clients between good and poor client outcomes and found that high arousal was associated with successful outcomes. Similar results were obtained by Missirlian et al. (2005), who found that high emotional arousal was associated with successful
treatment outcome in different phases of the counseling process among 33 clients. Furthermore, the results from large reviews (Orlinsky et al. 1994, Orlinsky et al. 2004) of process-outcome studies have shown that focusing on clients’ affects and troubling aspects of client’s experience are critical in good outcomes.

1.3.2 Reinforcement of adaptive emotion regulation

One essential factor that may explain the effect of emotional focus is that it reinforces adaptive emotion regulation strategies (Barlow et al. 2004). Emotion regulation strategies refer to actions that aim to influence the intensity, occurrence, experience and expression of emotions (Barlow et al. 2004, Gross & Thompson 2009). Emotion regulation strategies can include situation selection, situation modification, attentional deployment, cognitive change, and response modulation (Gross & Thompson 2009). However, there are maladaptive strategies that aim to irrationally suppress or avoid thoughts, emotions or events that can cause even greater dysfunction in a person’s behavior (John & Gross 2002, Muraven & Baumeister 2000).

Associated with maladaptive emotion regulation strategies, Barlow et al. (2004) identify avoidance behavior as a common basis for a variety of nonclinical and clinical psychological problems. This idea of associating maladaptive emotion regulation with psychological problems has gained empirical support from studies that have emphasized maladaptive strategies such as suppression or experiential avoidance as a common aspect of psychological distress and most common disorders (Barlow et al. 2004). The conception of a common basis for most common psychological disorders including depression, obsessive-compulsive disorder and anxiety disorders has gained substantial evidence with the results from treatment outcome research including psychological and pharmacological treatment of patients with comorbid diagnoses. There is evidence that psychological treatments for anxiety disorders produce significant improvement in additional comorbid anxiety or mood disorders that are not specifically addressed in the treatment (Borkovec et al. 1995, Brown et al. 1995, Davis et al. 2010, Newman et al. 2010). For example, in a study in which 76 adults with generalized anxiety diagnosis, of whom 46 had a comorbid diagnosis, treating the generalized anxiety disorder led to a reduced number of comorbid diagnoses including, for example, depression and social phobias (Newman et al. 2010). Furthermore, depression and anxiety disorders both respond to antidepressant medication, indicating a similar pathophysiology behind the disorders (Belzer & Schneier 2004, Gardner & Boles 2011, Lepola et al. 2003, Hudson et al. 2003).
Furthermore, results from this research area indicate that avoidance behavior and stressful, anxious or depressive episodes are not limited to clinical populations but are also found in nonclinical populations as a response to negative life events (Craske et al. 1995, Gold et al. 2005, Meiser-Stedman et al. 2012, Reynolds & Brewin 1998). For instance, Gold et al. (2005) compared the intensity of psychological distress and trauma symptoms among 450 adults between 2 groups (1. after a non-traumatic negative event and 2. after a traumatic negative event) and found that the levels of psychological distress were the same in both groups. Similar findings were obtained in studies in which a high prevalence of depressive and anxious episodes and a high frequency of intrusive thoughts were found among healthy subjects after negative events (Craske et al. 1995, Meiser-Stedman et al. 2012, Reynolds et al. 1998). The differences between clinical and nonclinical populations are mainly in the intensity and stability of the symptoms that are mediated by each individual’s mood stability (Bowen et al. 2012).

On the basis of the common behavioral features associated with psychological distress and the most common psychological disorders, Barlow et al. (2004) identified 3 principles in helping clients with avoidance-based problems: 1. modifying the action tendencies associated with a specific emotion, 2. increasing the sense of controllability and predictability, and 3. decreasing the focus on the non-task-related consequences of excessive emotional activity. The first principle focuses on preventing avoidance behavior (e.g., suppression, rationalization, observation of body sensations, repetitive behaviors) by initiating behavior that exposes the client to his fears while he performs non-avoidant behaviors (Barlow et al. 2004). This behavior can be initiated in counseling conversations or outside of the counseling practice. The second principle focuses on gaining control by intentionally examining the feared catastrophe (Barlow et al. 2004). The third principle focuses on reducing the avoidance of emotional experiences by reinforcing their acceptance (Barlow et al. 2004).

1.3.3 Measurement of client emotions

When examining the evidence of the significance of emotions in counseling and therapy, it can be noted that valid measures of emotions can be used as indicators of counseling effectiveness (Sloan & Kring 2007). Earlier studies show that emotions have been measured with a variety of self-report, observational and psychophysiological measures (Sloan & Kring 2007). Typically, the emotion measures used in counseling and psychotherapy studies have focused on clients’ self-
reports of their emotions after counseling sessions (Kaplan et al. 2013, Sloan & Kring 2007). Some self-report measures can be used during counseling, for example, subjective unit of distress (SUD) ratings and the Self-Assessment Manikin (SAM), and these were developed in the context of exposure therapy (Kaplan et al. 2013, Sloan & Kring 2007). Additionally, electronic feedback devices have been used during client sessions (Kaplan et al. 2013, Sloan & Kring 2007). However, the use of self-report measures during conversational counseling is problematic because of the client’s need to focus on two tasks (Sloan & Kring 2007, Hill & Lambert 2004).

Often, counselors have not been willing to integrate psychophysiological measurement into natural settings (Sloan & Kring 2007). Instead, psychophysiological measures have been often used when comparing brain or autonomic responses between nonclinical and clinical groups in out-of-session settings or during specific tasks (e.g., Sloan & Kring 2007, Roffman et al. 2005, Baxter et al. 1998, Oathes et al. 2008). Many of these studies have supported the hypothesis that counseling and psychotherapy produce changes in clients’ nervous systems with pretest-posttest comparisons. For example, a study by Roffman et al. (2005) reviewed 14 neuroimaging studies on psychotherapy outcomes in which changes in the brain activity of depressed and anxious patients were examined. Their results showed similar changes among depressed and anxious patients as a result of therapy (Roffman et al. 2005). Furthermore, changes in autonomic activity during exposure therapy have been measured in studies that have shown graded decreases in sympathetic activity when clients are focused on feared stimuli (e.g. Grayson et al. 1982, Grey et al. 1981, Grey et al. 1979, Watson et al. 1972).

However, a small number of studies exist in which client emotions have been measured with autonomic measures in conversational settings during sessions (Yuan et al. 2010, Levenson & Rueff 1997, Levenson & Gottman 1983, Stanek et al. 1973). These studies have focused on husband-and-wife relationships in couples therapy or on therapist and client emotions during sessions. In this context, the covariance of the dyads’ emotional variables has been interpreted as an indicator of emotional synchrony connected to the development of alliance and agreement (Yuan et al. 2010, Levenson & Rueff 1997, Levenson & Gottman 1983, Stanek et al. 1973).

Although in-session measurements of emotions have been uncommon, they are crucial for obtaining information on emotional changes during counseling sessions (Hill 2004, Sloan 2007). Furthermore, more accurate response-level measurements are needed when comparing counselor responses and client emotions during sessions, but to the best of my knowledge, no studies have been conducted in this area. When examining emotions during counseling sessions, measurements are needed during
clients’ verbal reports of troubling experiences (Hill & Lambert 2004). However, results from recent studies have shown that verbal reporting might have a regulatory influence on clients’ emotions (Campbell-Sills et al. 2006, Foland et al. 2008, Lieberman et al. 2007, Hariri et al. 2000, Hariri et al. 2003; Tabibnia et al. 2008). Although these findings may question the way measurements of emotions during verbal reporting should be interpreted, to my knowledge, the effect of verbally reporting troubling or avoided experiences on client emotions during counseling sessions has not yet been examined. Instead, this notion has gained support from the results of recent experiments indicating that the verbal reporting of unpleasant experiences is associated with positive emotional changes in the subject (Campbell-Sills et al. 2006, Foland et al. 2008, Lieberman et al. 2007, Hariri et al. 2000, Hariri et al. 2003, Tabibnia et al. 2008). In a study by Hariri et al. (2003), 11 healthy subjects’ brain activity was measured with blood-oxygen-level dependent functional magnetic resonance imaging (BOLD fMRI) while they were viewing and labeling unpleasant pictures from the International Affective Picture System (IAPS). The results showed that viewing unpleasant stimuli was associated with a bilateral amygdala response and that labeling IAPS stimuli was associated with attenuation of this amygdala response and increased activity in the right prefrontal cortex and the anterior cingulate cortex, indicating regulatory influences on subjects’ emotions (Hariri et al. 2003).

Similar results have been found with autonomic measures (Campbell-Sills et al. 2006, Eifert & Heffner 2003, Hariri et al. 2003, Hoffman et al. 2009, McRae et al. 2009, Tabibnia et al. 2008). For example, in a study by Tabibnia et al. (2008), 27 healthy adults were exposed to aversive images with and without labeling while their skin conductance and heart rate were measured as indicators of autonomic reactivity. Exposure to images with labeling led to more attenuation of autonomic reactivity than did exposure alone (Tabibnia et al. 2008). Affective labeling may help to dampen autonomic reactivity, as indicated by skin conductance and heart rate deceleration (Tabibnia et al. 2008). Similar results were obtained by McRae et al. (2009) in their study, in which they measured the autonomic reactivity of 40 healthy subjects with skin conductance while the subjects were viewing emotional IAPS pictures. They found that skin conductance values were lower during the verbal labeling of the pictures than during their viewing (McRae et al. 2009). Furthermore, studies on the acceptance vs. suppression of emotions have indicated higher autonomic responses and more negative subjective ratings of emotional stimuli when emotions are suppressed (Campbell-Sills et al. 2006, Hoffman et al. 2009, Eifert et al. 2003). For example, Campbell-Sills et al. (2006) studied 60 subjects divided into acceptance and
suppression groups while they were exposed to emotional film clips. The acceptance group had lower heart rates and less-anxious subjective experiences of the films than did the suppression group (Campbell-Sills et al. 2006).

Although brain activity and skin conductance during the reporting of unpleasant emotional experiences have been studied, only a small number of studies have measured cardiac responses while clients labeled or reported unpleasant experiences (Campbell-Sills et al. 2006, Tabibnia et al. 2008). For example, in a study by Tabibnia et al. (2008), 27 healthy adults were exposed to aversive images with and without labeling while their heart rates were measured. Exposure to images with labeling led to more attenuation of autonomic reactivity than did exposure alone (Tabibnia et al. 2008). However, to my knowledge, spectral analysis of heart rate variability (HRV), which can be used to extract the sympathetic and parasympathetic components of cardiac response, have not yet been examined regarding data on verbally labeling or reporting unpleasant experiences. This may be attributable to the problems connected to HRV measurement in interactional settings, in which the control of respiratory effects is critical in obtaining reliable estimates of HRV powers (Beda et al. 2007).

The respiratory parameters that can influence HRV powers are respiratory frequency (Schipke et al. 1999, Pöyhönen et al. 2004), tidal volume (Pöyhönen et al. 2004), end-tidal carbon dioxide partial pressure (Henry et al. 1998; Pöyhönen et al. 2004), expiration/inspiration time ratios (Strauss-Blasche 2000) and respiratory dead space (Furutani et al. 1997). Control of breathing with a mask or mouthpiece is the only way to obtain all of these parameters. Other ways to gain partial control of respiratory effects are metronomic breathing or the use of plethysmographic recordings for obtaining respiration frequency. However, the chest bands used in plethysmographic recordings are sensitive to motion artifacts and to band stability problems and cannot be considered highly accurate (Noponen et al. 2012). There is also evidence that the measurement of respiratory effects itself affects breathing patterns and HRV (Hirsch & Bishop 1982).

The parasympathetic high-frequency (HF) component of HRV is associated with a typical respiration frequency of 1 breathing cycle in 2- to 7-second intervals (0.14–0.5 Hz), and frequent deviation from this breathing frequency will affect results (Cerutti et al. 1995). Furthermore, speaking tasks can change respiration frequency and can especially cause low-frequency (LF) bias in spectral analysis (Beda et al. 2007). This is why the minimal control of respiratory effects requires estimating respiratory frequency. Thayer et al. (2002) developed a non-invasive method for obtaining respiratory frequency estimates from the HRV spectrum that used a peak HF value as an indicator of typical breathing frequency. This is the frequency point at
which most of the power is located in the HF area. Thayer et al. (2002) tested this technique in comparison with plethysmographic measurements and obtained highly accurate results. This method is best used in settings in which full-scale respiratory measurements are not possible (Thayer et al. 2002).

1.4 Critical measurement components for assessing counseling interaction

The results from counseling outcome research have shown that client-centered counselor behaviors that aim to guide the client to approach his troubling and avoided experiences are critical factors in successful client outcomes (Orlinsky et al. 1994, Orlinsky et al. 2004). Furthermore, substantial evidence has suggested that facilitating approach behaviors helps clients use adaptive strategies to regulate and create positive changes in their emotions (Barlow & Gross 2002, Muraven & Baumeister 2000, Borkovec et al. 1995, Brown et al. 1995, Davis et al. 2010, Newman et al. 2010). Despite these results, in-session measurements of counselor responses and client emotions have been uncommon (Cahill et al. 2007, Rantanen & Soini 2013, Hill & Lambert 2004, Sloan & Kring 2007). However, when developing evidence-based counseling and therapy, it is critical to make specific, response-level measurements of counselor actions, client emotions and their relationship to better understand what types of counselor actions are connected to positive emotional changes in clients (Hill & Lambert 2004, Sloan & Kring 2007). Fulfilling this goal requires valid observational measures that can be used to measure counselor responses, counselor client-centeredness and changes in clients’ emotions during sessions (Hill & Lambert 2004). By observing counselor actions and client emotions at a response level, it is possible to deepen our understanding of counseling interaction and to develop more effective counseling communication strategies (Sloan & Kring 2007).

When examining emotions during counseling sessions, measurements are needed while clients are verbally reporting troubling experiences (Hill & Lambert 2004). However, results from recent studies have shown that verbal reporting might have a regulatory influence on client emotions (Barlow et al. 2004). Although these findings may question how measurements of emotions during verbal reporting should be interpreted, to my knowledge, the effect of verbally reporting troubling or avoided experiences on client emotions during counseling sessions has not been examined to date. Instead, this notion has gained support from the results of recent experiments indicating that the verbal reporting of unpleasant experiences is associated with positive emotional changes in subjects (Campbell-Sills et al. 2006, Foland et al. 2008, ...
Lieberman et al. 2007, Hariri et al. 2000, Hariri et al. 2003, Tabibnia et al. 2008). These findings need closer examination and more experiments to confirm them, and they may have a fundamental influence on the way emotional reactions should be interpreted during verbal reporting in counseling sessions and in relation to counselor responses.
2 Aims of the study

The general purpose of this study was to develop methodology to assess counseling interactions during sessions by developing measures for assessing two key components in successful counseling interactions: client-centered counselor responses and client emotions. The first main objective of the research was to develop measures for assessing counselor responses and client-centeredness (Study I). Specific aims connected to the first objective were to assess the content validity and inter-rater reliability of the Counselor Response Coding System (CRCS) and the content validity, construct validity and inter-rater agreement of the Skilled Verbal Responding Scale (SVRS), which are measures included in the Counselor Response Observation System (CROS). The second main objective was to assess the applicability of heart rate variability (HRV) to the measurement of emotions during clients’ verbal reports of unpleasant experiences in a setting consistent with client-centered counseling aims (Studies II and III). The setting was constructed to facilitate emotions in the subjects using emotional pictures and to guide the subjects to verbally report their observations and thoughts associated with the experience. The specific aims connected to the second objective were to examine the differences in subjects’ emotions between viewing and reporting unpleasant experiences and to examine the differences in their emotions between reporting pleasant and unpleasant experiences with HRV. The data from studies II and III were analyzed as one data set in this thesis.
3 Methods

In this chapter, I will present the methods of the study. First, I will describe the methods that were used in the development of the CROS. Finally, I will describe the methods that were used in the HRV experiment.

3.1 Development of the CROS

In this chapter, I will present the methods used in the development of the CROS. First, I will give a detailed description of the participants and the research procedure used in the study. Finally, I will present the CROS measures that were tested.

3.1.1 Participants

Participants in the reliability testing of CROS instruments consisted of 6 educational psychology master’s students participating in an advanced counseling research course. Participants in the content validity testing consisted of 3 researchers, 2 Ph.D. students and 1 Ph.D. researcher, conducting counseling research, counselor training and counseling work at the clinic of educational psychology. The source data for the reliability testing consisted of video recordings taken from 85 counseling sessions. Eighty-five educational psychology master’s students in a counseling psychology course participated in these counseling sessions as counselors and clients in authentic sessions in which clients’ true worries and problems were addressed. These counseling sessions were supervised by 1 Ph.D. researcher who was experienced in counselor training.

3.1.2 Procedure

The 3 participants in the content validity testing evaluated the content validity and the operational criteria of the instruments by examining the adequacy of the CRCS categories and the SVRS items in relation to their purposes. The 6 participants in the reliability testing received 16 hours of training in using the CRCS and SVRS instruments. After participants were trained, the inter-coder reliability of the CRCS was tested using videoed data from counseling sessions. The 6 participants coded 30 counselor responses from a single counselor in a counseling session. The inter-coder reliability was tested using Fleiss’ kappa (Fleiss 1971), which is a statistical measure that is appropriate for testing observer agreement among multiple observers when
making assessments within a multiple-category system. Sample size adequacy for kappa was assessed by using minimum standard error for 6 raters and 30 targets, .05 α-level, kappa = .70 (Fleiss 1971). Estimation showed that 6 raters and 30 targets were sufficient to obtain kappa of 0.70 at .05 alpha level.

The inter-rater reliability of the SVRS was tested using an ICC (Shrout & Fleiss 1979). The test targets consisted of 30 videoed counseling sessions that were 30–40 minutes long and were rated by the 6 observers. Sample size adequacy was tested with a power analysis for ICC (Walter et al. 1998). Power was estimated by testing the hypothesis for an ICC > .50, which is considered a large effect according to Cohen’s criteria (1988). Power was estimated for 6 raters and 30 targets, .05 α-level, H0: ICC = .50, and H1: ICC = .70, which resulted in an adequate power of .85.

The construct validity of the SVRS was assessed using confirmatory factor analysis (Jöreskog 1969), and model fit was tested using a χ² test, the comparative fit index (CFI), the goodness of fit index (GFI) and the root mean square error of approximation (RMSEA). This analysis was executed on SVRS scores obtained from the one trained observer’s assessments of the 85 videoed counselor performances. The maximum likelihood (ML) estimation method was used in this analysis. Although the general least squares estimation method would possibly have been more accurate for small samples, ML was chosen to obtain better power estimates. Sample size adequacy was estimated using a cases-per-parameter ratio (Bentler & Chou 1987) and a power analysis for GFI and RMSEA (MacCallum, Browne & Sugawara 1996, MacCallum & Hong 1997). In the construct validity testing, data from 85 videoed counseling sessions in which 85 different counselors and clients were used and 14 parameters were estimated. This produced a ratio of 6 cases per 1 parameter, which is considered an adequate ratio for confirmatory analysis (Bentler & Chou 1987). Power analysis for GFI and RMSEA was performed using Cohen’s .80 criterion for adequate power. The power analysis for GFI was performed with a .05 α-level, df = 14, H0: .95, H1: .90, which resulted in a power of .79. The power analysis for RMSEA was performed with .05 α-level, df = 14, H0: .04, H1: .10, which resulted in a power of .41. The sample size for construct validity testing fulfilled Bentler and Chou’s criteria (1987) and was very near adequate for GFI power criteria. However, the RMSEA power was below the adequate .80 limit because of the small sample size; it would have required n=190 to be fulfilled. Our data were limited to 85 subjects, and obtaining large additional data from counseling sessions was not possible at the time of this study. Because of the small sample size, confirmatory factor analysis was considered to be a starting point in the confirmation of the SVRS structure.
3.1.3 Measures

The CRCS is a systematic observation and coding method that was developed initially from Ivey’s micro-counseling theory (1971, 2007), and it enables highly accurate behavior quantification from videoed counseling sessions. A detailed description of the pretesting procedures and of the CRCS’s development can be found in the original publication (Study I). The CRCS is applied by coding each of the counselor’s speech sequences in relation to the client’s messages. The counselor’s speech sequence or response was determined as follows: from the beginning of the counselor’s speech to the beginning of the client’s speech. The CRCS was developed to obtain the response frequencies of an entire counseling session, but it also enables the creation of time-adjusted matrices after measurement by selecting the speech sequences needed for the analysis. Total frequency measures per subject are used in the statistical analysis of the obtained data. The CRCS instrument is composed of two categorical variables. The first component of the CRCS is the counselor response variable. Counselor response comprises 5 categories: reflection, specifying question, conclusion, suggestion and self-disclosure. The response variable can be used to explore what the counselor is doing in relation to the client’s message and to what degree. The second component in the CRCS is the counselor response focus variable. The foci in the CRCS comprise 5 categories: feelings, explanations, actions, context and strengths. Counselor response focus can be used to determine which speech content the counselor is focusing on. The response focus indicates how many times the counselor directs the conversation to specific contents during the counseling session. The prevalence of the CRCS categories was assessed in the pretesting phase to create a balanced structure that would produce observations. The minimum prevalence limit for each category was set at 1 out of 30 observations. A detailed description of the pretesting phase can be found in the original article (Study I). Furthermore, detailed presentations of the CRCS categories, operational definitions and scoring procedure are provided in Appendix A.

The SVRS measures a counselor’s skilled responding based on videoed counseling sessions. It was developed from Okun’s (2002) and Truax and Carkhuff’s (1967) ideas of facilitative, client-centered counseling, which were heavily based on the results of counseling outcome research (Orlinsky et al. 1994, 2004). Counseling in this context is defined as an action in which the counselor helps the client to explore and understand his problem, and therefore, it is of utmost importance that the counselor focus on the client’s message in a personal, consistent and thorough way (Okun 2002, Truax & Carkhuff 1967). On the SVRS, these qualities are considered to
be the foundation of skillful responding. These ideas were described in chapter 1.1. The SVRS is scored based on an entire session and is composed of 7 observational items, all of which are rated on a 4-point Likert scale (0 = not at all, 1 = a little, 2 = somewhat, 3 = a lot). More detailed scoring criteria are presented in Appendix A. The SVRS is applied after the counseling session has been observed. In short, the SVRS measures the degree to which client-centered aspects are fulfilled in the counselor's actions.

3.2 Assessment of applicability of HRV to the measurement of emotions during verbal reports

In this chapter, I will present the methods used in the assessment of HRV during verbal reporting. First, I will give a detailed description of the participants used in the study. Finally, I will describe materials and measurement equipment applied and the experimental setting and procedure used in the study. The setting of the experiment was analogous with client-centered counseling aims and was constructed to facilitate emotions in the subjects using emotional pictures and to guide the subjects to verbally report their observations and thoughts associated with the experience. Studies II and III are part of a larger research project that focuses on multimodal emotional analysis using highly non-invasive measures; thus, the experiment consisted of multiple measures including HRV, eye movement, pupil size, facial expression and speech signal data. This study focuses only on the HRV data obtained from the experiment, but I will describe the full setting with its measurements to create an overall picture of the experiment.

3.2.1 Participants

A total of 39 right-handed, (native fluency in Finnish and no reported history of speech disorder) Finnish-speaking undergraduate students (25 females and 14 males) from the University of Oulu participated in the experiment. The students participated voluntarily in the experiment during their psychology studies. Eyesight was tested using the Snellen card, and anxiety was measured using the State Trait Anxiety Inventory (STAI) (Spielberger et al. 1970). In addition, possible alexithymia was tested with the TAS-20 (Bagby & Taylor 1994), the validity of which has been tested in Finland as well (Joukamaa et al. 2001). All subjects had normal eyesight (≥ 1.0), and neither anxiety (STAI score < 35) nor alexithymia (TAS-20 score < 51) were
found. After the explanation of the experimental protocol, the subjects gave written consent.

3.2.2 Apparatus and materials

The IAPS pictures (International Affective Picture System) (Lang et al. 2005) were presented on the screen (17”) of a computer with an Intel Pentium 4 processor that was connected to a Tobii 1750 eye tracking system (Tobii Technologies AB, Sweden). The sample rate was 50 Hz, and the spatial resolution was 0.25 degrees. The eye tracking system located every fixation point and measured the duration of fixation, the pupil size variation and the distance of the eye from the computer screen. The heart rate variations were measured using beat-to-beat RR intervals with a Polar S810i heart rate monitoring system (Polar Oy, Finland). The Polar monitor was chosen because the overall research project aims to apply highly non-invasive measurements in natural settings. The Polar S810i has been validated for R spike recognition with electrocardiographic recordings (ECG) in multiple studies (Gamelin 2006, Weippert et al. 2010, Vanderlei et al. 2008). Facial expressions were recorded with an IEEE 1394 firewire camera (Sony DFW-VL500, Japan). In addition, each subject’s speech was recorded using a wireless microphone system (Sennheiser HSP2, Denmark). A total number of 48 International Affective Pictures (Lang et al., 2005) were used in the experiment. The pictures were divided into three different groups: 16 pleasant, 16 neutral and 16 unpleasant pictures (see also Nummenmaa et al. 2006). The overall luminance levels of the pictures were adjusted with Adobe Photoshop 6.0 software.

3.2.3 Procedure

The subjects were interviewed and the STAI (Form 2) and TAS-20 questionnaires were presented before the experiment. Subsequently, each subject was able to practice the experimental procedure from the paper version with the experimenter. Thereafter, the subjects practiced the procedure with the computer. Before the actual experiment, each subject rested for 60 seconds while the heart rate monitoring, audio and camera systems were combined with the eye tracking system. The subject’s eye movements were also calibrated into the eye tracking system. Respiratory parameters connected to HRV measurements were not directly measured in this study. Instead, Thayer's et al. (2002) method was used to estimate respiratory frequencies from the power spectrums.
The experimental procedure is presented in Figure 1. In the experiment, the pictures were presented on the computer screen and the distance of the subject from the screen was 65 cm. At first, the subject had to look at the letter X, which appeared in the middle of the screen, for 30 seconds. Sequentially, either a pleasant, a neutral or an unpleasant picture appeared on the screen for 20 seconds in random order.

Immediately after the 20 seconds, the SAM scale (Self-Assessment Manikin, Lang 2005) appeared. The subject’s task was to orally report the valence and arousal of the picture according to the SAM scale (1–9 categories). These categories were grouped according to their valences, giving new categories: “pleasant” (1–3), “neutral” (4–6), and “unpleasant” (7–9). This grouping was made to improve comparison with the IAPS pictures (pleasant, neutral or unpleasant).

After the report, the subject had to press the enter button to darken the screen. In this phase, the subject’s task was to orally report what had been observed, what was happening and what was going to happen in the picture to the experimenter, who was sitting behind the computer screen. After the report, the subject had to press the enter
button for the next picture to appear. After 48 pictures, the letter X appeared for 30 seconds. Finally, the STAI (Form 1) questionnaire was presented. The experimentation was approved by the ethics committee of the Faculty of Education, University of Oulu (see also Laukka et al. 2008).

The HRV data from the viewing and reporting phases with the pleasant and unpleasant pictures were analyzed. The HRV data that related to the pleasant and unpleasant pictures in different phases were obtained using a MATLAB 7.12.0.635 (Mathworks, USA) script. The script was constructed to compare the timelines of the Tobii eye tracker system matrix, which contained time indices for different conditions, with the HRV timeline. The timelines were synchronized using the HRV timestamp recorded in Tobii video. The obtained RR data were examined for abnormal beats, and any found artifacts were removed. Four subjects were removed from the HRV analysis because of missing data. The HRV analysis was executed using Kubios HRV analysis software (University of Eastern Finland). The data were interpolated to 4 Hz to fulfill the requirement for spectral analysis. Spectral analysis applying the Fast Fourier Transform was performed on the data using Welch’s (1967) periodogram (window size 256, overlap 50%) to extract the powers of the spectral components relating to sympathetic and parasympathetic activity per subject in each condition. Variance analysis for repeated measures (2x2 ANOVA) was used to compare HRV differences between the pleasant and unpleasant pictures. Power analysis for repeated measures ANOVA was performed with .05 α-level, df1 = 2, df2 = 102, η² = .25, which resulted in a power of .90. The normality of dependent variable in each condition was examined, and significant outliers were filtered out (|z| ≥ 2.5). Sphericity was assumed because there were only 2 levels in each factor. Estimates of respiratory frequency were obtained with Thayer’s et al. (2002) method using HF power peak location as an estimate of respiratory frequency. This method has been validated with plethysmographic recordings (Thayer et al. 2002). The mean values of respiratory frequency estimates in each condition were: 0.26 Hz (viewing pleasant pictures), 0.27 Hz (viewing unpleasant pictures), 0.25 Hz (reporting on pleasant pictures), and 0.25 Hz (reporting on unpleasant pictures). A criterion for acceptable respiratory frequency was set to HF peak location greater than 0.153 Hz based on the recommendation by Thayer et al. (2002) to place HF peak clearly over the border value between the components. Three of 140 respiratory frequency estimates were found to be lower than the criterion, indicating possible low-frequency bias in the results connected to these conditions and subjects. These cases were filtered out.
4 Summary of the results

In this chapter, I will present summary of the results. First, I will describe the results from the validity and reliability analyses of the CROS. Finally, I will describe the results from the analyses of the HRV experiment.

4.1 Development of the Counselor Response Observation System

In this chapter, I will present the results from the validity and the reliability tests of the CROS measures. First, I will describe the results from the content validity and inter-rater reliability tests of the CRCS. Finally, I will describe the results from the content validity, inter-rater reliability and construct validity tests of the SVRS.

4.1.1 Content validity and inter-rater reliability of the CRCS

The validity of the CRCS variables and their categories was evaluated through expert review of the content and criteria of the coding system. The experts were 3 researchers, two Ph.D. students and one Ph.D. researcher, who conducted counseling research, counselor training and counseling work at the educational psychology clinic. The experts evaluated both the relevance of the CRCS categories for describing counselor responses and the consistency of the categories and their criteria. All categories and their criteria were considered to be essential to the coding system.

The reliability of the CRCS was assessed with Fleiss’ kappa, which is a statistical measure of inter-observer agreement between multiple observers assessing data within a multiple-category system (Fleiss 1971); 6 observers rated 30 targets. The significance of the obtained kappa values was interpreted according to criteria set forth by Landis and Koch (1977). Fleiss’ kappa was calculated separately for the counselor response and counselor response focus variables. For both response and response focus, the Fleiss’ kappa values showed substantial, $\kappa > .70$, inter-observer agreement, meeting the criteria set by Landis and Koch (1977). The counselor response variable had high observed agreement, .91, and a high kappa value, $\kappa = .86$, SE = .05, 95% CI [.77, .94]. The counselor response focus variable had high observed agreement, .97, and a high kappa value, $\kappa = .96$, SE = .02, 95% CI [.92, .99]. Both CRCS variables indicated high reliability.
4.1.2 Content validity and inter-rater reliability of the SVRS

The content validity of the SVRS was evaluated by 3 researchers. The researchers evaluated the relevance of the SVRS items in relation to the main construct (skilled verbal responding). All items were considered essential for the scale. The reliability of the SVRS items was tested by 6 observers using ICC (Shrout & Fleiss 1979) and Cronbach’s alpha (Cronbach 1951). Observer agreement was assessed using ICC for a single measure and absolute agreement. Statistical significance was tested for an ICC > .50, which can be considered a large effect size according to Cohen’s criteria (1988). All items had high and statistically significant ICC values, ICC > .71, p < .01; these are presented in Table 2. The SVRS items were also tested for scale consistency using Cronbach’s alpha. The SVRS scale showed high internal consistency, a = .86, which meets Nunnally’s (1978) criterion of .70 for acceptable alpha.

Table 2. Intra-class correlation coefficients of the SVRS items.

<table>
<thead>
<tr>
<th>SVRS items</th>
<th>ICC</th>
<th>95 % CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The counselor speaks of things that the client has not mentioned</td>
<td>.95 ***</td>
<td>.91, .97</td>
</tr>
<tr>
<td>2. The counselor uses open questions</td>
<td>.80 ***</td>
<td>.70, .89</td>
</tr>
<tr>
<td>3. The counselor focuses on exploring the client's problem</td>
<td>.91 ***</td>
<td>.86, .95</td>
</tr>
<tr>
<td>4. The counselor does not share opinions or give advice</td>
<td>.85 ***</td>
<td>.76, .91</td>
</tr>
<tr>
<td>5. The counselor does not act mechanically</td>
<td>.91 ***</td>
<td>.85, .95</td>
</tr>
<tr>
<td>6. The counselor uses specifying questions</td>
<td>.71 **</td>
<td>.58, .83</td>
</tr>
<tr>
<td>7. The counselor ignores the client's message</td>
<td>.78 **</td>
<td>.67, .87</td>
</tr>
</tbody>
</table>

*** p < .001, ** p < .01, H0: ICC = .50
Construct validity was tested with a confirmatory factor analysis (Jöreskog 1967) for a one-factor model using AMOS 19.0 software (IBM). Data screening was performed for univariate and multivariate normality. Univariate normality was tested by examining histograms, standardized skewness, standardized kurtosis and outliers of the variables. The histograms of the variables indicated approximately normal distributions. Items 1, 2, 3, 4, 5, and 7 were not significantly skewed (|z| < 2.0, p > 0.05, two-tailed). Item 6 was slightly skewed, |z| = 2.39, p > .01, two-tailed, but was within an acceptable range of normal. None of the variables had significant kurtosis, |z| < 2.0, p > .05, two-tailed, or had outliers, |z| < 2.57, p > .01, two-tailed. Multivariate normality was assessed with Mardia’s coefficient of multivariate kurtosis, which was found to be non-significant, c.r. = .301, p > .05, two-tailed. Multivariate outliers were assessed using the Mahalanobis distance, which was below the critical value, $\chi^2 (14, N = 85) < 36.12, p < .001$. Linearity was confirmed by examining scatter plots. Multicollinearity and singularity were assessed by examining the determinant of the covariance matrix. The assessment indicated problems with multi-collinearity and singularity (det. = 0). Many of the standardized covariance residuals had values near 2 or more, indicating that the model might not be adequate. All factor loadings were large, $\beta > .60$, and statistically significant, p < .001. Model fit was tested using $\chi^2$, CFI, GFI and RMSEA. The criteria for acceptable model-fit values were as follows: $\chi^2$ was non-significant at a .05 threshold (Barrett, 2007), CFI > .95 (Hu & Bentler 1999), GFI > .95 (MacCallum & Hong 1997) and RMSEA < .10 (MacCallum 1996). All indices indicated that the model was inadequate: $\chi^2 (14, N = 85) = 110.66, p < .001$; CFI = .66; GFI = .71; RMSEA = .29, 90% CI [.24, .34].

Modification indices were examined, and they suggested that the model could be improved by allowing residuals d1 and d4 to correlate. However, this suggested modification would not have been adequate or useful for our construct. The covariance matrix of the model was examined to determine the causes of the inadequacy and to improve the model. Item 5 (the counselor does not act mechanically) of the SVRS had a low correlation with items 1 and 6 and was removed.

Large residuals indicated problems with the structure of the model; therefore, a 2-factor model was created. The factors were constructed by grouping items that strongly correlated with each other and had similar content. Factor 1: Client-centeredness was composed of items 1, 2 and 4 (1. The counselor speaks of things that the client has not mentioned, 2. The counselor uses open questions, and 4. The counselor does not share opinions or give advice). Factor 2: Consistency was
composed of items 3, 6 and 7 (3. The counselor focuses on exploring client’s problems, 6. The counselor uses specifying questions and 7. The counselor ignores the client’s message). Sample size adequacy was determined again because of the changes in the SVRS model. The modified model comprised 85 cases and 13 parameters, producing a 6.5:1 ratio that fulfilled Bentler and Chou’s criteria (1987) for adequate sample size. The power for GFI reduced to .78 and the power for RMSEA reduced to .30, both owing to the decrease in degrees of freedom. Data screening was performed for univariate and multivariate normality. Univariate normality was tested by examining histograms, standardized skewedness, standardized kurtosis and outlier of the variables. The histograms of the variables indicated approximately normal distributions. Items 1, 2, 3, 4, and 7 were not significantly skewed, |z| < 2.0, p > .05, two-tailed. Item 6 was slightly skewed, |z| = 2.39, p > .01, two-tailed, but was within an acceptable range of normal. None of the variables had significant kurtosis, |z| < 2.0, p > .05, two-tailed, or had outliers, |z| < 2.57, p > .01, two-tailed. Multivariate normality was assessed with Mardia’s coefficient for multivariate kurtosis, which was found to be non-significant, c.r.= .114, p > .05, two-tailed. Multivariate outliers were assessed with the Mahalanobis distance, which was below the critical value, χ² (8, N = 85) < 26.13, p = .001. Linearity was confirmed by examining scatter plots. Multi-collinearity and singularity were assessed by examining the determinant of the covariance matrix. The covariance matrix indicated no multi-collinearity or singularity (det. > 0). The standardized covariance residuals were between -1.379 and 0.692, indicating good fit. Factor loadings are presented in Table 3. All loadings were large and statistically significant, β > .68, p < .001, and the 2 factors correlated as hypothesized, r = .54, p < .001. The model fit was tested using χ², CFI, GFI and RMSEA. The criteria for acceptable model-fit values were as follow: χ² was non-significant at a 0.05 threshold (Barrett, 2007), CFI > 0.95 (Hu & Bentler 1999), GFI > 0.95 (MacCallum & Hong 1997) and RMSEA < 0.10 (MacCallum 1996). All indices indicated that the model was a good fit, χ² (8, N = 85) = 9.454, p = .305; CFI = .99; GFI = .96; RMSEA = .05, 90% CI [.00, .14]. The complete 2-factor model is presented in Figure 2. Cronbach’s alphas were computed for the new subscales. The alpha values were .85 for factor 1 and .81 for factor 2.
<table>
<thead>
<tr>
<th>Factor</th>
<th>Item</th>
<th>Factor loadings</th>
<th>Standardized factor loadings</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1.00 (fixed)</td>
<td>.89</td>
<td>.79</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>.68 ***</td>
<td>.75</td>
<td>.56</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>.89 ***</td>
<td>.80</td>
<td>.64</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>1.00 (fixed)</td>
<td>.68</td>
<td>.47</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>1.09 ***</td>
<td>.76</td>
<td>.58</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>1.41 ***</td>
<td>.88</td>
<td>.77</td>
</tr>
</tbody>
</table>

*** p < .001, two-tailed
Fig. 2. The 2–factor SVRS model showing the standardized factor loadings and multiple correlation coefficients of the items.
4.2 HRV during the oral reporting of unpleasant experiences

The HRV dynamics as subjects viewed and reported on pleasant and unpleasant pictures were assessed with a repeated-measures ANOVA. A 2 (Phase: Viewing vs. Reporting) x 2 (Valence: Pleasant vs. Unpleasant) ANOVA was performed on the HRV power components. The ANOVA results for LF power indicated that neither the valence factor, i.e. F(1,30) = 0.332, partial $\eta^2 = .01$, nor the phase factor, i.e. F(1,30) = 2.41, p = .13, partial $\eta^2 = .07$, had a statistically significant impact on LF power. However, the interaction of the valence and phase factors had a statistically significant (F = 6.23, p = 0.02, $\eta^2 = .17$) effect on LF power. The ANOVA results for HF power showed that neither the valence factor, i.e. F (1, 30) = 0.60, p = .44, partial $\eta^2 = .02$, nor the phase factor, i.e. F (1, 30) = 1.57, p = .22, partial $\eta^2 = .05$, had a statistically significant impact on the HF power. However, the interaction of the valence and phase factors had a statistically significant (F = 5.76, p = 0.02, $\eta^2 = .16$) effect on HF power. The ANOVA results for LF/HF power indicated that neither the valence factor, i.e. F(1,28) = 2.50, p = .125, partial $\eta^2 = .08$, nor the phase factor, i.e. F(1,28) = 3.34, p = .08, partial $\eta^2 = .11$, had a statistically significant impact on LF/HF power. However, the interaction of the valence and phase factors had a statistically significant F (1, 28) = 9.59, p = 0.00, $\eta^2 = .26$ effect on LF/HF power.

The nature of the interactions is presented in Figures 3a, 3b and 3c with the mean LF/HF values for each condition. Detailed LF and HF values in each condition can be found in Appendix B. Figure 3a shows that the LF/HF power was higher as the subjects viewed unpleasant pictures than it was when they reported on them. In contrast, figure 3b shows that LF/HF power was lower as the subjects viewed pleasant pictures than it was when they reported on them. Figure 3c, in its turn, shows that LF/HF power was higher as the subjects reported pleasant pictures than it was when they reported on unpleasant pictures. Comparison of figures 3a and 3b show that viewing of unpleasant stimuli is associated with higher LF/HF than viewing pleasant stimuli indicating higher sympathetic reactivity to unpleasant than pleasant stimuli. Overall, our results indicated that as the subjects reported on unpleasant pictures, sympathetic activity was lower and parasympathetic activity was higher when compared with their viewing of unpleasant pictures or their reporting on pleasant pictures. Furthermore, the results indicated that during the subjects’ reporting on pleasant pictures, sympathetic activity was higher and parasympathetic activity was lower compared with their viewing on pleasant pictures. Overall, the results indicate that the emotional dynamics while viewing or reporting on unpleasant pictures were the opposite of the dynamics observed with pleasant pictures.
Fig. 3. LF/HF during viewing and reporting of pleasant and unpleasant stimuli.
5 Discussion

The general purpose of this study was to develop methodology for assessing counseling interactions during counseling sessions. In the development work, the focus was on two key components of successful interactions: client-centered counselor responses and client emotions. The first main objective of the research was to develop measures for assessing counselor responses and their client-centeredness (Study I). Specific aims connected to the first objective were to assess the content validity and inter-rater reliability of the Counselor Response Coding System (CRCS) and the content validity, construct validity and inter-rater agreement of the Skilled Verbal Responding Scale (SVRS), which are measures included in the Counselor Response Observation System (CROS). The second main objective was to assess applicability of HRV to measuring emotions during client verbal reports in a setting consistent with client-centered counseling aims (Studies II and III). Specific aims connected to the second objective were to examine the differences in subjects’ HRV and in their emotions between their viewing and reporting on unpleasant pictures and between their reporting on pleasant and unpleasant experiences. The data from studies II and III were analyzed as one data set in this thesis.

The results from study I supported the content validity and inter-rater reliability of the CRCS and the content validity and inter-rater reliability of the SVRS. However, because of the small sample size in the confirmatory factor analysis, the SVRS construct validity results are better interpreted as preliminary. Overall, the results from the CROS testing were considered to be a good starting point in the future development of the system. The results from the integrated analysis of the HRV data from studies II and III were consistent with earlier findings indicating that sympathetic activity was higher and parasympathetic activity lower as clients viewed unpleasant pictures than it was as they reported on them and that sympathetic activity was lower and parasympathetic activity higher when subjects were reporting on unpleasant pictures than when they were reporting on pleasant pictures. These findings indicate that emotional reactions change during reporting and that it may be necessary to interpret emotions during reporting in relation to initial emotional reactions during viewing. A detailed examination of these findings is discussed in chapters 5.1 and 5.2.
5.1 The Counselor Response Observation System

The results supported the content validity and inter-rater agreement of the CRCS and the SVRS, which are measures included in the Counselor Response Observation System (CROS). The content validity of the CRCS variables and their categories were evaluated by expert review of the content and criteria of the coding system. Category content was developed from Ivey’s (1971) micro-counseling model with improvements on the structure; specifically, response and its focus were separated into their own variables. Furthermore, the prevalence of the CRCS categories was assessed in the pretesting phase to create a balanced structure. The experts evaluated both the relevance of the CRCS categories in describing counselor responses and the consistency of the categories and their criteria. All categories and their criteria were considered to be essential to the coding system by the researchers. The content of the CRCS categories was built on an earlier structure that has been widely used and accepted (Ivey & Ivey 2007) and that was assessed by counseling researchers; thus, its content validity can be concluded to be sufficient. The inter-rater reliability of the CRCS was assessed with Fleiss’ kappa (Fleiss 1971). The results indicated high inter-rater reliability for counselor response and counselor response focus, indicating that trained educational psychology master’s students can produce the same observations from the same observed counselor actions. Although the coverage of the CRCS categories was assessed in the pretesting phase, it should be further tested in different settings to assess its accuracy.

The content validity of the SVRS was evaluated by three researchers. The SVRS items were strongly built on the results from process-outcome research (Orlinsky et al. 1994, Orlinsky et al. 2004) and Okun’s (2002) helpful responding model. The researchers evaluated the relevance of the SVRS items in relation to the main construct, and all items were considered essential for the scale. The content of the SVRS was built on client-centered behaviors supported by the findings from process-outcome research (Orlinsky et al. 1994, Orlinsky et al. 2004) and was assessed by counseling researchers; thus, its content validity can be concluded to be sufficient. Observer agreement was assessed using ICC for a single measure and absolute agreement (Shrout & Fleiss 1979), and all items had high and statistically significant ICC values. The results suggest that educational psychology students who have been trained to use the SVRS can produce equivalent observations from the same targets. The SVRS items were also tested for scale consistency using Cronbach’s alpha, which indicated high internal consistency. Confirmatory factor analysis (Jöreskog 1969) was
performed on the data from 85 counselor performances, and 2 factors were extracted: client-centeredness and consistency. The construct validity for the modified 2-factor SVRS model was supported by the analysis, but due to the small sample size, the evidence obtained is better interpreted as preliminary in confirming the SVRS structure. Furthermore, the stability of the structure in different contexts and with different subjects is unclear at this point and requires analyses with more data from educational psychology students and from other populations.

In comparison with earlier developed measures, the CROS measures have many advantages. In our review of earlier measures, we found three coding systems with basic information on their validity and reliability: the Verbal Response Category System (HCVRCS) (Hill 1978), the HCVRCS-revised (Friedlander 1982) and the Motivational Interviewing Skills Code (MISC) (Miller 2000). The HCVRCS and the HCVRCS-revised had overlapping categories and undifferentiated use of response and response focus. One of the key features of the CRCS improves on this because it allows for freely combining response and its focus. Furthermore, the time-consuming structure of the MISC is not an issue with the CRCS, which has a simple structure with 5 response and 5 response focus categories in its variables. In the review, three rating scales that focused on client-centered aspects were found: Carkhuff’s empathy scale (Carkhuff 1969), the Accurate Empathy Scale (Truax & Carkhuff 1967) and the Skilled Counseling Scale (Crews et al. 2005). It was observed that they all had some variables with problematic operational definitions whose identification from observable behaviors could be difficult. To improve on this, all SVRS items were constructed for assessing directly observable behaviors. At this point, the success of the CRCS and the SVRS can be best assessed by the results obtained from the validity and reliability analyses, which indicated that master’s-level educational psychology students who have been trained to use the system can produce the same observations of the same events with the CRCS and the SVRS. Furthermore, these observations are connected to concepts that were assessed by the three counseling experts and that have a foundation in the results of counseling outcome research. This evidence indicates that the CROS is a potential measurement system for counseling research and that it is important to continue its development in the future.

The development of the CROS has produced a theoretical model of counselor behaviors and skilled verbal responding based on empirical findings from the development of the CRCS and the SVRS as well as the critical factors that influence client outcome, which were obtained from a vast body of counseling and psychotherapy research (Martin et al. 2000, Horvarth & Symonds 1991, Horvarth et al. 2011, Orlinsky et al. 1994, Orlinsky et al. 2004). On the CRCS, ideas adapted
from Ivey’s (1971) micro-counseling approach were refined into a new construct of counseling responses. This construct defines responses in terms of the relationship between counselor responses and the client’s message. A distinct response is identified depending on the nature of this relationship. The CRCS construct separates a response from its focus, and it treats responses and their foci as freely combining units. In this context, individual counselor responses are not observed as a measure of skillfulness but rather as different ways of relating to the client’s message. The SVRS was constructed based on Okun’s (2002) and Truax and Carkhuff’s (1967) ideas of skilled client-centered responding, which are well supported, in particular by the results from the process-outcome research by Orlinsky et al. (1994) and Orlinsky et al. (2004). The extracted factors client-centeredness and consistency are connected to those aspects of communication that help to activate the client, build the therapeutic bond and focus on the worries that the client finds most problematic.

At this point, the CRCS and SVRS measures included in the CROS enable the measurement of counselor responses and their client-centeredness and have many uses in counseling research and counselor training. The CRCS enables exploring the counselor’s use of different responses and identifying client-centered responses. This exploration can be performed from a full session or from specific time-limited segments. The CRCS provides a concrete profile of the counselors’ actions within a counseling session and can be used in response-level analysis in which different response types and their foci can be identified. The SVRS measures the client-centered aspects of counselor behaviors with 2 subscales and 6 items to provide information on the counselor’s ability to listen, specify and explore the client’s problematic experiences. At this point, the SVRS can be best used to assess the fulfillment of the client-centered aspects of counselor actions during counseling sessions. CROS measures can be used in counselor training as a feedback tool based on observed behaviors and can help observers recognize development areas and observe changes in trainee performance.

5.2 HRV and emotions during verbal reporting on unpleasant experiences

In the integrated analysis of the HRV data from studies II and III, the differences in the subjects’ emotions between viewing and reporting on unpleasant pictures and between reporting on unpleasant and pleasant experiences were examined. This setting was analog with client-centered counseling aims in that subjects produced verbal reports of their unpleasant experiences. Subjects’ emotional responses were
measured with HRV. The results from the HRV analysis showed that in the viewing phase, sympathetic activity was higher and parasympathetic activity was lower with unpleasant than with pleasant pictures. These results were consistent with the findings of earlier studies that showed that increased sympathetic activity is connected to unpleasant emotions and parasympathetic activity to pleasant emotions (e.g. Thayer & Lane 2009, Elliot et al. 2011, Bosch et al. 2009, Thayer et al. 2012). In examination of the reporting phase, the results showed that sympathetic activity was higher and parasympathetic activity lower as subjects viewed unpleasant pictures than it was as they reported on them and that sympathetic activity was lower and parasympathetic activity higher when subjects were reporting on unpleasant pictures than it was when they were reporting on pleasant pictures. These results are consistent with recent findings from studies that measured brain activity or ANS response in similar settings and that indicated that the verbal reporting of unpleasant experiences may have a regulatory effect on emotions (e.g., Campbell-Sills et al. 2006, Foland et al. 2008, Hariri et al. 2000, Hariri et al. 2003, Lieberman et al. 2007, Tabibnia et al. 2008). This dynamic was not observed during reports of pleasant experiences. However, there are other factors that might explain the changes from the viewing phase to the reporting phase. Although only a few seconds passed from viewing the stimuli to reporting on them, the time effect might explain some of the changes in HRV between viewing and reporting. This process may have also been influenced by the emotional intensity of the initial experience; high-intensity emotions may last longer than low-intensity emotions. However, comparison within phases shows that viewing and reporting on unpleasant stimuli clearly differ from viewing and reporting on pleasant stimuli. Viewing unpleasant stimuli showed higher sympathetic activity and lower parasympathetic activity than did viewing pleasant stimuli. Furthermore, the results from reporting on unpleasant stimuli showed lower sympathetic activity and higher parasympathetic activity than that for reporting on pleasant stimuli. This evidence supports the conclusion that verbally reporting on unpleasant experiences may have regulatory effects on subjects’ emotions that can be observed as an increased parasympathetic effect and a decreased sympathetic effect.

Additionally, the subjects’ respiratory frequencies were estimated with spectral peak location to control possible biases in the spectral components of HRV (Thayer et al. 2002). The respiratory frequency estimates were within acceptable limits for nearly all analyzed sequences, but 3 subjects were filtered out from the analysis for biased spectral powers. Overall, oral reporting did not produce spectral bias in most of the sequences. However, the 3 biased sequences in the estimation confirm the importance of respiratory frequency estimates in obtaining accurate HRV results. In
spite of the estimation of respiratory frequency, other respiratory influences could not be controlled because controlling them would have required the use of a mask. This was not possible because the task involved speaking and also because of the specific research aims.

The results from the integrated analysis of the HRV data from studies II and III were consistent with earlier findings indicating that sympathetic activity was higher and parasympathetic activity lower as subjects viewed unpleasant pictures than it was as they viewed pleasant pictures and that sympathetic activity was lower and parasympathetic higher as subjects were reporting on unpleasant pictures than it was as they were reporting on pleasant pictures. Respiratory frequency estimates showed that most of the sequences analyzed were within acceptable limits, and only a small number of biased sequences were found. From the motivational point of view (Lang & Bradley 2010), the effect of reporting unpleasant experiences can perhaps best be explained with approach behavior that is facilitated by the reporting task. In the viewing phase, stimuli are experienced as unpleasant with concordant subjective evaluation and autonomic activity. In the reporting phase, the subjects’ focus is targeted directly at making observations and reporting their thoughts and observations fulfilling the primary goal of client-centered counseling. This process is analogous to Barlow's et al. (2004) treatment principles, which aim to reinforce approach behavior in avoidant subjects. The client-centered approach produces a special way of reporting unpleasant experiences whereby the focus is on observing one’s problematic experiences and the associated events. These two factors (exposure and non-avoidant task focus) relate to Barlow's et al. (2004) ideas on treating avoidance-based distress and disorders, which emphasize preventing avoidance behavior and initiating constructive tasks. One of the fundamental ideas behind Barlow's et al. (2004) principles is enhancing adaptive emotion regulation by shifting attention and awareness toward the distressing or avoided parts of a client’s experience. In client-centered counseling, these are both realized using a method that aims to facilitate the client in approaching difficult aspects of his experience by verbally reporting on them (Okun 2002).

### 5.3 General conclusions

The general purpose of this study was to develop methodology to assess counseling interactions during sessions by developing measures for assessing two key components in successful counseling interactions: client-centered counselor responses and client emotions. The CROS was developed to assess counselor responses and
their client-centeredness. HRV experiments were conducted to obtain information on HRV’s applicability to measuring emotions during verbal reports of unpleasant experiences. The general idea behind the developments was to construct instruments and methods for assessing the relationship between counselor actions and client emotions. Overall, the CROS measures can be considered versatile tools in observing the types of responses counselors use and assessing the client-centeredness of counselor actions. The results from the HRV experiments were consistent with earlier findings on the effect of verbally reporting on unpleasant experiences on ANS responses that indicated decreased sympathetic activity during the reporting. This dynamic was not observed with pleasant pictures. The results suggest that verbal reporting may have regulatory effects on emotions and that emotional reaction during reporting may depend on the initial emotional reactions during viewing and thus that verbal reporting should be interpreted accordingly. Together, the CROS and HRV are potential measures for examining the relationship between counselor actions and client emotions. At this point, the current state of CROS development can be considered a starting point that enables its further development. Furthermore, more HRV experiments are required before it will be possible to accurately interpret client emotions during sessions because of the possible regulatory effects.

Earlier studies have shown that many client problems are connected to avoidance behavior and maladaptive emotion regulation, which manifests as avoidant behavior and associated emotions (Barlow 2004). When counselor actions focus on the client’s message and lead the client to give accurate descriptions of his experience, the client faces the troubling, initially avoided parts of his experience and adaptive regulatory strategies for emotions are supported (Okun 2002; Truax & Carkhuf 1967, Barlow et al. 2004). One important aspect in examining this emotional dynamic is that client emotions can be measured and interpreted reliably in relation to counselor actions. However, recent findings have indicated that verbally reporting on unpleasant experiences, which is crucial in conversational client-centered counseling, may have instant regulatory effects on emotions. These findings may have important consequences from the counseling point of view, implying that HRV during reporting should be interpreted in relation to a client's emotional reactions during a counselor’s response. For example, it can be hypothesized that a question initiated by the counselor that is focused on an unpleasant aspect of the client’s experience will increase sympathetic influence but that when the client starts to report observations and thoughts associated with this experience, parasympathetic activity will increase and sympathetic activity will decrease. Reporting shifts the focus to the content that was initially avoided but does so in a non-preventive way by observing it. This focus
shift may have implications for the way that emotions should be interpreted during conversation and may suggest that stimulus and reporting should be interpreted as one unit. In this approach, the client’s psychophysiological reactions during reporting depend on the focus of the counselor’s response and the focus of the client’s verbal response. When the counselor focuses on troubling, unpleasant aspects of a client’s experience and guides the client to describe related content, this should be reflected in an increase in the client’s sympathetic activity. When the client responds accordingly to the counselor’s initiative and describes his unpleasant experience, the sympathetic activity should decrease. The interpretational frame is highly dependent on recognizing and measuring the client’s emotions during the counselor’s response and the client’s emotions during reporting, which cannot be interpreted without this initial sequence. Although only the first steps have been taken in the development of the CROS and the assessment of using HRV to measure emotions during verbal reports, as the development work continues, these measures can be perhaps used together in the integrated exploration of client-centered counselor actions and client emotions. Eventually, this will enable the extraction of those counselor actions that are associated with positive changes in client emotions and successful outcomes.

5.4 Limitations and future developments

There are some aspects of the CROS development that require further testing. Use of the CRCS and the SVRS should be extended to additional populations and alternative settings. For instance, the coverage of the CRCS categories and the SVRS items and their usability in different contexts would provide information about the CROS’s applicability and limitations in different settings. Because of the small sample size, the SVRS 2-factor model should be tested with a larger sample to obtain more evidence of its accuracy. Construct validity should also be strengthened with convergent and discriminant validity analyses that compare related and unrelated measures. Furthermore, inter-rater reliability should be further tested with an even greater number of targets in different settings and with different trained observers. This would provide information of the stability of the measurement in different contexts and raters. Furthermore, it will be important to apply CROS measures to more natural counseling settings and to process-outcome measurement in studies in which the entire counseling process, including multiple sessions, is measured and analyzed. The SVRS should also be tested for criterion validity with a validated client outcome measure. This would validate the SVRS as an indicator of an effective counselor and would allow for easily executed interactional measurements that would
provide information on a counselor’s ability to help a client achieve positive outcomes.

There are aspects of the HRV measurement during verbal reporting that should be addressed. Respiration may have produced HRV effects that could not be controlled because a mask was not used in this study, and this may have affected the extracted spectral components. Furthermore, extraction of the parasympathetic HF component is connected to a typical breathing frequency of 1 breathing cycle per 2- to 7-second interval (0.14–0.5 Hz), and frequent deviation from this breathing frequency will affect the results (Cerutti et al. 1995). This is why it is critical to control respiratory frequency. Although respiratory frequency was not directly measured, a respiratory estimation method was used that has been validated with plethysmographic recordings; thus, the estimates can be considered sufficient when considering respiratory frequency influences (Thayer et al. 2002). Furthermore, there are other factors that might explain the changes from the viewing phase to the reporting phase. Although only a few seconds passed from viewing the stimuli to reporting on them, the time effect might explain some of the changes in HRV between viewing and reporting. This process may have also been influenced by the emotional intensity of the initial experience; high-intensity emotions may last longer than low-intensity emotions. This suggests that most reliable comparison were done within the reporting phase.

In the future, measurement of HRV and emotions during verbal reporting should be expanded to more natural settings and eventually to counseling sessions in which they can be compared with counselor actions obtained with the CROS. Furthermore, additional autonomic measures should be applied to the setting to obtain more reliable assessments of client emotions. However, this requires more experiments in addition to having a detailed understanding of emotional dynamics in a setting that examines client emotions at certain response levels and at different phases of a conversation. This is crucial because the interpretation of emotions during reporting may be associated with the primary emotional valence of the experience. As the development continues, together, CROS and HRV measurement can help to create a framework in which counselor actions and client emotions can be accurately measured and interpreted.
References


Appendix A

Table 4. The Counselor Response Coding System.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Operational definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Counselor response</td>
<td></td>
</tr>
<tr>
<td>Reflection</td>
<td>The counselor repeats client words or phrases or summarizes the essence of what the client has said.</td>
</tr>
<tr>
<td>Specifying questions</td>
<td>The counselor asks a question from the client that intends to specify what the client has said.</td>
</tr>
<tr>
<td>Conclusion</td>
<td>The counselor presents a hypothesis that is constructed by connecting elements of the client’s message; the counselor completes the client’s message by adding an element to it.</td>
</tr>
<tr>
<td>Suggestion</td>
<td>The counselor asks a question that leads the client to answer according to the counselor’s assumptions; the counselor tells the client how to act, think or feel.</td>
</tr>
<tr>
<td>Self-disclosure</td>
<td>The counselor describes his own experience or opinion to the client.</td>
</tr>
<tr>
<td>Counselor response focus</td>
<td></td>
</tr>
<tr>
<td>Feelings</td>
<td>The counselor focuses on the client’s emotional words, phrases or explores with questions what kind of feelings the client is experiencing or has experienced.</td>
</tr>
<tr>
<td>Explanations</td>
<td>The counselor focuses on the way how the client perceives reasons and motives behind his thoughts, feelings and actions.</td>
</tr>
<tr>
<td>Action</td>
<td>The counselor focuses on the client’s actions, aims, plans, or strategies.</td>
</tr>
<tr>
<td>Context</td>
<td>The counselor focuses on the way how the client generally perceives his thoughts, persons, facts or places related to the client’s message.</td>
</tr>
<tr>
<td>Strengths</td>
<td>The counselor focuses on the client’s positive traits, successes, skills.</td>
</tr>
</tbody>
</table>

Table 5. Skilled Verbal Responding Scale.

<table>
<thead>
<tr>
<th>Factor</th>
<th>SVRS items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Client-centeredness</td>
<td>1. The counselor speaks of things that the client has not mentioned(^a)</td>
</tr>
<tr>
<td>Client-centeredness</td>
<td>2. The counselor uses open questions ()</td>
</tr>
<tr>
<td>Client-centeredness</td>
<td>4. The counselor does not share opinions or give advice (^b)</td>
</tr>
<tr>
<td>Consistency</td>
<td>3. The counselor focuses on exploring the client’s problems (^a)</td>
</tr>
<tr>
<td>Consistency</td>
<td>6. The counselor uses specifying questions (^a)</td>
</tr>
<tr>
<td>Consistency</td>
<td>7. The counselor ignores the client’s message (^b)</td>
</tr>
</tbody>
</table>

Items are scored as estimated proportion of each behavior during the whole session. Scoring criteria: estimated frequency proportion of the item \(x\) \((0–3)\).

\(^a\) Positive scale: \(0 = [x = 0] \), \(1 = [0 < x \leq \frac{1}{4}] \), \(2 = [\frac{1}{4} < x < \frac{1}{2}] \), \(3 = [\frac{1}{2} \leq x]\)

\(^b\) Negative scale: \(3 = [x = 0] \), \(2 = [0 < x \leq \frac{1}{3}] \), \(1 = [\frac{1}{3} < x < \frac{1}{2}] \), \(0 = [\frac{1}{2} \leq x]\)
## Appendix B.

### Table 6. Mean LF, HF and LF/HF values during viewing and reporting of pleasant and unpleasant stimuli.

<table>
<thead>
<tr>
<th>Variables</th>
<th>LF (%)</th>
<th>HF (%)</th>
<th>LF/HF</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Viewing pleasant</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>54.77</td>
<td>39.12</td>
<td>1.37</td>
</tr>
<tr>
<td>Standard error</td>
<td>2.97</td>
<td>2.83</td>
<td>.13</td>
</tr>
<tr>
<td>95% CI</td>
<td>48.71, 60.83</td>
<td>33.35, 44.89</td>
<td>1.10, 1.63</td>
</tr>
<tr>
<td><strong>Viewing unpleasant</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>57.62</td>
<td>36.48</td>
<td>2.04</td>
</tr>
<tr>
<td>Standard error</td>
<td>3.02</td>
<td>2.98</td>
<td>.27</td>
</tr>
<tr>
<td>95% CI</td>
<td>51.45, 63.79</td>
<td>30.39, 42.58</td>
<td>1.49, 2.58</td>
</tr>
<tr>
<td><strong>Reporting pleasant</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>55.52</td>
<td>37.59</td>
<td>1.56</td>
</tr>
<tr>
<td>Standard error</td>
<td>2.42</td>
<td>2.47</td>
<td>.15</td>
</tr>
<tr>
<td>95% CI</td>
<td>50.57, 60.46</td>
<td>32.55, 42.63</td>
<td>1.25, 1.86</td>
</tr>
<tr>
<td><strong>Reporting unpleasant</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>50.93</td>
<td>42.51</td>
<td>1.36</td>
</tr>
<tr>
<td>Standard error</td>
<td>2.59</td>
<td>2.49</td>
<td>.17</td>
</tr>
<tr>
<td>95% CI</td>
<td>45.65, 56.22</td>
<td>37.43, 47.59</td>
<td>1.02, 1.70</td>
</tr>
</tbody>
</table>
List of original publications


Reprinted with permission from Sage (I) and Elsevier (II and III).

Original publications are not included in the electronic version of the dissertation.
134. Erkkilä, Tuomas (2013) Pedagogiikka Tapiolan kuorossa ja Kari Ala-Pöllänen yhteistoiminnallisena lapsikuoronjohtajana
143. Virkkula, Esa (2014) “Soittaminen ammatillaisen kanssa on paras tapa oppia” : työpajapäivänä työssä oppiminen muusikoiden ammatillisessa koulutuksessa
144. Lutovac, Sonja (2014) From memories of the past to anticipations of the future : pre-service elementary teachers’ mathematical identity work
147. Suorsa, Teemu (2014) Todellisuista on mahdollinen : systeeminen sekä subjektietiivistelmä ruokavarsituspsykologiseen kokemuksen tutkimukseen

Book orders:
Granum: Virtual book store
http://granum.uta.fi/granum/
Antti Rantanen

DEVELOPMENT OF METHODOLOGY FOR ASSESSING COUNSELING INTERACTIONS

DEVELOPING THE COUNSELOR RESPONSE OBSERVATION SYSTEM AND ASSESSING APPLICABILITY OF HEART RATE VARIABILITY TO THE MEASUREMENT OF CLIENT EMOTIONS DURING VERBAL REPORTING