

Comprehensive Summaries of Uppsala Dissertations
from the Faculty of Social Sciences 140



Nondirective Counseling

*Effects of Short Training and Individual
Characteristics of Clients*

BY

ERIK RAUTALINKO



ACTA UNIVERSITATIS UPSALIENSIS
UPPSALA 2004

Dissertation presented at Uppsala University to be publicly examined in Sal X, Universitetshuset, Uppsala, Friday, November 12, 2004 at 13:15 for the degree of Doctor of Philosophy. The examination will be conducted in Swedish.

Abstract

Rautalinko, E. 2004. Nondirective Counseling. Effects of Short Training and Individual Characteristics of Clients. Acta Universitatis Upsaliensis. *Comprehensive Summaries of Uppsala Dissertations from the Faculty of Social Sciences* 140. 68 pp. Uppsala. ISBN 91-554-6036-4

Nondirective counseling is to listen, support, and advise, without directing a client's course of action. It has been influenced by humanistic theories in the tradition of Carl Rogers, but techniques used in nondirective counseling are common in many forms of psychological counseling and treatment today. There are, however, few conclusions as to what the results of training nondirective counseling are. The purpose of the present thesis is to examine effects of nondirective counseling training, and to analyze how such effects are moderated by the characteristics of clients. Three quasi-experimental or experimental studies (Paper I-III) are presented. In Paper I, trained and untrained insurance company employees were compared on their Reflective listening (RL; a subskill of nondirective counseling) skills before and after a training program. Training increased RL, and the skills were transferred to authentic settings. Trained employees were, however, not evaluated differently than untrained. In Paper II, psychology students were compared before and after RL training of three time lengths. All training times increased skills equally, but clients disclosed more information to those with longer training, the students remembered the information better, and external judges perceived the therapeutic relationship as better, especially if the judge was socially competent. In Paper III, two nondirective counseling techniques, RL and open-ended questions, were evaluated by judges who differed in social skills and cognitive ability. RL received positive ratings, whereas open-ended questions did not, and the judges' ratings were moderated by their social skills and cognitive ability. In the Discussion, it is proposed that even short training has effects, that trained skills generalize to authentic contexts, but that the usefulness of the examined subskills of nondirective counseling depends on client characteristics such as social skills and cognitive ability.

Keywords: nondirective, counseling, reflective listening, open-ended question, training, social skills, cognitive ability

Erik Rautalinko, Department of Psychology, Box 1225, Uppsala University, SE-75142 Uppsala, Sweden

© Erik Rautalinko 2004

ISSN 0282-7492

ISBN 91-554-6036-4

urn:nbn:se:uu:diva-4551 (<http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-4551>)

LIST OF PAPERS

Paper I: Rautalinko, E., & Lisper, H.-O. (2004). Effects of training reflective listening in a corporate setting. *Journal of Business and Psychology, 18*, 281-299.

Paper II: Rautalinko, E., Lisper, H.-O., & Ekehammar, B. (2004). Training reflective listening in personal counseling: Effects of time and client social skills. *Manuscript submitted for publication.*

Paper III: Rautalinko, E. (2004). Nondirective counseling: For the socially skilled and intelligent? *Manuscript submitted for publication.*

Paper I is reprinted with permission from Kluwer Academic.

Contents

INTRODUCTION	9
Verbal behavior and language.....	9
Speech style and accommodation	10
Language research.....	11
Nativist, behavioral and interactionist theories.....	11
From the descriptive to the applied.....	12
Importance of theory for changing communication	14
Counseling psychology	15
History and theoretical contributions.....	16
Counselors and techniques	17
Nondirective counseling.....	19
Carl Rogers and client-centered therapy.....	19
Thomas Gordon and reflective listening.....	22
Research on reflective listening.....	24
EMPIRICAL STUDIES.....	28
Research questions and method summary.....	28
Participants	29
Communication skills training.....	29
Taping, transcription and response categorization	29
Dependent measures	30
Moderating factors.....	32
Design and statistical analyses.....	33
Paper I	34
Study 1	35
Study 2 and Study 3	36
Paper II.....	37
Study 1	38
Study 2.....	39
Study 3.....	40
Paper III.....	41
Study 1	42
Study 2.....	43

Effects of sex — a supplementary analysis.....	45
Paper I.....	45
Paper II	46
Summary.....	46
DISCUSSION.....	47
Main findings and implications.....	47
More training is needed	48
Communication skills vary.....	48
Methodological issues	50
Social validity.....	50
Defending the randomized experiment.....	53
Content analysis and quantification.....	54
Rating procedure.....	55
Applications	56
Psychological interventions.....	57
Call-center activity.....	59
Future research	60
ACKNOWLEDGEMENTS.....	62
REFERENCES	63

Abbreviations

CST	Communication skills training
MI	Motivational interviewing
RL	Reflective listening
SPM	Raven's Standard Progressive Matrices
SSI	Social Skills Inventory
WAI	Working Alliance Inventory

INTRODUCTION

The word *communication* sums up the process of conveying information from one person to another. It is an important aspect of social interaction, and much of everyday behavior, such as hugging someone, saying “Hello”, or ordering an airplane ticket on the Internet, involves communication. These examples readily show, too, that communication varies in complexity, can be automatic or produced, intentional or unintentional, and that the information sent from one person to another takes on many forms: written, oral, or nonverbal.

Communication requires a sender, who sends information to a receiver, through their shared channel of communication. Needless to say, the process of sending and receiving information can take part between two or several people, and more than one message can be sent at the same time. This variety can make communication a complex event, and the study of it is a task that can be undertaken from within many scholarly disciplines, such as psychology, sociology, or linguistics, to name a few.

One aspect relevant to psychological study of communication is how professional helpers (e.g., counselors) communicate with help-seekers. The purpose of this thesis is to investigate how the communication skills of professional helpers can be enhanced. In focus is a particular communication skill that is sometimes used in counseling and other psychological interventions: so called nondirective counseling.

Verbal behavior and language

Language is a subset of verbal behavior, which is a subset of communication. Language is governed by morphological, syntactic, and semantic rules that are shared by people belonging to a certain social group. This is how seemingly meaningless combinations of sounds result in meaningful sequences of words (in the present text called verbal responses), that make language a fast medium of communication. A mastery of this medium, however, does not only require knowledge of words. It also requires knowledge of the cultural rules guiding what is appropriate to say in

a specific context. For example, Searle (1979) identified five sorts of meaning that can be communicated through language: saying how something is, getting someone to do something, expressing feelings and attitudes, making commitments, and accomplishing something directly. These meanings may be present individually or simultaneously, rule out one another etc., adding to the complexity of the interaction. Consequently, this complexity transfers to the investigation of language.

There has been debate about the origin of human capacity for learning language. Chomsky (1959) argued that the rules of grammar are basically innate, learning being activated in the process of normal development. Others (e.g., Lock, 1980) have proposed that language learning stems from prelinguistic interaction between a child and its caregivers. This discrepancy exemplifies the nature-nurture debate in psychology: is language ‘within’ the individual at birth? Or is it learnt in interaction with caregivers and the environment? Such questions can be tantalizing if one wants to fulfil some kind of *truth* criterion, but are nevertheless interesting both in theory and in practice. Later on in the text, I shall discuss the implications of nature or nurture theories to someone who undertakes the challenge of changing verbal behavior.

Speech style and accommodation

Nonlinguistic accompaniments of speech, such as volume, stress, pitch, speed, and voice tone, are sometimes referred to as paralanguage or nonverbal skills (Ralph & Thorne, 1993). Timing, pitch, and loudness, the so-called prosodic features of language (Argyle, 1975), are important aspects of verbal behavior too, because they can change the meaning of a verbal response. They are important cues to the speaker’s emotions: a low pitch can communicate sadness or boredom, while a high pitch can communicate anger, fear, or surprise (Frick, 1985). This line of research exemplifies that the social psychology of language has focused on *how* something is said, as well as on *what* is said. Furthermore, results of nonlinguistic research accentuate the influence of variables other than those readily seen in a transcript of a conversation.

Most people master a variety of ways to say something, or speech styles, and tailor these with consideration to context. We tend to speak slowly and use simple words to children and others who do not understand complicated speech (Elliot, 1981). Hence, speech style is influenced by the subjective perception of the context, such as purpose, topic, participants and their relationships, their moods, and appearance (Brown & Fraser, 1979). This means that after having formed a subjective impression of a context, we adjust our speech style, or even seek out a context that matches the

impression that we want to create. And vice versa, speech style itself may tell us things about the person talking: her or his social class, ethnicity, sex, and age. The vital point is, however, that verbal behavior can never be judged at face value because it is so dependent on its context. This perspective is called the contextual approach.

People accommodate their speech according to motives of being understood and obtaining social approval. If the speakers' styles shift toward each other, then speech is said to converge, and when the opposite happens, it is said to diverge. Convergence increases liking (Bourhis, Giles, & Lambert, 1975), in particular when it is seen as intentional (Simard, Taylor, & Giles, 1976). This suggests that generally the parties in a conversation would want to converge their responses. The process of convergence and divergence can, however, be affected by other motives as well, such as someone wanting to clarify that she or he belongs to a certain social group.

Another example of speech accommodation is our disclosing approximately equal amounts of personal matters to one another when entering a relationship (Cozby, 1973). This does not hold for most professional conversations, but there is speech accommodation in such interactions too. For example, a therapist's verbal behavior may accommodate to that of a client, such as when a therapist speaks softly and slowly to a client who is overtly depressed. Later on, it is suggested that when a client is not aware of the therapist's accommodation, both intentional and unintentional speech feel essentially the same.

Language research

Language research in the linguistic sense goes far back, perhaps to the time when written language was invented, consisting of literary analyses of plays, novels, and speeches. Scientific research on verbal behavior, however, is a product of the 20th century. Consequently, psychological investigations of verbal behavior have followed the main theoretical and methodological perspectives of the last century.

Nativist, behavioral and interactionist theories

There are different theoretical perspectives on verbal behavior. A major difference is between the nativist and behaviorist approaches, whose conflicting ideas led to a spirited debate between Noam Chomsky and B. F. Skinner in the 1950s (Chomsky, 1957; Skinner, 1957).

Nativist theories, such as those proposed by Chomsky, advocate that language acquisition is a uniquely human process with origins in the structure of the brain. There is support for the nativist perspective from, for example, languages that are invented by deaf children or children with different cultural backgrounds who initially lack common language (Gentner & Goldin-Meadow, 2003). These invented languages may be structured in a way suggesting that humans have some kind of built-in language acquisition device. Other support for the nativist perspective comes from research in the spirit of Jean Piaget showing that language development in children goes through specific stages that are essentially the same over many different cultures (e.g., Piaget, 1926).

Behavior theory, on the contrary, claims that language is acquired through modeling and operant conditioning. Skinner and others have focused on showing how parents selectively reinforce language in their children. Caregivers provide examples of verbal behavior to children, who imitate verbal operations that are either ignored or reinforced. This perspective distinguishes itself from nativist theories by a continuity strategy, proposing that the behavior found in humans and less developed species is determined by the same rules. Support for the behaviorist perspective comes, for example, from language training of people with developmental disabilities (e.g., Griffiths, Feldman, & Tough, 1997).

Although behaviorism dominated psychology between 1920 and 1970 (Metcalf, 2000; Nelson, 1996), not many theorists on communication retain a behaviorist perspective today. Instead, research is focused on nativist/cognitive issues, and there has also been a consolidation between theoretical approaches into an interactionist theory of language. It states that there is an interaction between innate predispositions in the child and the social environment. Genetic factors set the stage for language learning, but the social context is important as well. This perspective emphasizes that we must know who the interactants in a conversation are before we can assume the meaning of an utterance. Interrupting someone, for example, can be considered rude, or signal dominance, but it can also signal involvement and support (Ng, Bell, & Brooke, 1993).

From the descriptive to the applied

One way to categorize research has already been mentioned, namely according to whether findings are theoretically more related to nativist, behavioral, or integrative approaches. Another way to categorize research is according to the applicability of the results. Such a categorization could result in the three domains descriptive, predictive, and applied research.

The investigations that fall into the descriptive category consist of case studies or anecdotes of verbal behavior. One example is a reanalysis of a 30-min videotaped session between Carl Rogers and a female client, focusing on nonexpert language and meta-statements (Wickman & Campbell, 2003). Their main research question was what different kinds of language there are. Descriptive investigations focus on what verbal behavior exists in specific groups, such as teachers, humanistic counselors, or managers, and has advantages as well as disadvantages. An advantage is that the behavior under study is often complex, interesting, and ecologically valid. The approach is if possible non-intrusive, and may be the only way to study unique verbal behavior that cannot be subject to customary sampling. A disadvantage of descriptive research is that it may generate knowledge that cannot be applied, thus lacking one aspect of external validity.

Investigations in the predictive category often comprise studies of process and outcome in language, for example effects of specific verbal responses on perceived expertness, attractiveness, and trustworthiness (Watkins, Savickas, Brizzi, & Manus, 1990), or the effects of the sex of the speaker on competence (Steckler & Rosenthal, 1985). They answer questions such as: what does this language lead to? In order to demonstrate causality, quantified data and experimental manipulation are common, either in a controlled setting, or in naturalistic experiments where determinants are analyzed using regression techniques. There is often an external criterion, and though the research seldom prescribes certain behavior, the results may implicitly do so. An advantage of predictive research is that the findings often have explanatory power. An external criterion such as shyness is causally explained from other factors, in ways that often seem logically convincing. A disadvantage is that as measures are made more reliable, they can become oversimplified and devoid of meaning. Sometimes they lack causal explanatory power as well, if relations are supported only statistically in cross-sectional data. The investigations in the present Paper III are examples of predictive research.

Investigations in the applied category consist of reports on education or communication skills training. Such investigations answer questions such as: how do we obtain a language that does this? Here it is common to use quantified data and experimental manipulation, although the circumstances for such research are sometimes constrained by practical issues. There is often an external criterion, and the goal of the research is to prescribe a process or behavior that has some positive consequence. Like the other two methodologies, applied research has both advantages and disadvantages. One advantage is that the presumptive results are usually high on one aspect of external validity, that is, usefulness. They may match the goal of applied research which is to benefit the community. Common disadvantages are intrusive procedures, and that the external criterion against which usefulness

is evaluated is implausible. The investigations in the present Paper I and II fall into the applied category.

Importance of theory for changing communication

What makes the discussion of theory and methodology relevant? The first obvious reason is that it is often advocated that psychological research should be theory-bound or at least theorized (e.g., Kiesler, 1973, cited in Marmar, 1990). Findings that are incorporated in a theory may have more explanatory power than detached empirical evidence or 'loose' findings. Moreover, an investigator who fails to make theoretical assumptions about her or his findings may have difficulty publishing the results.

Another argument is that a debate between theoretical paradigms can boost research. Debates catch attention, which inspires to research programs, which leads to more funding. Attention drawn to a debate between two theoretical paradigms can also have the intentional or unintentional side effect of attention being drawn away from a third paradigm. An example of this process is how the debate between behavioral, cognitive, and humanistic approaches during the 1950s and 60s drew the attention of scientific research on professional communication from psychoanalytic theory.

My dividing point in how theory and methodology influence research, however, is that they affect how one regards the deliberate change of communication skills. Purely nativist theories can never give implications for changing verbal behavior of an individual. If behavior is innate, then it cannot be manipulated after birth. Purely descriptive theories undoubtedly reason in the same way. If behavior is seen as a unique entity, then it cannot be generalized outside its specific context, and the prospects of changing small elements of behavior in order to attain a certain goal are small. On the contrary, the interactionist perspective in the domain of applied research supplies the pre-requisites for enhancing people's communication skills.

According to an interactionist perspective on language, however, it is impossible to define a communication style that is always suitable (i.e., 'good communication'; Hayes, Follette, & Follette, 1995, pp. 159-160). Instead, the dispositions and behavior of the individual and the environment interact in a manner that may be unpredictable. If one sets out from a defined relationship and context, however, it can be possible to state what characterizes good communication in that specific instance. Definitions might focus on outcomes of an interaction, including goals that one wants to attain in the relationship.

This thesis is based on the axiom that communication skills are learned behaviors that affect an individual's relations with others. They are situation-specific and affected by individual characteristics such as age, sex, and

social status of the interactants. They are both verbal and nonverbal, and can be specified and targeted for intervention (Elliott, Sheridan, & Gresham, 1989). There is no standard method for assessing communication skills. Rather, the hypotheses under investigation determine how an assessment of behaviors is done, from the microscopic level to the global. Moreover, it is important to remember that estimating people's communication styles amounts to making value statements, and generally includes some analysis of the people who are interacting and the purpose of their communication.

Systematic procedures to change communication skills are called communication skills training (CST). Although the training is conducted in a specific context, it is assumed to generalize across time, settings, individuals, and behaviors. Research on CST, however, has not always produced consistent results. For example, assessment data obtained from different sources often correlate modestly if at all (Elliott et al., 1989). Therefore, multiple sources of information in, and replications of, a training program are required to increase the external validity of findings. This has sometimes been called the social validation approach (van Houten, 1979). Such a reductionist view has raised critique towards CST. Critics have claimed that, although training sometimes has been found to have positive effects, it is unethical, flawed in delivery, and manipulative in its impact (Hargie & Tourish, 1994). CST supposedly can be a covert method to change the participants' opinions of what good communication is, and there can be no guarantees that training is for the good of those with whom the trained people interact.

Counseling psychology

Counseling is a superordinate term for many helping activities. The term itself means to listen, support, advise, or recommend professionally on social problems (Guralnik, 1984; Longman dictionary of contemporary English, 1995; Sykes, 1982). It emphasizes "facilitating the normal development and growth of the individual in coping with important problems of everyday living, as initially contrasted with clinical psychology" (retrieved from <http://www.dictionarybarn.com>, on March 17, 2004). Thus, counseling psychology focuses on everyday life, while clinical psychology essentially deals with psychopathology. This difference may, however, be gradually fading out (Tyler, 1992).

American Psychological Association considers for its Journal of Counseling Psychology articles on assessment, interventions, consultation, supervision, training, prevention, and psychological education, as well as career development and other professional issues. Evidently, the term is

broad, and the research area is perhaps best defined by the inclusion criteria for publications in the field's main peer-reviewed journals, such as the *Journal of Consulting Psychology* (founded in 1937), *Journal of Counseling and Development* (1923), *Journal of Counseling Psychology* (1954), and *Counseling Psychologist* (1973).

A procedure that seems closely related to counseling is consultation. In a study about consultation in schools, Lin, Kelly, and Nelson (1996) defined it as the interaction between a counselor and significant others (e.g., spouse, parent, or teacher) with the purpose of assisting someone to function more effectively. Their results did not amplify the practical differences between counseling and consultation, instead it seemed that the verbal behaviors used by counselors vs. consultants were practically the same. However, giving advice was more commonly used in consultation.

History and theoretical contributions

Psychological counseling is probably as old as humankind itself. Some people may still associate counseling with psychoanalysis of the old school, imagining cushioned couches in 19th century Vienna. But there are differences. First, the field of psychotherapy is itself remarkably different today from what it was a hundred years ago, although this is not discussed further in the present text. Second, there are vast differences between the accurate study of individual patients by Freud and his followers, and the empirical study of psychotherapy and counseling that took off in the 1930s and 40s.

Two factors precipitated the study of counseling psychology. One was a need emerging at the end of World War II, when the U.S. founded the Veterans' Administration (Hill & Corbett, 1993). Its purpose was to integrate veterans back into civilian life, a lengthy task creating a huge demand for services such as psychological and vocational guidance. The number of returning war veterans was so high that the authorities providing psychotherapy and counseling, mainly consisting of psychotherapists and psychiatrists, could not accomplish the task by themselves. The remedy was to educate psychologists and other professionals for counseling services, usually in short time compared with traditional analytic therapy. This education created a need for the scientific study of counseling psychology. The other factor was the means for doing the research, which had come a decade earlier as audio recording became available to larger audiences in the 1930s.

The first large-scale research program started at Ohio State University some years before the war, and was conducted by Frank Robinson (Hill & Corbett, 1993). Carl Rogers, one of the icons of humanistic psychology,

joined the clinical psychology program of that university, and contributed to the already existing research by developing theory. It has later been suggested (Gelso & Fretz, 1992) that Rogers' theory fitted especially well into the practice of counseling psychology and the Zeitgeist of the American 1950s. It put emphasis on persons' strengths and assets, instead of the conventional emphasis on their shortcomings. Consequently, help-seekers in counseling are often called clients, while in more traditional therapies they are called patients.

Counseling psychology was strongly influenced by humanistic theories. Eventually behavioral and cognitive theories began to impact counseling psychology. They fitted well into the humanistic tradition because of their emphasis on maladaptive learning instead of psychopathology. Many behavioral treatments have been developed through the years, some of them influenced by the humanistic heritage in counseling psychology, although it is not always given credit (e.g., Motivational interviewing, see further below). After the cognitive revolution of the 1960s, in particular, thoughts were regarded as intermediate behaviors linking external stimuli and the client's overt behavior. It renewed the interest in the client's thoughts, the so-called cognitive therapy approach, and also justified the scientific study of therapist attitudes and intentions.

While psychoanalytic theory had laid the foundations for psychotherapy through the work of Freud, it was not involved in counseling psychology until much later. This coincided with the advent of self-psychology theories (Kohut, 1971, cited in Hill & Corbett, 1993) that were more compatible with the humanistic-oriented counseling theories than earlier psychoanalytic theories had been. There are some differences, however. Kohut assumed that the goal of therapy was to strengthen the self, while Rogers assumed that the self was already established and the goal of therapy was to widen consciousness of the self. Still, the contribution of psychoanalytic theory to counseling psychology has been large, due to its focus on the therapeutic relationship, a nowadays pantheoretical construct that is seen to predict positive change in the client in all forms of counseling.

Counselors and techniques

There are two ways to produce counselors: selection or training (Jennings & Skovholt, 1999; Riggio & Taylor, 2000). Selection procedures, for example ranking counseling program applicants according to their grade point average or by interviews, are a proactive way to get the 'best' people. The difficulty is to find variables that reliably predict later performance. Yet, one of these could be listening skills which are typically covered in CST programs (Hargie & Tourish, 1994). The second possibility for training

assumes that people's communication skills have a baseline that can be elevated through repeated practice of relevant microskills. This is the approach suggested in the present text.

Since the time when counseling psychology was founded, the field has crudely been divided into process and outcome research. The former investigates therapist and client behaviors and their interaction, the latter refers to changes as a result of a psychotherapy (Marmar, 1990). Hill (1992) wrote a review on therapist techniques in brief individual therapy. She criticized effectiveness studies that correlate frequency of a technique with treatment outcome, because the correlation method is not sensitive to issues such as quality or timing. This is the reason why many investigators have started to study the immediate effects of counseling techniques. It is believed that therapist verbal responses influence the client directly (e.g., interpretation, which has been found to be effective across many studies), although other variables such as therapist and client individual characteristics, their relationship, the context, all interact to influence process and outcome of therapy (Marmar, 1990). The dividing point, as I remarked earlier, is that many individual characteristics of a counselor cannot be targeted for change during any short time.

The most common method to operationalize therapist techniques has been through response modes. Hill and others (1988; 1992) came up with a category system comprising nine topic-independent counselor techniques that facilitate positive change in clients. They were approval (providing support or reinforcement), information (giving facts or answering questions), direct guidance (e.g., advice), closed question (inquiring limited information), open question (without purposely limiting the response), paraphrase (restating content, reflecting feeling, or summarizing), interpretation (giving new meaning to behaviors, feelings, or thoughts), confrontation (highlighting discrepancy or contradiction), and self-disclosure (sharing something personal about oneself).

Another review of therapist characteristics and techniques positively impacting liking (Ackerman & Hilsenroth, 2003) presented a list: being flexible, experienced, honest, respectful, trustworthy, confident, interested, alert, friendly, warm, and open. Among techniques with positive effects were exploring, seeking depth, reflecting, being supportive, noting past therapy success, interpreting, facilitating expression of affect, being active, affirming, understanding, and attending to patient's experience. The reader can see that much on the list belongs to a superordinate category of 'being in focus'. Moreover, communicating a sense of hope, being open-minded, referring to common experiences, and clarifying, were also regarded as positive. A difficulty with lists is that the distinction between personal attributes and techniques is sometimes unclear. For example, among techniques is noted "communicating a sense of hope", which is similar to the

attribute “hopeful”. Furthermore, the message of hope can be accomplished in many different ways (e.g., saying “It will get better” or smiling). This implies that being on a ‘good’ list does not mean that the attribute or behavior is trainable.

Nondirective counseling

If counseling means to listen, support, advise, or recommend professionally on social problems, then nondirective counseling should be an equivalent process, albeit with less emphasis on advice. Thomas Gordon residing at the University of Chicago Counseling Center (Grummon & Gordon, 1948) articulated this already in the 1940s. In that work place, the *philosophy* (the term used in the book; see also Gordon, 1950) of nondirective counseling was applied to all activities, from counseling practice to staff meetings. Although Gordon used the terms counseling and therapy interchangeably (Grummon & Gordon, 1948), the low mean number of counseling sessions at the center (5.2) clearly shows that the help given in nondirective counseling at that time was much briefer than any traditional psychotherapy.

The nondirective counselor is more concerned about the client’s perceptions of her or his problems, than the problems *per se*. The counselors task is to handle the client’s emotionalized attitudes and feelings non-judgmentally, leading to exploration, insight, and change in behavior (Grummon & Gordon, 1948). Such an approach places much responsibility on the client, who generally was allowed to use the Counseling Center according to own needs and wishes. For example, the client set the number of sessions and the time-interval between them, although most clients came once or twice a week. Nor were any formal records kept about the client, with the exception of information needed for billing.

Nondirective counseling has always been an important component of client-centered therapy, which in its turn is closely related to Carl Rogers’ person-centered theory of personality. The following text describes Rogers’ influence on nondirective counseling, and how this influence through Thomas Gordon, a co-worker and follower of Rogers, made techniques used in nondirective counseling known to a wider audience.

Carl Rogers and client-centered therapy

Having a background in developmental and clinical child psychology, Carl Rogers (1902-1987) worked as an executive secretary at the Counseling Center at the University of Chicago in the 1940s. His most productive period

was 1945-1957, during which he became well established in American psychology, and his so called client-centered (later person-centered) approach gained in reputation. Rogers had been elected president of the American Psychological Association in the mid-40s, and in 1956 he received the APAs Scientific Contribution Award (Cissna & Anderson, 1994). In 1957, he left his position as director of the Counseling Center for an appointment at the University of Wisconsin, and later Ohio State University.

Rogers' career went through two theoretical phases, although he always put more emphasis on helping people than on theory development. The first nondirective/mirroring phase (Rogers, 1942) proposed that the client moves from negative to positive feelings, and reaches insight as a result of the therapist's recognizing, clarifying, and accepting all of the client's feelings. Rogers' second and more attitude-focused phase (Rogers, 1951, 1957, 1975) shifted attention from techniques to the therapist's attitudes to the capacity of the individual, making the client's phenomenological experience increasingly important. This 'softer' variant of the counseling process culminated in Rogers' (1957) treatise on the three necessary and sufficient conditions of therapy: congruence, unconditional positive regard, and empathy. Of these constructs, empathy has received most attention. In the 1960s, Rogers left Ohio State University to do work with schizophrenic patients, and later to become a spokesperson for nondirective education, conflict resolution, and global peace.

Nondirective counseling and client-centered therapy received criticism from contemporary proponents of psychoanalytically oriented psychotherapies. They regarded nondirective counseling and psychotherapy as very distinguishable processes (e.g., Hathaway, 1948/2000), and feared that if limits between counseling and therapy were not set, other professional interaction could soon be indiscriminately classified as psychotherapy as well. Rogers, who had been trained in psychoanalytic methods, reacted calmly to the objections. He articulated that most clients could strive in the right direction themselves, if only there was someone who was prepared to share their experience. He claimed that he had discovered that simply listening to a client was helpful. So when he did not know what to do (in the psychoanalytic sense), he simply listened, especially to the client's emotions (Rogers, 1975). He stated that usually it was very limited verbal responses such as reflecting feelings that were responsible for change to the better or worse, and that the psychoanalytic emphasis on interpretation and clarification was sometimes on a 'too high level'.

The first phase of Rogers' teaching came to be known as a technique. Reflecting the client's feelings was coined 'nondirective therapy', and was sometimes presented — perhaps to amuse those who had criticized him — as a parrot, a caricature of what he had originally intended (Cissna & Anderson, 1994). Therefore, even today it is sometimes insisted that the

therapeutic and the commonplace use of reflective listening are essentially different (Sundararajan, 1995), although there is no definition on what makes reflective listening 'authentic'. Rogers himself insisted on the importance of personality characteristics, especially empathy, presenting research showing that a majority of therapists rate empathy as the most important characteristic of a therapist, being correlated with self-exploration, and predicting success in counseling (Rogers, 1975). Rogers suggested that empathy is a unique quality in a relationship, and that therapists offer more of it than even helpful friends can, depending on the self-integration of the therapist, and her or his experience.

Carl Rogers defined empathy as "entering the private perceptual world of the other /.../ being sensitive /.../ living in his/her life /.../ sensing meanings of which he/she is scarcely aware /.../ communicating your sensings of his/her world /.../ checking with him/her as to the accuracy of your sensings" (Rogers, 1975, p. 4). This, and earlier references (e.g., Rogers, 1957) imply that empathy needs to be communicated, and he provided examples of empathic communication, at so concrete level that they could be described in a manual. Rogers held that "an empathic way of being, though highly subtle conceptually, can also be described in terms which are perfectly understandable" (1975, p. 5), and that "the ability to be accurately empathic is something which can be developed by training" (1975, p. 6).

There is certainly more than one view on empathy (Riggio & Taylor, 2000). An example of the differences between perspectives on empathy concerns the distinction between cognitive empathy (taking another's perspective) and emotional empathy (experiencing another's emotions). Riggio, Tucker, and Coffaro (1989) argued that empathy could be defined as basic communication skills as well as personal characteristics of an individual. For example, some skills help us to sense another person's emotions, and other skills help us to communicate this understanding. Women scoring higher than men on most empathy scales, while also more skilled in nonverbal behavior, have supported this view on empathy. Further support comes from correlations between empathy scales and the Social Skills Inventory (SSI; Riggio, Tucker, & Coffaro, 1989). A view of empathy consisting of social skills implies that empathy can be trained, and this might have been the *primus motor* of Carl Rogers' emphasis on communication skills before he started to focus on the individual characteristics of counselors. It may also have influenced Rogers' former collaborator and follower, Thomas Gordon, when he set out to adapt nondirective counseling skills for everyday use.

Thomas Gordon and reflective listening

Thomas Gordon (1918-2002) was an assistant professor and counselor at the Counseling Center at the University of Chicago. He was a student of Carl Rogers and later worked together with him. Whereas Rogers successively became more and more concerned with the inner attitudes of therapists, Gordon remained with the earlier technique-oriented views of nondirective counseling. Accordingly, during the 1960s and 70s, Gordon started training parents in reflective listening (although he used the term active listening; Gordon, 1970), and later extended his 'mission' to include other groups as well. Through the expansion of his training corporation, and his and his wife's writings (e.g., Adams, 1989), the "Gordon method" has now been taught in 43 countries and 26 languages to over a million people.

Rogers had been puzzled by the impact of the degree to which the counselor overtly communicates empathy to the client (Rogers, 1957). He stated that a means to operationalize the communication would be to let external judges rate the communication. This is what Thomas Gordon eventually laid his emphasis on, namely the outside manifestation of empathy in skills, and how these skills could be trained. And while Rogers eventually did not regard techniques such as 'reflecting feelings' as essential to therapy (Rogers, 1951), Gordon evidently held another opinion. Another difference between them was that whereas Rogers concentrated his efforts on client-centered therapy, the professional use of nondirective counseling, Gordon decided to adapt the method to more commonplace relationships, such groups, parents, leaders, and teachers (Gordon, 1977).

Reflective listening

Thomas Gordon used the term active listening instead of reflective listening. I have not found any written material suggesting that these two constructs denote different processes or counselor behaviors. It can be speculated that Gordon chose another term in order to put emphasis on the active role of the listener, or that he simply wanted to give it a name not so strongly associated with Carl Rogers' theory. However, through the work of Gordon, reflective listening has become a widely known ingredient in a variety of dyadic contexts that involve gathering information and solving problems, for example between counselors and clients, parents and children, and spouses. In this way, counseling skills described by Rogers have reached a wider audience than perhaps originally intended.

Reflective listening (RL) occurs when someone in a conversation tries "to understand what it is the sender is feeling or what his message means. Then he puts his understanding into his own words (code) and feeds it back for the

sender's verification" (Gordon, 1970, p. 50). Another definition of RL is that someone repeats in own words what another person has said, including both factual and emotional information, and adding nothing (Lindh & Lisper, 1990). This definition shifts focus from the listener's inherent attitudes to her or his overt behavior, and has also been the foundation of the CST conducted at the Department of Psychology at Uppsala University, where the present empirical studies were conducted.

Professional helpers, such as counselors and psychologists, commonly use RL, and it is emphasized that the content of another person's story should not to be altered. As Gordon (1970) put it, "The receiver does not send a message of his own — such as an evaluation, opinion, advice, logic" (p. 53). By giving such advice, Gordon disqualified many of the verbal responses typically found in sessions conducted by counselors who are regarded as professional by their peers (e.g., interpretation; Hill et al., 1988; Nagel, Hoffman, & Hill, 1995). The purpose of RL is to encourage someone to talk by verifying the message in a non-judgmental fashion, and in this way it is nondirective. Both Rogers and Gordon have given written examples of RL, but more detailed descriptions have been given elsewhere. For example, Lindh and Lisper (1990) operationalized RL as seven response modes (categories of utterances): Minimal encourager, Direct encourager, Reflecting fact, Reflecting emotion, Recapitulation, and open- and closed-ended Question on fact and Question on emotion. It is essentially these categories that have been used in the present research, and some of them need explaining.

Minimal encourager, sometimes called minimal response, is a short utterance that most listeners do automatically, such as saying "Uh-um" or "Yes", or nodding (e.g., Davis, 1986; Mansfield, 1991; Ralph & Thorne, 1993). Direct encourager is to explicitly encourage the other to continue a story, for example saying "Go on", "Continue", or "Tell me more". Reflecting, sometimes called paraphrasing, restating, or formulating (e.g., Hill, 1992; Horne, Vatmanidis, & Careri, 1994; Knippen & Green, 1994; Lin, Kelly, & Nelson, 1996; Phillips, 1999), is restating the other's utterance in one's own words. The difference between open- and closed-ended questions is that the latter requires only a short answer, such as "Yes" or "Twelve weeks ago", while open-ended questions encourage telling a story. Of these seven categories of verbal responses: Reflecting fact, Reflecting emotion, Question on fact, and Question on emotion can be regarded as the core skills of RL, because they involve exploring and understanding the sender's message, and putting it into own words according to Gordon's (1970) definition. If there were a paragon RL skill, however, then it would still not be easy to define, because successful listening involves a mixture of many different verbal responses.

Carl Rogers stated that RL changes a relationship for the better for all interacting parties. The listener experiences little rebellion from the client, because there is mutual respect. The communication opens up, and the listener learns to accept both positive and negative feelings toward the other. Moreover, the person engaging in RL does not make the other's problems his own, but facilitates the other's problem solving. The other feels honest, independent, accepted, and in charge of her or his own life (Rogers, 1951, 1961).

It has been proposed that a counselor must get the client to collaborate in formulating treatment goals and plans, and that no client will examine the proposals and ideas of an unlikable therapist (Hayes, Follette, & Follette, 1995). A good relationship has been considered important in physiological treatment as well (e.g., Horne, Vatmanidis, & Careri, 1994), with claims such as “[a]ctive listening is the single most important skill needed by the clinician in communication” (Mansfield, 1991, p. 488). The Rogerian viewpoint implies that one should not view a client's undesirable behavior primarily as something inherent in the client. Instead, one should attribute it to the interaction between the client's and the counselor's verbal behavior, goals, and attitudes. The best example of counselor verbal behavior is, of course, what words are used.

Research on reflective listening

Much previous research on RL has been embedded in training packages of nondirective counseling (as defined by Rogers), and active listening or Parent Effectiveness Training (as defined by Gordon). A large proportion of this research seems not to have been published, apart for some general conclusions reported in books that describe and advocate the method (e.g., Gordon, 1970; Rogers, 1957). Some research, however, has been published as meta-analyses and individual reports.

Reviews and meta-analyses

Rinn and Markle (1977) performed one of the earliest reviews on the “Gordon method”. It was a qualitative investigation of 14 studies on Parent Effectiveness Training, a CST usually consisting of eight 3-hr sessions. It was not an outcome study. Instead it reviewed the methodological shortcomings of the studies. Rinn and Markle concluded that outcomes were contaminated by nonrandom assignment of participants to conditions, and that objective behavioral criteria were seldom used to measure change. Moreover, they reported that many Parent Effectiveness Training studies did not include control groups or follow-up assessment. More than a third of the

studies had investigated single-group outcomes, making the results vulnerable to threats to internal validity. Nine of the studies included control-groups, the control usually being a waiting list or some kind of discussion group. The external validity of the findings was criticized: would the trained participants ever use the skills in a natural setting?

In an attempt to use Glass' meta-analytical method, Cedar and Levant (1990) reviewed some 60 studies on Parent Effectiveness Training, performed between 1975 and 1990. They too suggested that many of the studies did not fulfill important methodological criteria, which is why it was difficult to make decisive statements about the effectiveness of RL. They concluded, however, that the largest effects were found in parent course knowledge, attitudes, and child self-esteem, while the effects in parenting behavior were smaller. In other words, in these studies training mainly influenced attitudes instead of behavior.

A couple of years earlier, Dembo, Sweitzer, and Lauritzen (1985) had compared 18 Parent Effectiveness Training studies with 30 other group parent education programs. They stated that all programs, regardless of theoretical orientation, tried to target both awareness and skills, although not all evaluations included control groups, random assignment, specified procedures, or multiple dependent measures. The standard programs for Parent Effectiveness Training addressed children up to the mid-teens but only two of the 18 reviewed studies included independent observer data. Furthermore, only three of the 18 studies were published at the time of the review. Generally, however, the studies reviewed by Dembo et al. (1985) did not report significant effects of Parent Effectiveness Training.

In summary, reviews of Parent Effectiveness Training have yielded few conclusions as to whether the training gives any results.

Individual reports

Graybill (1986) hoped to shed light on weaknesses in Parent Effectiveness Training, such as nonrandom assignment, no control groups, no multiple-outcome criteria, and overreliance on self-reported data. He introduced yet another term for the Rogerian method, reflective consulting, which denotes a parent's understanding of a child's feelings, acceptance, and feedback. He investigated 32 parents with their children during the course of a 12-hr training, and concluded that the parents who had trained were less anxious than controls, and that they used more RL. Unfortunately, there were no effects on their children. Unfortunately as well, Graybill did not account for how RL was measured, although the reported use of questionnaires suggests that it was not done by observation.

A study by Darsten, Lisper, and Sohlberg (1979) compared ward psychologists with RL-trained psychology students. Both groups had

assessment interviews with psychiatric patients, after which the patients were asked about their impression. The results showed that RL made the patients feel more understood, although they felt a larger benefit from the interviews with the real psychologists. The authors concluded that questions in an interview might feel good, because they give the patient a structure to her or his problems. They also concluded that a 20-hr training of RL could teach empathic understanding.

A later study (Lisper & Nilsson, 1982), involving 2- to 6-year old children in kindergarten, observed how crying was affected by different verbal responses from the staff. The staff was trained to respond to the children's crying according to different methods: RL, detraction (e.g., saying "Be a big boy"), belittling ("It'll soon be over"), and distraction ("Let's see what the other kids are doing"). The results showed that when the staff used RL the children stopped crying after a shorter period ($M = 10$ sec) than when they used other verbal responses ($M = 90$ sec). The authors proposed that the child's relationship to the person who responds may be of less importance than the response itself.

In 1987, Wood and Davidson pointed out the importance of evaluating Parent Effectiveness Training because of its wide popularity. They compared nine trained parents to 10 matched control group parents. The Parent Effectiveness Training was the customary eight weekly 3-hour sessions. They found that parents did learn the skills at least to such a degree that they could write down what would normally constitute a RL-response. However, they also investigated the parents' attitudes to authoritative child rearing, and reported that they had not changed.

Nearly a decade later, Nugent and Halvorson (1995) suggested that RL does not consistently make the client feel understood. One factor impacting the impression of RL was named Type A and Type B reflective listening. Type A is when the listener takes no stance to whether the client's view is correct or not: it is neutral RL. Type B is when the listener validates the client's story, for example by eliminating words such as "possibly" or "may". The participants rated different types of RL in scenarios that provoked negative affect. The results showed that neutral RL gave lower ratings on depression and anger. The conclusion was that practitioners need to pay attention to how they phrase their RL-responses.

I was involved in a study that examined short CST of psychology students (Lisper & Rautalinko, 1996). The study compared supervised role-play, written exercises, and discussion conditions. The purpose of all conditions was that the students should learn to use RL in subsequent conversations with a fictive client. The results showed that even a 6-hr training increased the use of RL-responses in the trained participants compared to controls. Unexpectedly, the results did not show that more compared to less RL affected how the fictive client experienced the short counseling sessions.

Conclusions from previous research

Taken together, nearly all investigators have highlighted the methodological limitations of previous research on RL. When there has been effects of RL, it has mainly consisted of changes in attitudes in those who have received a counseling skills training. It has, however, been more difficult to show attitude or behavior differences in those who are listened to. This means that, generally, the support for RL has been sparse.

EMPIRICAL STUDIES

The first aim of this thesis is to report by what means the communication skills of professional helpers can be enhanced in a short time. The starting point was a CST that has been given to psychology students during their third term at the Department of Psychology at Uppsala University. As this CST has focused on nondirective counseling skills, the second aim of this thesis is to report how RL compares to more directive (e.g., untrained) communication skills, regarding how the skills are perceived by other people. The third aim, finally, is to provide some evidence on how the perception of more or less directive counseling skills is moderated by the people's individual characteristics.

Research questions and method summary

The thesis includes three empirical studies (Paper I-III) intended to shed light on the following research questions: What effects does a short training of RL have? Do the effects increase if the training time is increased? Do the effects generalize to naturalistic settings, and for how long do they persist? Is a professional helper who uses nondirective communication skills regarded as better than a helper who is more directive? If so, is this effect moderated by individual characteristics of a help-seeker?

Paper I and II include a short RL training, followed by an evaluation of the acquired communication skills. Paper III does not include any training, instead it focuses exclusively on the evaluation of possible effects of a short training.

Below is a summary of the procedures, measures, and statistical analyses used in the present papers. The final section is an analysis investigating Paper I and II for how the sex of the participants may have moderated the effects of training.

Participants

The participants can be divided in three groups: those who have participated in CST, those who have participated in interactions subsequent to CST, and those who have evaluated these interactions.

Those who have participated in CST have been insurance claim adjusters and psychology undergraduates. Those who have participated in interactions subsequent to CST (i.e., help-seekers) have been psychology undergraduates, insurance customers, clinical psychologists, and other less specified people. Those who have evaluated interactions have been psychology undergraduates, nonpsychology university students, high school students, vocational counselors, and people selected from a telephone catalogue. For details, consult the original papers.

Communication skills training

The CST described in Paper I and II resembled the training that students at the Psychologist program at Uppsala University receive during their third semester of education. Today, it is 15 hours consisting of RL, I-messages, and the No-Lose problem-solving method (see Gordon, 1970). All of the training activities somehow touch upon the nondirective techniques of reflecting fact and emotion. Moreover, it is taught that the participants choose open-ended instead of closed-ended questions. The training usually begins with a demonstration of a counseling skill, followed by role play in groups of 3 to 5 students. The total number of students per supervisor in the present studies has been approximately 15.

Taping, transcription and response categorization

The sound and picture quality in the audio- (Paper I-III) and videotapings (Paper II) was good throughout. Material must sometimes be discarded due to poor quality, but this was fortunately not the case in the present studies. The time-length of the tapings varied from 7 to 22.5 min, but the sequences used for transcription and content analysis were mostly a 5-min interval 'in the middle'. An exception was the Paper I follow-up Study 2, where eight employees had taped conversations that were 2 to 11 min long, in which all responses were analyzed.

I made transcriptions of all conversations, except the videotapes of Paper II Study 2 that were not transcribed at all. The transcriptions were subsequently content analyzed (Holsti, 1969; Krippendorff, 1980; Weber,

1990). This is a mix between the quantitative survey and the qualitative observation or interview, with the goal of classifying qualitative themes in written and oral materials. The common steps in content analysis is developing categories, deciding on a coding procedure that can be manualized, training coders, coding the material, analyzing the data, and perhaps assessing interrater and intersource reliabilities. It is common that themes are identified from earlier research or theory, and that reliabilities are measured with percent agreement or Cohen's kappa coefficient (Schneider, Wheeler, & Cox, 1992).

The content analyses in the present studies were carried out by me. The unit chosen for coding was a verbal response, defined as one dyad member's vocalization bounded by the other member's vocalization. The a priori RL categories found in the employees' responses were Minimal encouragement, Direct encouragement, Reflecting fact, Reflecting emotion, Question on fact, and Question on emotion. Questions were categorized into open- and closed-ended. The non-RL categories that emerged were Interpretation, Advice, and Offering help.

Repeated content in a response was coded only once. The first and last minutes of conversations were often excluded, as they included greeting- and goodbye-responses. Interrater reliabilities between the authors (Paper I-II) were calculated with Cohen kappa coefficients that ranged between .73 and 1.00. Category drift for the first author's coding was assessed after 10 months (Paper I), and 3 or 16 months (Paper II), with Cohen kappa coefficients that ranged between .87 and 1.00.

Dependent measures

There were three dependent measures that were used in the rating of the audio- or videotaped interactions. These were the Positive evaluation index, Open-ended descriptions, and the Working Alliance Inventory.

Positive evaluation index

The positive evaluation index was used in all three studies. Beginning in Paper I, the employees were rated on a 10-point scale ranging from 1 (*not at all*) to 10 (*very much*) on their competence, confidence, and compassion. These aspects were chosen, based on face-validity, as three variables that describe a good employee/counselor, and were eventually retained in the other two investigations. The ratings were, however, highly related across subjects ($r = .59$ or above), and were averaged to a Positive evaluation index ranging from 1 to 10. The Cronbach coefficient alpha of the Positive evaluation index was .91 or above.

Open-ended descriptions

Open-ended descriptions of the counselors were used in Paper II and III. They were blank lines on which the respondents were invited to continue the sentence “I think the counselor...” It was considered important to include an open-ended measure together with the confined measures, because it was difficult to know in advance what impression the conversations would make on the external judges.

All open-ended descriptions were categorized by me. The categories were set to positive (e.g., “calm”), negative (e.g., “in a hurry”), and neutral (e.g., “careful”). In a further content analysis of the positive valence descriptions, subcategories such as “Interested”, “Empathic”, and “Relevant” emerged. The negative valence descriptions belonged to subcategories such as “Insecure”, “Dominant”, and “Detached”. The neutral valence descriptions were not analyzed in detail, but a large proportion of them concerned body-posture or facial features, or behavior that is not positive or negative in itself, such as “The counselor poses questions”. There were also comments on the client such as “The client thinks out loud”. In Paper II, only the positive valence descriptions provided significant differences between the conditions. In Paper III, the correlations between the number of positive and negative open-ended descriptions were relatively high (r s around .50), so they were combined into an Open-ended index consisting of the number of positive subtracted by the number of negative descriptions.

Obviously, the categories could have been other than the positive, negative, and neutral valence categories. In order to estimate the reliability of the categorization, a philologist who was blind to the purpose and method of the studies categorized randomly selected subsets of the open-ended descriptions. Cohen’s kappa for the present author’s and the philologist’s independent categorizations was .79 or above.

Working Alliance Inventory

The Working Alliance Inventory (WAI; Horvath & Greenberg, 1989) was used in Paper II and III. It consists of 36 statements answered on a 7-point scale ranging from 1 (*never*) to 7 (*always*). It includes three subscales that assess the development of bonds (attachment), agreement on goals (outcomes), and assignment of tasks (behaviors and thoughts), giving a composite score between 36 and 252. A high score means that a counselor and a client agree on the problem areas to be worked on, and that there is mutual trust and approval in their relationship.

WAI was designed to measure some of the generic (nonspecific to theory or technique) variables affecting success in counseling. Its background is fourfold: Rogerian client-centered therapy, Strong’s social influence theory

(success is proportional to the degree the client perceives the counselor as trustworthy, expert, and attractive; Horvath & Greenberg, 1989), Freudian idea of client's realistic attachment, and Bordin's integrationist formulation that a patient's compliance to treatment depends on the relationship between her or him and the therapist (Bordin, 1979).

The internal consistency of the WAI was investigated by calculating Cronbach alpha coefficients. They ranged between .80 and .88, which is in line with previously reported reliability data (Horvath & Greenberg, 1989).

Moderating factors

Two measures were used to investigate whether the ratings of the audio- and videotaped interactions would be moderated by individual characteristics of the evaluators. These were the Social Skills Inventory, and Raven's Standard Progressive Matrices.

Social skills

Paper II and III attempted to measure the social competence of those who evaluated conversations. This was carried out using the Social Skills Inventory (SSI; Riggio, 1986), a 90-item assessment of six basic social skill dimensions: sending, receiving, and controlling communication in two separate domains, namely the emotional-nonverbal and social-verbal domains. Each item is answered using a 5-point Likert scale ranging from 1 (*not at all like me*) to 5 (*exactly like me*), and the total score, between 90 and 450, approximates the respondent's global social competence, with a higher score indicating greater social competence.

Social competence is one of the most popular concepts of social psychology. It can be defined as the appropriate use of responses in social interactions (Odom & McConnell, 1992). The definition implies that social competence depends on a match between behavior and the context. Social competence is measured as a personality characteristic, and is supposedly manifested in behavior that generalizes over many contexts.

One purpose of the SSI is to discriminate between social competence and general intelligence. It was not designed to measure any general skill, but a multidimensional compound of individual characteristics. Interrelations between these are complex, and so is the nonlinear relation between any of them and social effectiveness, that is, outcomes that are desirable for the individual (Riggio, 1986). That is why the composite scores of the SSI were used in the present studies.

The internal consistency of the SSI was calculated using the Cronbach alpha coefficient. These coefficients ranged between .84 and .89.

Cognitive ability

Paper III included a measure of cognitive ability, Raven's Standard Progressive Matrices (SPM; Raven, Court, & Raven, 1992). The SPM was originally introduced as a measure of nonverbal "eductive ability", and has been proposed to measure one of the two components in the Spearman g factor. The SPM consists of five series of 12 puzzles in which a piece is missing, and the respondents' task is to indicate which of the 6 or 8 suggested pieces is the missing one. Answers are scored as correct or incorrect, resulting in a total score between 0 and 60.

Reliabilities for the SPM were not calculated, because it is such a widely used instrument. Internal consistency of the English version of the SPM has been reported to lie between .97 and 1.00, that is, extremely robust (Raven, Court, & Raven, 1992).

Design and statistical analyses

The designs in Paper I and II were similar for many reasons. Both included CST that was preceded and followed by an evaluation where the trainees met with confederates. Both had a follow-up as well as an external evaluation of the conversations that had been audio- or videotaped. The design in Paper III was different. There was no CST; instead, selected conversations were manipulated, audiotaped, and evaluated. Consequently, the statistical analyses in Paper I and II were more complicated than the analyses in Paper III.

Paper I

The design in Paper I was quasi-experimental. As the participants were not randomly assigned to conditions, and as the confederate's roles at pre- and posttest were not the same, ANCOVA was used to test the effects of training on response categories and judgments, using corresponding pre-training data as covariates. The ANCOVA was preceded by a MANCOVA where the directive and nondirective response categories, respectively, were grouped together. For responses found only at posttest, the data from the participants in the two conditions were compared with unrelated t tests.

Paper I considers other statistical options for a pretest-posttest design than the ANCOVA (Pedhazur & Pedhazur Schmelkin, 1991). I argue that an analysis of change scores by a one-way ANOVA or a t test was not applicable, because pre- and posttest scales were not identical. The other option would have been a two-way repeated measures ANOVA with the trials as a within-group factor. Opinions on the choice between them differ

(for details see Paper I) so all data were analyzed with both procedures. As the conclusions are the same, only the ANCOVAs are presented.

Paper II

The design in Paper II was experimental with random assignment of participants. Here too, the directive and nondirective response categories, respectively, were grouped together in a MANOVA. As there were statistically significant differences, the response categories were further investigated in repeated measures ANOVAs.

Paper III

The two studies in Paper III were factorial experiments with two independent variables. The results were analyzed using factorial repeated-measures ANOVAs. In addition, there were effects of the moderating variables social skills and cognitive ability. These effects were included using the Statistica software's General Linear Model-analysis, with the moderating variables as continuous predictor variables.

Paper I

The research in Paper I took form when a large Swedish insurance company wanted to enhance its employees' communication skills, effectiveness, and customer satisfaction. In order to reach these goals they wanted a CST that was to be conducted in groups of 10 to 20 people. The training that eventually started was similar to the training received by the students at the Department of Psychology. Originally, it was to be evaluated using a large-scale customer satisfaction index collected through questionnaires distributed to the company's customers each year. This approach was later regarded as too intrusive, and was replaced by a quasi-experiment involving a fictive female client. The design of Paper I was comparable to a design used in another study some years earlier (Lisper & Rautalinko, 1996).

The general hypothesis was that training given to the insurance company employees would affect both the employees' and the customers' behaviors and emotions. More specifically, it was expected that trained employees use more RL as compared with untrained, and generalize this to an authentic customer setting. Moreover, it was expected that trained employees receive more positive evaluations of their communication skills than untrained do.

Study 1

The chosen training group consisted of 12 company employees. They were middle-age middle-level managers with long professional experience. The control group of 9 employees resembled the training group participants, although perhaps not enough (see Paper I for a discussion).

The training group practiced RL, I-messages, and the No-Lose problem-solving method (see Gordon, 1970) during four consecutive days, of which a total of 16 hrs were spent on RL. The authors supervised the training that consisted of modeling and role play in small groups which aimed at teaching the participants to use RL, and choosing open-ended instead of closed-ended questions when gathering information from customers. The Control group received no training.

Before and after training, all employees had short audiotaped conversations on the telephone with a fictive customer. The customer's roles covered two frequent topics that were different at pre- and posttest. Telephone contact was chosen for convenience, but was justified by the expanding number of insurance cases settled by telephone in the present company. The confederate evaluated the employees on the Positive evaluation index.

The conversations were transcribed and all participants' responses were categorized into RL and other verbal responses and quantified. Coding was content analytical (Holsti, 1969; Krippendorff, 1980; Weber, 1990), the categories being given by theory on nondirective counseling (Gordon, 1970) and previous research (Lisper & Rautalinko, 1996). The RL categories that were found were similar across all studies in the present thesis: Minimal encouragement, Direct encouragement, Reflecting fact, Reflecting emotion, Question on fact, and Question on emotion. The non-RL categories that emerged were Interpretation, Advice, and Offering help. Moreover, questions were categorized into open- and closed-ended, and the proportion of open-ended questions of all questions was calculated.

Table 1 presents the number of the employees' RL-responses per minute (percentages of all responses in parentheses)¹. A MANCOVA was carried out on the posttest *z* scores of the employees' nondirective and directive response categories, using corresponding pre-training data as covariates. There was a statistically significant difference between the Training and Control groups for the nondirective but not for the directive responses, as the training group displayed a relatively higher use of nondirective responses. A further examination showed statistically significant differences for

¹ In the original paper, the number of responses for the whole sequence was presented. Here, the number per minute and percentages are presented to make comparison easier. Observe that the percentages do not add up to 100.

Reflecting fact and Reflecting emotion. The descriptive statistics of Table 1 can give the impression that the ANCOVA main effect is due to an impairment in the Control group. The role play before and after training, however, should not be regarded as a repeated measure, but as two different role plays.

Table 1

Employees' RL-Responses at Pre- and Posttest: Number per Minute (Percentages of All Responses in Parentheses)

RL-response	Training group		Control group	
	Pre	Post	Pre	Post
Reflecting fact	0.3 (6)	0.8 (15)	0.3 (6)	0.2 (5)
Reflecting emotion	0.3 (5)	0.3 (6)	0.4 (7)	0.0 (1)
Question on fact	1.2 (22)	0.7 (12)	0.5 (9)	0.6 (11)
Question on emotion	0.1 (2)	0.1 (2)	0.1 (1)	0.0 (1)

The ANCOVA of the confederate's Positive evaluation index ratings of the employees at posttest, with pre-training rating as a covariate, showed a statistically significant difference ($p = .01$) in favor of the trained employees: the score difference between the Training and Control groups was greater after than before training, and in favor of the Training group.

Study 2 and Study 3

The follow-up Study 2 was carried out 5 months after training, by four employees in each of the Training and Control groups. They audiotaped approximately 25 min of conversations on the telephone with real customers, but only the employees' responses were analyzed. Response categorization was carried out by content analysis similar to that used in Study 1.

The employees had audiotaped between four and seven conversations each, adding up to 21 conversations in both groups. Table 2 presents the employees' use of RL-responses at follow-up. The employees in the Training group used significantly more Reflecting fact and Reflecting emotion. There were, however, no statistically significant differences between the groups regarding their use of directive responses.

Table 2

Employees' RL-Responses at Follow-Up: Number per Minute (Percentages of All Responses in Parentheses)

RL-response	Training group	Control group
Reflecting fact	2.0 (9)	0.4 (2)
Reflecting emotion	0.9 (4)	0.2 (0.7)
Question on fact	3.1 (14)	3.6 (16)
Question on emotion	0.2 (0.1)	0.0 (0.2)

The validation Study 3 investigated whether trained employees receive more positive evaluations of their conversation skills than untrained do. The purpose was to test the external validity of the confederate's assessments by using a larger sample of judges. It was also investigated whether employees using more RL as compared to less RL would be more positively evaluated.

Participants were 54 high-school students and psychology undergraduates. They were randomly divided into two samples that listened to and rated the audiotaped posttest conversations of Study 1. The first sample listened to the 11 employees (8 trained and 3 untrained) who used more RL. The second sample listened to the 10 employees (4 trained and 6 untrained) who used less RL. The median split of RL was carried out using the summed z scores across Direct encouragement, Reflecting fact, Reflecting emotion, Question on fact, and Question on emotion. The participants used the same rating scale as the confederate in Study 1. The results, however, showed effects of neither training nor the use of RL.

Paper II

The research for Paper II span off from two questions: is the use of specific nondirective counseling skills responsible for differences in the therapeutic relationship between a counselor and a client? If they are, how much time is required to train these nondirective counseling skills? The concept of therapeutic relationship or alliance has its origins in Sigmund Freud, who in his publication "On beginning the treatment" (1913; cited in Ackerman & Hilsenroth, 2003) believed that there was a positive transference from the patient to the therapist. Eventually, the therapeutic or working alliance has become to pantheoretically denote the conscious collaboration between a patient and a therapist. The relationship between alliance and outcomes has been studied extensively (Ackerman & Hilsenroth, 2003) whereas studies of the therapist's contribution to the alliance are lacking, with respect both to her or his attributes and techniques.

The purpose of Paper II was to compare a CST, at that time 14 hours long, with two other similar procedures that were twice and three times longer, respectively. The general idea was that training affects the students' as well as their clients' behaviors and attitudes. More specifically, it was expected that students who had trained for a longer time use more RL, get more information from their clients and remember it better, and retain their skills better at follow-up. Moreover, it was expected that those who had trained for a longer time would reach a better working alliance and receive more positive evaluations of their counseling skills.

Study 1

The participants were 28 psychology undergraduates who were randomly assigned to three training groups: 14, 28, and 42 hours of training RL. The three procedures were multiples of the same 14-hr training program focusing on RL, I-messages, and the No-Lose problem-solving method (see Gordon, 1970). There were seven sessions over two weeks during which the participants were taught to use RL and choose open-ended instead of closed-ended questions. The shortest of the programs (14 hrs) was almost identical to the training presented in Paper I. The intermediately long training program (28 hrs) was comparable in length to the customary 'Gordon method' counseling skills programs (24 hrs; Rinn & Markle, 1977).

Before and after training, the students met with two confederates who acted as clients. They had short counseling conversations with both confederates whose four roles covered ordinary topics in personal counseling. These counseling conversations were audiotaped, and the confederates evaluated the students on the WAI. The conversations were transcribed and the participants' responses were categorized into RL and other verbal responses and quantified. Quantification was preceded by a content analysis similar to the one presented in Paper I, and the categories that emerged were also the same.

Table 3 presents the number of the students' RL-responses per minute (percentages of all responses in parentheses). A repeated measures MANOVA was applied on the number of RL-responses, resulting in a statistically significant effect of time between pre- and posttest, but a nonsignificant effect of training group. The separate repeated measures ANOVAs of the separate RL-responses essentially replicated the result. This means that the students in all training groups increased their RL, although a longer training did not have larger effects than a shorter.

Table 3

Students' RL-Responses at Pre- and Posttest: Number per Minute (Percentages of All Responses in Parentheses)

RL-response	14-hr		28-hr		42-hr	
	Pre	Post	Pre	Post	Pre	Post
RF	0.7 (11)	1.0 (15)	0.6 (11)	1.3 (19)	0.7 (9)	1.5 (20)
RE	0.2 (3)	0.4 (6)	0.3 (6)	0.6 (9)	0.2 (3)	0.9 (13)
QF	1.7 (28)	1.1 (16)	1.7 (29)	0.6 (8)	1.8 (25)	0.7 (9)
QE	0.5 (8)	0.3 (5)	0.5 (8)	0.1 (2)	0.6 (8)	0.1 (1)

Note. RF = Reflecting fact, RE = Reflecting emotion, QF = Question on fact, QE = Question on emotion

The confederates' responses were analyzed as well. The results suggested that the confederates did not talk differently to the students in the three training groups as a function of the students' training. The analysis of the students' recall of information, however, showed that the students in the 42-hr training group remembered the information given by the confederates better than the other students did. There were no significant main or interaction effects on the confederates' WAI ratings, suggesting that the confederates did not experience any differences between the students in the three training groups.

Study 2

The purpose of Study 2 was to follow up the training in Study 1 after 12 months. It investigated whether the students who had trained RL for a longer time would retain their skills better, and whether the clients would give them more information. The procedure did not involve any additional training, instead the data were collected during another counseling skills course.

The follow-up participants belonged to the previous training groups with some attrition. Each of them videotaped a longer (compared with Study 1) counseling session with one of the fictive clients who acted in new roles. The clients did not, however, evaluate the students. The conversations were coded verbatim and the participants' responses were categorized into RL and other verbal responses and quantified. A content analysis similar to the one presented in Paper I was used, and the response categories were set identical to those in Study 1.

Table 4 presents the number of the students' RL-responses per minute (percentages of all responses in parentheses). A MANCOVA was carried out

on the follow-up scores of the students' RL-responses, using corresponding pre-training data as covariates.

Table 4

Students' RL-Responses at Pretest and Follow-Up: Number per Minute (Percentages of All Responses in Parentheses)

RL-response	14-hr		28-hr		42-hr	
	Pre	F-up	Pre	F-up	Pre	F-up
RF	0.8 (12)	1.1 (17)	0.7 (11)	0.9 (17)	0.7 (9)	0.9 (17)
RE	0.1 (2)	0.3 (4)	0.3 (5)	0.7 (13)	0.2 (3)	0.7 (13)
QF	1.8 (28)	1.4 (21)	1.7 (27)	1.1 (20)	1.8 (25)	0.4 (8)
QE	0.4 (6)	0.2 (4)	0.5 (8)	0.4 (8)	0.6 (8)	0.2 (4)

Note. RF = Reflecting fact, RE = Reflecting emotion, QF = Question on fact, QE = Question on emotion

The differences in RL between the three training groups were nonsignificant. However, clients who met students with the 42-hr RL training used nearly twice as many responses on emotion, as the other clients used.

Study 3

The idea for Study 3 was to let a larger sample of raters evaluate the students, and to explore if there was a relationship between the raters' evaluations and their own social skills. Study 1 and 2 had shown that the trainees had learnt nondirective counseling skills, but were still not generally more appreciated by the confederates — a result that contradicted the Rogerian assumption. One explanation could be that they had not learnt the skills well enough, another that people with different characteristics would not appreciate the skills equally.

Among numerous individual characteristics of clients, social skills seemed like a good starting point. Professional helpers working with challenged clients may find that directive responses, such as instructions combined with rewards, is an option to letting clients decide themselves what steps to take. There is also some evidence that client functioning can be related to the responses of the counselor (e.g., Hill et al., 1988). Hence, it was thought that especially raters with good social skills would be fond of a nondirective counseling style.

The external raters were 43 vocational counselors and psychology undergraduates without any CST. They watched and rated six conversations from each of the three training groups (14-, 28-, and 42-hr). They evaluated the students on the Positive evaluation index, and on an open-ended measure that described the students. Finally, the raters filled out the SSI. The nearly

2,800 open-ended descriptions were categorized as positive, negative, or neutral. Then they were quantified.

For all measures, a two-way mixed ANOVA was applied to describe the differences between the ratings given to the students in the three training groups, and how the ratings interacted with the raters' social skills (low, high). The results obtained from the Positive evaluation index, and the positive open-ended descriptions were the same with respect to the descriptive and inferential statistics. Therefore, only the results for the positive open-ended descriptions are presented (see Figure 1).

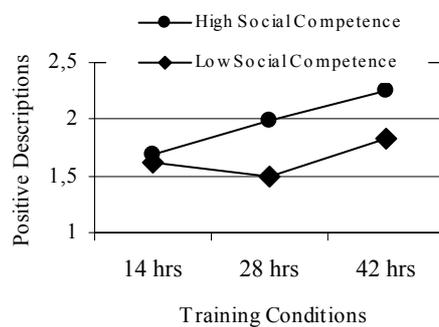


Figure 1. Mean number of positive descriptions given to the students in the three training conditions by raters with high and low social skills

There was a significant main effect of length of training for the positive descriptions. The students who had trained RL for 42 hrs received more positive descriptions than the students in the other groups. Moreover, there was a significant effect of length of training and the raters' social skills. Figure 1 illustrates that raters with high social skills gave consistently more positive evaluations to the students who had trained for a longer time, whereas raters with low social skills gave more inconsistent ratings.

Paper III

There are two main drawbacks with using a CST to evaluate communication skills: it takes long time, and it is difficult to decide whether it is the skills themselves, or some confounding variable, that leads to better ratings (Quinn, Sherman, Sheldon, Quinn, & Harchik, 1992). The second drawback was actualized in Paper I where the students in the 42-hr training group remembered the information given by the confederates better (Study 1), and

the clients of the students in the 42-hr training group talked more about their emotions (Study 2), although the use of RL changed equally among the three groups. So it seems that training led to other effects besides change in RL, and that these effects confounded the evaluation of the skills themselves.

The purpose of Paper III was to evaluate subskills of nondirective counseling, instead of evaluating effects of training. Thus, the conditions were not different trainings, nor effects of these. Instead, two of the verbal responses advocated in nondirective counseling, RL and open-ended questions, were manipulated in a factorial experiment. The general hypothesis was that nondirective responses (i.e., RL and open-ended questions) would be evaluated more positively than directive responses (i.e., less RL and closed-ended questions). Furthermore, it was expected that the raters' social skills and cognitive ability would moderate the evaluation, so that raters with higher social skills and cognitive ability would evaluate RL and open-ended questions more positively than others. Support for these hypotheses was obtained from the present Paper II, and from studies showing that cognitive ability is related to preferred levels of directivity (Hill, 1992) and structure (Holloway & Wampold, 1986).

Study 1

The starting point for Study 1 was the transcript of conversations from the pretest of Study 1 in Paper II. It was borne in mind that they were carried out by counselors in training, that is, would represent a normal achievement of people participating in CST. Four roles were chosen that were average with respect to RL and open- and closed-ended questions.

The transcripts were manipulated by adding or subtracting RL-responses and open- and closed-ended questions. For each transcript, two responses from the categories Reflecting fact, Reflecting emotion, and open- and closed-ended Question on fact and Question on emotion were either added or subtracted. This resulted in four conditions: Much RL and many open-ended questions, Little RL and many open-ended questions, Much RL and few open-ended questions, and Little RL and few open-ended questions.

I aimed at creating a close-to-authentic atmosphere in the evaluation. Therefore, the counselors in the subsequent taping of the manipulated transcripts were four clinical psychologists of different age and sex, who met with four fictive clients. The eight participants role-played in a total of 32 conversations that were audiotaped.

The 52 participants in the evaluation were high school students, psychology undergraduates, nonpsychology university students, and people selected from a telephone catalogue. They listened to the audiotaped conversations and gave open-ended descriptions of the psychologist.

Afterwards, they rated the conversations using the Positive evaluation index, and the WAI. Thus, three dependent variables were used in the evaluations. Finally, they self-reported their social skills using the SSI.

The positive and negative descriptions were combined into an Open-ended index consisting of the number of positive subtracted by the number of negative descriptions. The differences on the dependent measures between the four conditions were analyzed using factorial repeated-measures ANOVA. The typical result (e.g., the Positive evaluation index) is presented in Figure 2.

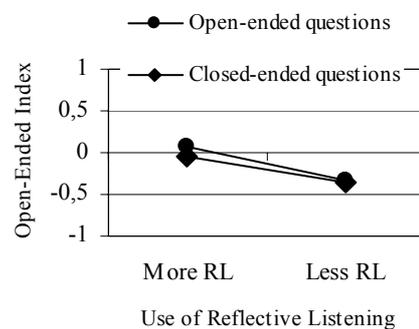


Figure 2. Mean scores on the Positive evaluation index given to the psychologists in the four conditions

In summary, the counselors who used more as opposed to less RL received more positive evaluations of their conversation skills and working alliance. However, the results showed that counselors who used more as opposed to less open-ended questions also received less positive evaluations. When the SSI was included as a continuous predictor variable, two of three dependent variables indicated that the judges' social skills affected how they evaluated conversation skills. Raters with higher social skills judged RL and open-ended questions more favorably than did the raters with lower social skills (for details, see Paper III).

Study 2

Study 2 utilized the transcripts of Study 1 without any audiotaping. I hoped to replicate the moderating effects of the evaluators' social skills, and incorporated also a measure of cognitive ability as a second variable that might moderate an evaluation. Adding another moderating variable would seemingly necessitate a larger sample of participants. That is why Study 2

utilized transcripts instead of audiotaped conversations. The idea thereby was to decrease the extraneous variance of prosodic features.

The 63 participants were psychology undergraduates, nonpsychology university students, and people selected from a telephone catalogue. The transcripts were presented as written dialogue, and the judges evaluated them in a manner similar to Study 1. After the evaluations, they completed one additional (compared with Study 1) self-report measure of cognitive ability, the SPM. The differences on the dependent measures between the four conditions were analyzed using factorial repeated-measures ANOVA. The typical result is presented in Figure 3.

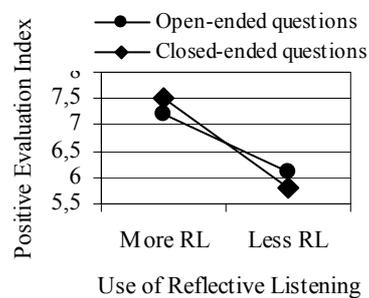


Figure 3. Mean scores on the Positive evaluation index given to the psychologists in the four conditions

Taken together, the counselors who used more as opposed to less RL received more positive evaluations of their conversation skills and working alliance. However, the results showed that counselors who used more open-ended questions received less positive evaluations. When the SSI and SPM were included as continuous predictor variables, one dependent variable, the WAI, indicated that the judges' social skills and cognitive ability affected how they evaluated conversation skills.

Raters with high social skills gave less positive ratings to less RL and fewer open-ended questions, while raters with low social skills gave about equal ratings. Moreover, all three dependent variables indicated that the judges' social skills and cognitive ability affected how they evaluated conversation skills (for details, see Paper III).

Effects of sex — a supplementary analysis

The effects of sex on role behaviors, including verbal behavior, have been extensively studied (Sells, Goodyear, Lichtenberg, & Polkinghorne, 1997; Steckler & Rosenthal, 1985). Early studies reported that men (are encouraged to) send more high-power messages, whereas the conversations of women address feelings and have an emphasis on relational concerns. This has been epitomized as women's conversations being more relationship-oriented, and men's being more task-oriented.

At the same time, research has shown that people's conception of female and male stereotypes is more divergent than the actual differences between the sexes justify (Eagly & Carli, 1981). And I have during eight years' of CST never felt that there were any sex differences between the participants with respect to nondirective counseling skills.

The purpose of this supplementary analysis was to investigate sex differences in the data presented in Paper I and II, where participants of both sexes were assessed before and after training. It was hypothesized that if there were differences between women and men before training, then these differences would become smaller. The rationale for this hypothesis is that training of specific conversation skills (e.g., nondirective counseling) make participants more homogeneous; those who do not use RL before training (whether it might be women or men) learn to.

Paper I

Study 1 had quantified nondirective and directive verbal responses. In a supplementary analysis of the effects of sex, a MANCOVA was applied to the posttest z scores of the employees' core RL-responses (Reflecting fact and emotion, and Question on fact and emotion) and directive responses (Advice, Information, and Offering help). Corresponding pre-training data were used as covariates. Considering the small number of women ($n = 14$) and men ($n = 9$) in the comparison, there was some difference between them regarding the RL-responses (Wilk's $\Lambda = .48$, Rao's $R = 3.24$, $p = .05$), but not regarding the directive responses (Wilk's $\Lambda = .73$, Rao's $R = 1.73$, $p = .21$).

The core RL responses were analyzed further with multiple unrelated t tests. There was one difference before training: the women used more Reflecting emotion ($z = .26$) than the men did ($z = -.50$), $t(30) = 2.17$, $p = .04$. This difference was smaller after training, however ($z = .28$ and $z = -.43$ for women and men, respectively, $t(21) = 1.73$, $p = .10$). None of the other

separate *t* tests on the responses before or after training were significant, *ts* between .02 and .49, *ps* between .10 and .93.

Paper II

In Study 1, nondirective and directive responses had been quantified. In a supplementary analysis of the effects of sex, two 2×2 (Sex: female and male; Time: pre- and posttest) repeated measures MANOVAs were conducted for the students' core RL-responses (Reflecting fact and emotion, and Question on fact and emotion), and directive responses (Advice and Interpretation). There were, however, no statistically significant differences between women and men, neither regarding the core RL-responses (Wilk's $\Lambda = .89$, Rao's $R = 0.57$, $p = .69$), nor the directive responses (Wilk's $\Lambda = .86$, Rao's $R = 1.57$, $p = .23$).

In Study 2, core RL- and directive responses in the follow-up had been quantified. In a supplementary analysis, the *z* scores were analyzed using 2×2 (Sex: female and male; Time: pretest and follow-up) MANCOVAs with the pre-training data as covariates. Considering again the small number of women ($n = 17$) and men ($n = 2$) in the comparison, there was a statistically significant difference between them regarding the RL-responses (Wilk's $\Lambda = .56$, Rao's $R = 2.76$, $p = .07$), but not for the directive responses (Wilk's $\Lambda = .97$, Rao's $R = 0.31$, $p = .74$).

Here too, the core RL responses were analyzed further in four separate ANCOVAs. None of them showed any significant differences between women and men, however (*F*s between .18 and 3.08, *ps* between .09 and .67).

Summary

To recapitulate the analysis of the effects of the trained participants' sex, the results supported some differences between women and men. More precisely, women seemed to use more Reflecting emotion than men did before training. This finding was, however, not consistent through Paper I and II.

DISCUSSION

In every piece of research there are strengths and weaknesses. The studies presented in this thesis belong to a category called analogue research. They have been conducted in the ‘laboratory’ or in other controlled environments that attempt to mimic real life. The designs have aimed at controlling confounding variables and testing causal relationships, although it has not been without difficulties. Like all research, the present studies suffer from threats to both internal and external validity. The purpose of this final section is to discuss the main findings, their shortcomings, and possible implications of the results.

Main findings and implications

The results presented above suggest that RL is a communication skill that can be learnt in a relatively short time (i.e., 14 to 42 hours). There was also support for the view that those who train RL know how to use the communication skill outside the training context, and at follow-up (i.e., after 5 to 12 months). The results suggest that those who have trained RL for a longer time remember more information from conversations than those who have trained for a shorter time. They did, however, not support the view that those who have trained RL are considered better than those who have not. Further investigations proposed that reflecting fact and emotion are generally considered as positive, while open-ended questions are not, with the addition that the judge’s own social skill and cognitive ability may moderate the evaluation. It was predicted that persons with more developed social skills and cognitive ability evaluate a longer training, and more RL, as better than others do. Taken together, the results support that people can learn specific communication skills, even though the effects of using RL and open-ended questions still need to be specified.

Even though the use of RL and questions during the training had been overcorrected (exaggeratedly encouraged), the participants lowered their use of the skills in the assessments. The trained participants used Reflecting fact in around 10% to 20% of their responses (approximately once to twice a

minute), and Reflecting emotion at around 5% to 15% (0.3 to 1 times a minute). I recall that during training these responses were used approximately twice as much, meaning that some cognitive process within the trained person regulates the use of these responses, unless they are forgotten. It might reflect an individual preference, but it might also be determined by external factors according to the contextual approach described earlier.

More training is needed

I propose that more CST is generally needed among helping professionals. It is justified by the opinion that training usually has little space in education today. For example, the psychologist education at Uppsala University is currently 5 years long, consisting of a total of 8,000 'work hours' (40 hours each 20-week semester). It is difficult to estimate how much of this time is made full use of, but if the whole is considered, then the longest time length of CST in the present studies (i.e., 42 hours) corresponds to a proportion of 1/2 thousandth! And although the students actually receive some additional CST later in their education, my estimate of 60 hours during the 5-year program is not much.

There are some obvious counter-arguments to increasing CST. One is the demands that it makes on resources; something else may have to be put aside. Another objection is that it is difficult to know what communication skills should be taught, as the earlier review and the present studies imply. This later objection is nevertheless justified for much of the content in education today. To my present knowledge it has not been demonstrated how psychological theories aid the practice of professional helping, either. The pros and cons in this decision may eventually become clearer through future research.

Communication skills vary

The last two studies (Paper III) in the present thesis focused on what effects directive and nondirective counseling have on liking and the working alliance. Earlier studies have demonstrated that counselors can be more or less directive and focused on problem solving. For example, Nagel, Hoffman, and Hill (1995) in their investigation about career counselors' verbal responses reported that they primarily used information (42%), direct guidance (18%), and closed question (9%). In contrast, Carl Rogers himself used more reflecting (42%), interpretation (16%), and less information (16%). It remains unanswered whether nondirective counselors are regarded

as more helpful or not. I suggest that it depends on the context, the problem, and the help-seeker's individual characteristics. Implications for training professional helpers might, however, be that they must learn basic skills of counseling before they focus on different contents (Nagel, Hoffman, & Hill, 1995).

The contextual approach suggests that it is difficult to find a generally acceptable model for conversations such as nondirective listening. Instead, it is proposed that a good communicator masters many communication skills, and varies them according to what is appropriate. For example, RL might not work if a help-seeker is very talkative. Instead, the listener might need to emphasize turn taking, or pose closed-ended questions so as to discourage story telling, or perhaps even tell the client to be quiet. It is thus proposed here that education of professional helpers should include a variety of tangible communication skills. CST should include nondirective listening, but should also contrast between other directive and nondirective communication skills, teach students how to increase other people's motivation, give them advice, interpretations, perhaps rehearse the skill of 'saying no', or how to confront lying. It has earlier been proposed that if components of communication skills can be operationalized — and often they can (Riggio, 1986) — then they can also be trained.

Teaching a *smorgasbord* of communication skills suggests that a skill such as RL should never be declared a paragon (Hargie & Tourish, 1994). As the present research proposes, reflecting fact and emotion is generally considered as positive, while open-ended questions are not, and cognitively more developed persons may prefer reflecting and open-ended questions to a larger extent than others do. On the other hand, already in the 1940s, Grummon and Gordon (1948) reported that the largest group of interrupted cases at the Counseling Center of the University of Chicago were individuals who felt that they needed direct advice from the counselor. If interrupted cases were individuals who wanted direct advice, then nondirective counseling is perhaps not suited for problems where such are needed. Then how do we know? One obvious answer is "Ask the client". Another option is to provide the helping professional with knowledge about such preferences through training, or as rules-of-thumb in books or manuals, if possible. It is the knowledge of such mediating or moderating effects that, according to the contextual approach, distinguishes between a helper with better or worse communication skills.

Methodological issues

Hathaway (1948/2000) judged psychotherapy outcome research harshly, although more than half a century ago. He claimed that the client's liking and trust of the helper is a common feature to all types of counseling or psychotherapy. Most clients feel pressure to present an improved condition after therapy. Therefore, no helper should assume that she or he could control these effects simply by an accepting manner. Instead, the helper should assume that nearly all forms of attention given to a patient's problems could result in improvement.

Hathaway was critical to early attempts to evaluate psychotherapy, including nondirective counseling. He pointed out that evaluations often failed to ask for the opinions of those who depended most on psychotherapy practice, namely the clients. Instead, for some published research it was customary that clinical observers (themselves experts of practice and theory) decided whether an intervention was effective or not. In the present studies, the trainees' opinions before and after training were not included as a dependent measure, which makes Hathaway's remark less valid. Nevertheless, other aspects of the present investigations merit objection.

Social validity

Experimental psychology is often — at least by its practitioners — regarded as a natural-scientific pursuit, rife with objective measurements and far away from introspection (Wolf, 1978). The implication is that research is *seemingly* without values, or that its values are plausibly defined. However, it has been stated, and history shows plenty of examples, that research is often associated with value statements, and a need to validate certain goals, procedures, and results. These three aspects together have been called social validity. Wolf articulated that although theories and experience can give guidance, it is always the target populations of research (participants, target groups, and financiers) that decide whether it is valuable or not.

The issue of social validity was raised during the pessimism of the 1970s concerning psychotherapy process research, which was complex and seldom programmatic, researchers lacked computer capacity, and the results were rarely illuminating. Holsti (1969) summarized the lack of findings as "If you can't count it, it doesn't count; if you can count it, that ain't it" (p. 112). Since then, the field has been characterized by a split with respect to the operationalization of theoretical constructs: easy-to-quantify discrete approaches (such as those used in the present thesis) vs. more abstract complex variables. Eventually, investigators seem to have concluded that

variables should to be studied at the middle level of abstraction (Marmar, 1990).

Social validation is a stepwise procedure that is undertaken to create a new intervention, or to increase the external validity of an intervention that already exists (van Houten, 1979; Hughes et al., 1998; Minkin et al., 1976). (In some cases, intervention means a treatment, but the term can be substituted by manipulation here.) The first step is to decide how the phenomenon is to be sampled and represented, which usually means defining ideal, normative, and deficient levels of target behaviors. In the present papers, the normative levels of RL have been the proportion of such responses before training. For example, the norm without a manipulation (i.e., without CST) is the amount of nondirective responses used by untrained psychology students.

The next step in social validation is the selection of judges who rate behavior. There is a consensus that it is valuable to include many sources in ratings, for example patients, therapists, and independent clinical judges (Marmar, 1990). In the present research, the judges came from relatively undiversified populations, such as psychology undergraduates and nonpsychology university students. This is an obvious limitation in the present studies.

When judges have been identified, it is decided which qualities are to be evaluated. Usually previous research or experts are consulted. Then follows scale construction, and typing of instructions, unless the research can rely on scales whose validity has been previously investigated. One validity issue is the number of items in a scale. It is dangerous to use single-item measures, or few-item measures, as was pointed out to me regarding the Positive Evaluation Index² in the review of Paper I. The internal consistency of this three-item measure was 'satisfactory' .91 or more, but the figure can be interpreted in several ways. One interpretation is that the measure is homogeneous, which is unlikely considering items such as competence and compassion. Another interpretation is that these three items *do not* capture the essence of competence, confidence, and compassion, respectively; instead they indicate some momentary response bias, an undefined impression that we may have of someone, positive or negative.

Looking back, the three items denoted competence, confidence, and compassion were originally chosen for practical reasons. When the female confederate in Paper I had her conversations on the telephone with the insurance claim adjusters, she followed a strict time schedule where she received a call from a new employee as soon as she had hung up with the former. In that chosen procedure, she gave rapid ratings (on the three items) so that she had time to prepare for the next conversation. When the

² The average of the three items on the listener's competence, confidence, and compassion.

drawbacks of this were revealed, the procedure was improved by including the validated measure WAI.

The goal of social validation is to find information about the participants' behavior that is socially valid, that is, discriminates between groups, or effects of time (such as before and after an intervention). The goal is also to find information about the effects on some external criterion variable. For example, in the present research more and less directive counseling were contrasted, and it was hypothesized that nondirective counseling would strengthen the therapeutic relationship. Furthermore, social validation implies that if the data is divergent, the manipulation must be modified in an ongoing process. This is why the variables social skills and cognitive ability were included in Paper II and III, to investigate if individual differences on these variables could explain the inconsistent ratings of RL.

Difficulties in social validation

The depth of the sought-for effects in an intervention vary: from the proximal (for example communicating clearly), to the intermediate (meeting friends more often), and to the distal (having more friends) (Fawcett, 1991). Likewise, there is seemingly a need to measure change with several methods (e.g., self-report, behavioral) and from several perspectives (e.g., counselor, client, significant other such as a relative). An often-cited question by Paul (1967, p. 111) is "What treatment, by whom, is most effective for an individual with a specific problem, and under which circumstances?"

In theory, there is always an optimal range of responding for a behavior (van Houten, 1979), and this range can be established through replication and iteration. A social validation approach may present difficulties, however. One is that if a behavior is overcorrected, it is not plausible that the contingencies in the environment adjust it to an optimal level. In the present research, it seems that overcorrection of verbal responses (e.g., reflecting fact and emotion were 'overdone' in training) did not lead to an exaggerated use of these responses in other contexts. Another difficulty with social validation is that it is a time-consuming process. Target behaviors interact with each other, making it difficult to estimate optimal response levels for them individually. Paper III suggests that reflecting fact and emotion are considered positive. But it is another question whether such a finding would generalize if levels of other verbal responses had been manipulated as well.

Both theory and practice have fuelled the social validation carried out in the present research. The most influential theory with respect to the skills themselves has been the client-centered approach, and the skills have been trained according to social learning theory that puts emphasis on vicarious learning, modeling, and reinforcement. Needless to say, when a training

program is put together it evolves *peu à peu* due to the participants' experiences.

Defending the randomized experiment

The randomized experiment has for a long time been a paragon of psychological research (Zajonc, 1989), even though in some settings it may be impractical (Sackett & Mullen, 1993; Tannenbaum & Woods, 1992). Another design, the so called nonequivalent control group design, is often used in these other instances, such as the present Paper I. The major difference between them is the probability of the experiment and control groups being different before a manipulation. The true experiment seeks to eliminate any prior differences by randomization, the nonequivalent design does not.

The distinction between the two designs is not crystal clear. For example, sometimes a procedure is no longer regarded as an experiment if the attrition from any group is high enough. This is the main objection that can be raised toward Paper I and II. The attrition in the training groups was between 0% and 40% ($M = 25\%$), which is undoubtedly a whole lot. Considering the small number of participants in the first place, between 11 and 17 in each condition, there is really not so much one can say to defend the external validity of the results. In the present studies, the differences between dropouts and remaining participants were described and compared with ANOVAs. But for such small groups, differences between them would have to be exceptionally large in order to be statistically significant.

A meta-analysis by Shadish and Ragsdale (1996) investigated differences between effect sizes of interventions in family therapy. Their sample of randomized experiments generally had larger effect sizes than nonequivalent designs. Moreover, nonequivalent designs generally resulted in a priori differences between the groups. They recommended that randomized experiments be used when possible, with which I agree. If randomization is not possible, the bias could be partly corrected by including pretest confounding variables as covariates. This is the solution used in Paper I.

Evaluating CST is often equated with measuring how participants' performance change. If this is to be done rigorously, it requires experimental design. Meanwhile, for many organizations it is impractical to create experiments. This means that if the standards of evaluation are set very high (requiring pretest, control group, and random assignment) then evaluation might not even be considered. Indeed, it has been argued that a less-than-best evaluation is better than no evaluation at all, at least if the investigator is aware of the limitations of interpretation (Sackett & Mullen, 1993).

Content analysis and quantification

Content analysis is a research method that draws conclusions from verbal material such as oral presentations or written text. The method has traditionally not been widespread among investigators of psychology. It has been said that it is biased, fuzzy, difficult, and painstaking, although its use is increasing (Lee & Peterson, 1997).

The literature on content analysis does not state exactly how it is done. The general advice is that the researcher should choose a mode of analysis that suits the hypotheses under investigation, and that she or he must be aware of own biases that can contaminate the findings. The present research has defined a verbal response so that it is always delineated by verbal behavior of another person, while other researchers (e.g., Ralph & Lee, 1994; Ralph & Thorne, 1993) have regarded verbal responses with an intermittent pause as two responses.

With large amounts of data, it is possible for the investigator to choose to analyze parts of it instead of the whole (e.g., the present investigations' 5-min intervals of conversations), provided that the investigator makes an effort toward an unbiased selection procedure. Friedlander, Ellis, Siegel, Raymond, and Haase (1988) investigated what fraction of an interview best represents the entire interview. They sampled a total of 31 counseling sessions, and concluded that the longest segment is generally the most representative. However, for individual interviews, even the longest segment is representative only 25% to 72% of the time, while for aggregated interviews even short segments are representative. This result held regardless of what segment of the counseling session was extracted. Their recommendation was that, if doing a case study the investigator should examine whole interviews. If comparing groups, the investigator may choose even small segments of counseling sessions, as has been done in the present investigations.

Another important question is the level of detail at which content analysis should be carried out. The present investigations have categorized responses out of transcripts. An objection might be that I have left out important prosodic features of the responses that influence their quality. Furthermore, it is considered important that the people who code the material are unaware of the research hypotheses, and that their coding is preceded by adaptation of rules for coding, and rules for reliability calculation (Lee & Peterson, 1997). It was not so in the present investigations; therefore it is possible that confounding responses in the transcripts have led the coder (i.e., me) to, for example, categorize a response as RL even if it was not. In Paper I and II, randomly chosen conversations were coded by another person as well. It is

more reliable than a single-person coding, but undoubtedly does not control for a theoretical bias if the other person also is fluent in Rogerian theory.

The quantification of the verbal responses in the present papers has been presented using the number of responses in the transcribed samples. The quantification in the present thesis is presented as the number of responses during a 1-minute period (e.g., see Table 1). Yet, someone acquainted with this area of research knows that the most common way to present the use of responses is by percentages (e.g., Nagel, Hoffman, & Hill, 1995). The reason why I describe responses during 1 minute is that such a presentation works as a rule-of-thumb for how many responses should be used, more than percentages would. It is easier to recommend reflecting “twice a minute” than to recommend a percentage that is dependent on other responses and speed of talking.

Rating procedure

Rater selection is essential to counseling process research. Some theoretical constructs, such as the working alliance, cannot be quantified without using human judgment, and one must consider that characteristics of judges can affect the reliability and validity of the data. In fact, the best measure of a construct may sometimes be the variance in scores obtained from different groups of raters. And there is variance within groups as well; it has been known for long that raters’ interpretations of a counseling session can be strikingly different (Moras & Hill, 1991). Other features that may covary in the rating process is the investigator’s influence, possible training, and type of judgment that is required.

Thought sampling is a procedure where participants are asked to write a narrative description of their experience, perhaps prompted by some signal. Its popularity for quantifying thinking is increasing (Hurlburt, 1997). Still, it has been said that humans are not skilled at describing their thinking retrospectively; they make substantial errors when they are asked to look back over some time. This is why the participants in the present research gave their ratings as soon as possible after having listened to audiotapes. In general, ratings made from transcripts yield the highest interrater reliabilities, followed by videotapes and audiotapes (McDaniel, Stiles, & McGaughey, 1981; Weiss, Marmar, & Horowitz, 1988; cited in Marmar, 1990), but this issue has not been investigated further in the present thesis.

So humans are imperfect measuring devices of their own cognition. The discrepancy between cognition assumed to underlie behavior and self-reports of that cognition is called distortion (Nelson, 1996). This metacognitive idea suggests that verbal reports on people’s cognition should include an independent means of assessing the validity of the report. Examples of such

measures could be individual characteristics such as cognitive ability. Other characteristics could be asked for in the procedure, for example asking the participant to identify the aspects of a conversation that she or he is paying attention to.

The present research has included some individual characteristics of the participants. Nevertheless, it would have been possible to ask for more. A review by Moras and Hill (1991) pointed out that the potential impact of individual characteristics of raters on ratings is rarely mentioned. Their review concluded that for low-inference judgments, more information is perhaps not needed. For high-inference judgments, comparable to those used in the present studies, rater selection merits closer attention. The present studies have provided sex, age, and basic demographic data on participants. On some participants in Paper II and III, social skills and cognitive ability were measured as well.

For social psychologists, the validity of verbal reports is important, because they often ask their participants why they behaved, chose, or evaluated as they did. Participants readily answer such questions, but the accuracy of their answers is generally unreliable (Nisbett & DeCamp Wilson, 1977). Instead, they base their answers on own ideas about causal connections. In the present research, this implies that participants who judge conversations may develop ideas about what makes them give low or high ratings, which then may influence their ratings.

Applications

People who have gone through a normal cognitive development acquire communication skills readily through observational learning and interactions in the natural milieu. For others, and for those who need extraordinarily good skills (e.g., professional helpers), explicit instruction and reinforcement may be needed. This in turn raises the question of whether trained communication skills generalize to real-life situations (Griffiths, Feldman, & Tough, 1997). One way to maximize generalization is to provide multiple exemplars, fading reinforcement, and problem-solving training. All these aspects were provided in the CST in the present thesis.

I feel that a focus on CST is constructive. Emphasizing the counselor's verbal behavior means to focus on something that can be changed. It is a pragmatically diametrical view to emphasizing the individual characteristics of a professional helper.

Psychological interventions

Many popular and empirically validated psychological interventions have their origins in nondirective counseling. The best known is perhaps motivational interviewing, but there are others too, such as relationship-centered counseling. There is currently a controversy about whether positive change in counseling or psychotherapy is due to specific ingredients of the treatments, or to factors that are common in all therapies. Specific ingredients are, for example, exposure and cognitive restructuring. Instead, models that articulate common factors focus on the working alliance, context, and the client's belief in the treatment.

Advocators of common factors in treatment suggest that there is no evidence that specific ingredients are responsible for positive outcomes of counseling and psychotherapy. Instead, they claim that training should focus on other things such as interviewing skills, how to establish a good working alliance, and basic facilitative conditions in the client's life (Ahn & Wampold, 2001). Bordin (1979) operationalized the alliance as something that does not necessarily lead to good outcomes, but makes it possible for the client to accept and follow the treatment. In line with this idea, I suggest that the relation between therapeutic alliance and outcomes is, in fact, mediated by the counselor's and the client's behavior. The chain might look as follows: counselor behavior \Rightarrow working alliance \Rightarrow client behavior \Rightarrow outcomes. This is the basic idea behind motivational interviewing.

Motivational interviewing

Motivational interviewing (MI) is a technique mainly used in treatment of alcohol- and drug addiction, and other health-related problems. The earliest individual reports on MI date back to the 1980s (Miller, 1983), but it has today become increasingly popular among cognitive-behaviorally oriented counselors and therapists.

The term MI suggests that it is used to enhance a client's motivation to change her or his lifestyle. In treatments of addictions, therapists have traditionally met much resistance from clients who may have started treatment by decision of a court of justice, or been pushed by relatives. These clients might not themselves feel the need to change. Instead, their problem is that others are annoyed by their behavior. It may also be that the client's bad habits have not yet become an issue, but pose a risk for future health problems. Thus, the motivational interviewer sets out to reflect and clarify the client's present motivation for change, engaging the client in a discussion of pros and cons of changing behavior.

MI is a Rogerian approach. It sees the client's motivation not as a stable trait, but as something that evolves in the interaction between the client and the professional helper. Moreover, it suggests that the helper must empathize with the client, seeing behavior as it is, and understand what factors make it difficult for the client to deal with life. The MI professional does not condemn or preach; instead she or he retains curiosity for the client's experiences, collecting information that helps the client to make a decision.

Miller and Rollnick (2002) described the basic principles of MI: showing empathy with the client, pointing at the discrepancy between the current state and the goal, not engaging in argumentation, but instead giving in to resistance, and supporting self-efficacy. These principles operationalize into a discussion with many open-ended questions, clarifications, RL, and refraining from giving advice. These principles are also an essential part of client-centered therapy (Rogers, 1957).

There is, however, a contradiction-in-terms in MI. One way to define it is "a client-centered, directive method for enhancing intrinsic motivation to change by exploring and resolving ambivalence" (retrieved from <http://www.motivationalinterview.com>, on April 26, 2004). The definition says that it is client-centered, which in the Rogerian sense would imply nondirective. However, the clinician always has the goal that the client ultimately becomes free of drugs. Looking at the elements of MI (Bien, Miller, & Tonigan, 1993), it has both similarities and dissimilarities to the nondirective counseling of Carl Rogers. The similarities are feedback, emphasis on client's responsibility, communicating empathy, and increasing self-efficacy. The dissimilarities are giving direct advice, and providing alternatives for behavior change.

The similarities seem to outweigh the differences, especially if MI is compared with a traditional drug dependence treatment where the client is confronted (see Miller & Rollnick, 2002). And as MI has yielded promising results with respect to alcoholism, its collaborative approach is now *à la mode* in some other treatments as well. One example is treatment of eating disorders, and it has been reported that collaborative interventions were rated as more acceptable and more likely to produce positive outcomes than directive interventions. Moreover, clients who were least ready for change rated directive interventions as less acceptable than did clients who were more ready (Geller, Brown, Zaitsoff, Goodrich, & Hastings, 2003). This may mean that a thorough training in nondirective counseling may be a well-invested preparation for counselors working with addiction. It needs to be taken into consideration, however, that the results of the present studies imply that preference for more or less directive counseling style may be influenced by the client's cognitive functioning. This may in turn be impaired due to addiction.

Relationship-centered counseling

Kelly presented (1997) an intervention that integrates both humanistic and technical perspectives, which he called relationship-centered counseling. He recapitulated the polemics between the traditionally humanistic and relationship-oriented interventions, and the behavioral and task-oriented interventions. Moreover, he argued that the essence of being human should be the foundation for all therapies. By this integrative approach, Kelly argued that there is traditionally a reluctance from the humanist side to integrate technical components into counseling, as well as a denial of the significance of the relationship from the cognitive-behavioral side.

The relevance of Kelly's statement for the present thesis is that nondirective counseling, too, can be regarded either as a relationship- (e.g., the counselor is not trying to rule anyone's life), or a technical (e.g., the counselors 'uses' RL) issue. Kelly (e.g., 1974) has done research in the nondirective tradition of Rogers and Gordon, and as a consequence suggested that the basic principle of counseling is that being human is relational, and the therapeutic relationship is the core of counseling. Hence, techniques are secondary instrumental components of counseling, although they are legitimate, and the therapeutic relationship cannot exist without them. The primary purpose of counseling is to enhance clients' sense of freedom, purpose, and responsibility for life. In this way, the ideas in relationship-centered counseling echo what Carl Rogers and Thomas Gordon had articulated some decades earlier.

Kelly states that the relationship and technical dimensions in counseling are inseparable. This means that the counselor who expects the solution of human problems to be technical is wrong: there are problems in human life that are impossible to solve, regardless of the quality of counseling. Kelly's proposal also implies that there must never be pressure for quick results, and that a beginning counselor must be taught to shift between relational responsiveness and task-oriented probes (Kelly, 1997).

The relationship-centered counseling approach has not yet resulted in much research. Perhaps it will not; I have the impression that its purpose has rather been to clarify issues such as attitudes, goals, and responsibility, than to function as a platform for new interventions.

Call-center activity

Counseling skills are relevant in other contexts than counseling, too. Many helping professions share the task of providing comfort, support, and giving advice to people in need. It may be claimed that the shorter the intervention, the more important it is that it is generic, due to the helper's little insight

into, and control over, the change process of the help-seeker. Perhaps counseling skills should be considered especially important in situations where the professional helper and the help-seeker interact for only a short time, where there is no second interaction to correct the first. Such interactions can readily be found in commercial settings.

The success of employees at all levels depends on their ability to communicate effectively. Thus, when an organization invests in CST, they invest in human resources (Tannenbaum & Woods, 1992). The most compelling argument for training is that good communication skills might increase an organization's effectiveness. A questionnaire study by McEwen (1997) described the status of communication training in corporate settings in the U.S., identifying elements from programs and their participants. The results implied that managerial and interpersonal communication were the most frequently targeted areas. Of forty topics, team building, problem solving, and coaching skills were taught most frequently. Listening and counseling skills came 8th and 19th respectively, on the 40-topic list. This description is in line with Paper I, in which research took off when a company wanted to enhance its employees' communication skills, effectiveness, and customer satisfaction.

McEwen (1997) found that larger companies offered CST more frequently than smaller ones, and training was more common among service and management personnel than among production workers and administrative staff. Most companies spent up to 25% of their in-house education on CST. Generally, the impact of academic research was small: training was most frequently provided by in-house trainers and outside consultants. The least frequently used trainers were university professors. In summary, this means that corporate CST costs money, and that evaluation might either be lacking, or biased.

The results of studies investigating relationships between managerial communication skills and productivity have been inconsistent (Papa & Graham, 1991; Papa & Tracy, 1988). Nevertheless, CST is popular in corporate settings, although of varying sophistication. Some training is fairly randomly administered, and some assess managers' strengths and weaknesses prior to training. Such assessments can then be used to place the employee into one or another kind of training.

Future research

Counseling techniques seem to have significant effects. A problem with research on therapist technique, however, has been its focus on frequencies or proportion, without enough attention to quality of the operationalized skill

(Hill & Corbett, 1993). This is valid critique regarding the present investigations. Another problem has been that individual differences among therapists and clients have been ignored. The current thesis at least presents two individual characteristics among clients that moderate how they regard nondirective listening. Of course, there are other individual characteristics too that may be fruitful to examine, such as the help-seeker's global functioning, measures of anxiety, personality traits, and so on.

It was mentioned earlier that the studies in this text belong to a category called analogue research. It is evident throughout that it is not easy to simulate relevant events in counseling: confederates do not necessarily respond in the same manner as actual clients would. It has been suggested that future research should move from the aggregation of frequencies to more naturalistic case studies, sequential analysis, and the study of significant events (Hill & Corbett, 1993). Both exploratory and theory-driven research is needed, however. An interesting research question would be to further operationalize counselor behavior that enhances the therapeutic relationship.

The CST in the present studies has been short. I would happily design a longer and more 'sophisticated' CST for undergraduate psychology students. It seems that the students are ambitious, clever, and full of ideas, and that a short training does not fully make use of their learning potential. One issue regarding sophistication, but not investigated in the present thesis, is the modalities used in training. The CST presented here has been relatively simple and robust. Its purpose has been to teach basic nondirective counseling skills to psychology undergraduates, resulting in skills that are plainer than those investigated among master therapists (e.g., Jennings & Skovholt, 1999). Nevertheless, it would be possible to make training more fun, by video, discussion, and participation of to the students unknown people.

The aim of the present research has been to attain some applicable knowledge about RL. Needless to say, RL responses are only a small subset of all counseling responses that could be examined (see Hill, 1992). There are many challenging questions: what dimensions of counseling responses can be operationalized as more or less directive, and in what contexts are they considered helpful by clients? It would, for example, be interesting to investigate whether directivity can be operationalized pantheoretically, from within other theoretical orientations than the person-centered.

ACKNOWLEDGEMENTS

Äntligen! Finishing this thesis is an important event for me. As I see the terminus, I am very grateful for others' kind support. My former supervisor Hans-Olof Lisper is to be thanked for his deep knowledge of communication skills training, and for giving me interesting ideas for research. Without him, this work would have had no beginning. My current supervisor Bo Ekehammar is to be thanked for giving apposite feedback along the path that I already had chosen, a watchful eye, and some thorough advice on the investigation process. Without him, this work would have had no ending. I would also like to thank Staffan Sohlberg for his valuable comments on the manuscripts.

Time flies when you're having fun! Appreciation is expressed to family, friends, and beloved. And to those who have participated in the studies, be it as counselors or clients. Much of the fun has taken place at the Department of Psychology. Thanks to former and current PhD-students, especially my long-time companions Ingrid Olsson and Nazar Akrami, to those with whom I have shared room, the splendid library, and the support of the management, the teachers, and the technical-administrative staff.

Parts of the investigations have been subsidized by the Helge Ax:son Johnson Foundation, whose invaluable aid has been much appreciated.

Uppsala, September 2004

Erik Rautalinko

REFERENCES

- Ackerman, S. J., & Hilsenroth, M. J. (2003). A review of therapist characteristics and techniques positively impacting the therapeutic alliance. *Clinical Psychology Review, 23*, 1-33.
- Adams, L. (1989). *Be your best: Personal effectiveness in your life and your relationships*. New York: Putnam.
- Ahn, H., & Wampold, B. E. (2001). Where oh where are the specific ingredients? A meta-analysis of component studies in counseling and psychotherapy. *Journal of Counseling Psychology, 48*, 251-257.
- Argyle, M. (1975). *Bodily communication*. London: Methuen.
- Bien, T. H., Miller, W. R., & Tonigan, J. S. (1993). Brief interventions for alcohol problems: A review. *Addiction, 88*, 315-336.
- Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research, and Practice, 16*, 252-260.
- Bourhis, R. Y., Giles, H., & Lambert, W. E. (1975). Social consequences of accommodating one's style of speech: A cross-national investigation. *International Journal of the Sociology of Language, 6*, 55-72.
- Brown, P., & Fraser, C. (1979). Speech as a marker of situation. In K. R. Scherer, & H. Giles (Eds.), *Social markers in speech* (pp. 33-108). Cambridge: Cambridge University Press.
- Cedar, B., & Levant, R. F. (1990). A meta-analysis of the effects of Parent Effectiveness Training. *American Journal of Family Therapy, 18*, 373-384.
- Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.
- Chomsky, N. (1959). Verbal behavior. *Language, 35*, 26-58.
- Cissna, K. N., & Anderson, R. (1994). The 1957 Martin Buber-Carl Rogers dialogue, as dialogue. *Journal of Humanistic Psychology, 34*, 11-45.
- Cozby, P. C. (1973). Self-disclosure: A literature review. *Psychological Bulletin, 79*, 73-91.
- Darsten, M., Lisper, H.-O., & Sohlberg, S. (1979). Inläring av aktivt lyssnande och en jämförelse mellan blivande och erfarna psykologers empatiska förmåga. [Learning active listening and a comparison of empathic ability of future and experienced psychologists.] *Scandinavian Journal of Behaviour Therapy, 8*, 13-21.
- Davis, K. (1986). The process of problem (re)formulation in psychotherapy. *Sociology of Health and Illness, 8*, 44-74.
- Dembo, M. H., Sweitzer, M., & Lauritzen, P. (1985). An evaluation of group parent education: Behavioral, PET, and Adlerian programs. *Review of Educational Research, 55*, 155-200.
- Eagly, A. H., & Carli, L. (1981). Sex of researcher and sex-typed communications as determinants of sex differences in influenceability: A meta-analysis of social influence studies. *Psychological Bulletin, 90*, 1-20.
- Elliot, A. J. (1981). *Child language*. Cambridge: Cambridge University Press.

- Elliott, S. N., Sheridan, S. M., & Gresham, F. M. (1989). Assessing and treating social skills deficits: A case study for the scientist-practitioner. *Journal of School Psychology, 27*, 197-222.
- Fawcett, S. B. (1991). Social validity: A note on methodology. *Journal of Applied Behavior Analysis, 24*, 235-239.
- Frick, R. W. (1985). Communicating emotions: The role of prosodic features. *Psychological Bulletin, 97*, 412-429.
- Friedlander, M. L., Ellis, M. V., Siegel, S. M., Raymond, L., & Haase, R. (1988). Generalizing from segments to sessions: Should it be done? *Journal of Counseling Psychology, 35*, 243-250.
- Geller, J., Brown, K. E., Zaitsoff, S. L., Goodrich, S., & Hastings, F. (2003). Collaboration versus directive interventions in the treatment of eating disorders: Implications for care providers. *Professional Psychology: Research and Practice, 34*, 406-413.
- Gelso, C. J., & Fretz, B. R. (1992). *Counseling psychology*. New York: Harcourt Brace Jovanovich.
- Gentner, D., & Goldin-Meadow, S. (Eds.) (2003). *Language in mind: Advances in the study of language and thought*. Cambridge, MA: MIT Press.
- Gordon, T. (1950). What is gained by group participation? *Educational Leadership, 7*, 220-226.
- Gordon, T. (1970). *PET: Parent Effectiveness Training*. New York: Wyden.
- Gordon, T. (1977). *LET: Leader Effectiveness Training*. New York: Wyden.
- Graybill, D. (1986). A multiple-outcome evaluation of training parents in active listening. *Psychological Reports, 59*, 1171-1185.
- Griffiths, D., Feldman, M. A., & Tough, S. (1997). Programming generalization of social skills in adults with developmental disabilities: Effects on generalization and social validity. *Behavior Therapy, 28*, 253-269.
- Grummon, D. L., & Gordon, T. (1948). The counseling center at the University of Chicago. *American Psychologist, 3*, 166-171.
- Guralnik, D. B. (Ed.). (1984). *Webster's new world dictionary of the American language* (2nd college ed.). New York: Simon & Schuster.
- Hargie, O., & Tourish, D. (1994). Communication skills training: Management manipulation or personal development. *Human Relations, 47*, 1377-1389.
- Hathaway, S. R. (1948/2000). Some considerations relative to nondirective counseling as therapy. *Journal of Clinical Psychology, 56*, 853-859.
- Hayes, S. C., Follette, W. C., & Follette, V. M. (1995). Behavior therapy: A contextual approach. In A. S. Gurman & S. B. Messer (Eds.), *Essential psychotherapies: Theory and practice* (pp. 128-181). New York: Guilford.
- Hill, C. E. (1992). Research on therapist techniques in brief individual therapy: Implications for practitioners. *Counseling Psychologist, 4*, 689-711.
- Hill, C. E., & Corbett, M. M. (1993). A perspective on the history of process and outcome research in counseling psychology. *Journal of Counseling Psychology, 40*, 3-24.
- Hill, C. E., Helms, J. E., Tichenor, V., Spiegel, S. B., O'Grady, K. E., & Perry, E. S. (1988). Effects of therapist response modes in brief psychotherapy. *Journal of Counseling Psychology, 35*, 222-233.
- Holloway, E. L., & Wampold, B. E. (1986). Relation between conceptual level and counseling-related tasks: A meta-analysis. *Journal of Counseling Psychology, 33*, 310-319.
- Holsti, O. R. (1969). *Content analysis for the social sciences and humanities*. Reading, MA: Addison-Wesley.

- Horne, D. J. de L., Vatmanidis, P., & Careri, A. (1994). Preparing patients for invasive medical and surgical procedures 2: Using psychological interventions with adults and children. *Behavioral Medicine, 20*, 15-21.
- Horvath, A. O., & Greenberg, L. S. (1989). Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology, 36*, 223-233.
- Houten, R. van (1979). Social validation: The evolution of standards of competency for target behaviors. *Journal of Applied Behavior Analysis, 12*, 581-591.
- Hughes, C., Lorden, S. W., Scott, S. V., Hwang, B., Derer, K. R., Rodi, M. S., Pitkin, S. E., & Godshall, J. C. (1998). Identification and validation of critical conversational social skills. *Journal of Applied Behavior Analysis, 31*, 431-446.
- Hurlburt, R. T. (1997). Randomly sampling thinking in the natural environment. *Journal of Consulting and Clinical Psychology, 65*, 941-949.
- Jennings, L., & Skovholt, T. M. (1999). The cognitive, emotional, and relational characteristics of master therapists. *Journal of Counseling Psychology, 46*, 3-11.
- Kelly, E. W. Jr. (1974). Classroom discussions for personal growth and democratic problem-solving. *Elementary School Journal, 75*, 11-15.
- Kelly, E. W. Jr. (1997). Relationship-centered counseling: A humanistic model of integration. *Journal of Counseling and Development, 75*, 337-345.
- Knippen, J. T., & Green, T. B. (1994). How the manager can use active listening. *Public Personnel Management, 23*, 357-359.
- Krippendorff, K. (1980). *Content analysis: An introduction to its methodology*. Beverly Hills, CA: Sage.
- Lee, F., & Peterson, C. (1997). Content analysis of archival data. *Journal of Consulting and Clinical Psychology, 65*, 959-969.
- Lin, M.-J., Kelly, K. R., & Nelson, R. C. (1996). A comparative analysis of the interpersonal process in school-based counseling and consultation. *Journal of Counseling Psychology, 43*, 389-393.
- Lindh, G., & Lisper, H.-O. (1990). *Samtal för förändring*. [Conversation that changes.] Lund: Studentlitteratur.
- Lisper, H.-O., & Nilsson, K.-Å. (1982). En experimentell prövning av aktivt lyssnande som tröst för små barn som gråter. [An experimental investigation of active listening as comfort for small children who cry.] *Scandinavian Journal of Behaviour Therapy, 11*, 183-193.
- Lisper, H.-O., & Rautalinko, E. (1996). Effects of a six-hour training of active listening. *Scandinavian Journal of Behaviour Therapy, 25*, 117-125.
- Lock, A. (1980). *The guided reinvention of language*. London: Academic Press.
- Longman dictionary of contemporary English* (3rd ed.). (1995). Harlow, England: Longman.
- Mansfield, F. (1991). Supervised role-play in the teaching of the process of consultation. *Medical Education, 25*, 485-490.
- Marmar, C. R. (1990). Psychotherapy process research: Progress, dilemmas, and future directions. *Journal of Consulting and Clinical Psychology, 58*, 265-272.
- McEwen, T. (1997). Communication training in corporate settings: Lessons and opportunities for the academe. *Mid-American Journal of Business, 12*, 49-58.
- Metcalfe, J. (2000). Metamemory: Theory and data. In E. Tulving (Ed.), *The Oxford handbook of memory* (pp. 197-211). New York: Oxford University Press.
- Miller, W. R. (1983). Motivational interviewing with problem drinkers. *Behavioural Psychotherapy, 11*, 147-172.
- Miller, W. R., & Rollnick, S. (Eds.) (2002). *Motivational interviewing: Preparing people for change* (2nd ed.). London: Guilford.

- Minkin, N., Braukmann, C. J., Minkin, B. L., Timbers, G. D., Timbers, B. J., Fixsen, D. L., Phillips, E. L., & Wolf, M. M. (1976). The social validation and training of conversation skills. *Journal of Applied Behavior Analysis, 9*, 127-139.
- Moras, K., & Hill, C. E. (1991). Rater selection for psychotherapy process research: An evaluation of the state of the art. *Psychotherapy Research, 1*, 113-123.
- Nagel, D. P., Hoffman, M. A., & Hill, C. E. (1995). A comparison of verbal response modes used by master's-level career counselors and other helpers. *Journal of Counseling and Development, 74*, 101-104.
- Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist, 51*, 102-116.
- Ng, S. H., Bell, D., & Brooke, M. (1993). Gaining turns and achieving high influence in small conversational groups. *British Journal of Social Psychology, 32*, 265-75.
- Nisbett, R. E., & DeCamp Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*, 231-259.
- Nugent, W. R., & Halvorson, H. (1995). Testing the effects of active listening. *Research on Social Work Practice, 5*, 152-175.
- Odom, S. L., & McConnell, S. R. (1992). Improving social competence: An applied behavior analysis perspective. *Journal of Applied Behavior Analysis, 25*, 239-244.
- Papa, M. J., & Graham, E. E. (1991). The impact of diagnosing skill deficiencies and assessment-based communication training on managerial performance. *Communication Education, 40*, 368-384.
- Papa, M. J., & Tracy, K. (1988). Communicative indices of employee performance with new technology. *Communication Research, 15*, 524-544.
- Paul, G. (1967). Strategy in outcome research in psychotherapy. *Journal of Consulting Psychology, 31*, 109-118.
- Pedhazur, E. J., & Pedhazur Schmelkin, L. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Erlbaum.
- Phillips, B. (1999). Reformulating dispute narratives through active listening. *Mediation Quarterly, 17*, 161-180.
- Piaget, J. (1926). *The language and thought of the child*. New York: Harcourt, Brace.
- Quinn, J. M., Sherman, J. A., Sheldon, J. B., Quinn, L. M., & Harchik, A. E. (1992). Social validation of component behaviors of following instructions, accepting criticism, and negotiating. *Journal of Applied Behavior Analysis, 25*, 401-413.
- Ralph, A., & Lee, E. (1994). The effects of interpersonal verbal behaviour of varying participant role in initial dyadic conversations. *Scandinavian Journal of Behaviour Therapy, 23*, 155-175.
- Ralph, A., & Thorne, E. (1993). Defining and enhancing competent interview behaviour using the Verbal Interaction Analysis System. *Scandinavian Journal of Behaviour Therapy, 22*, 65-87.
- Raven, J. C., Court, J. H., & Raven, J. (1992). *Standard Progressive Matrices Manual*. Oxford: Oxford Psychologists Press.
- Riggio, R. E. (1986). Assessment of basic social skills. *Journal of Personality and Social Psychology, 51*, 649-660.
- Riggio, R. E., & Taylor, S. J. (2000). Personality and communication skills as predictors of hospice nurse performance. *Journal of Business and Psychology, 15*, 351-359.
- Riggio, R. E., Tucker, J., & Coffaro, D. (1989). Social skills and empathy. *Personality and Individual Differences, 10*, 93-99.

- Rinn, R. C., & Markle, A. (1977). Parent effectiveness training: A review. *Psychological Reports, 41*, 95-109.
- Rogers, C. R. (1942). *Counseling and psychotherapy*. Boston, MA: Houghton Mifflin.
- Rogers, C. R. (1951). *Client-centered therapy*. Boston, MA: Houghton Mifflin.
- Rogers, C. R. (1957). The necessary and sufficient conditions of therapeutic personality change. *Journal of Consulting Psychology, 21*, 95-103.
- Rogers, C. R. (1961). *On becoming a person: A therapist's view of psychotherapy*. London: Constable.
- Rogers, C. R. (1975). Empathic: An unappreciated way of being. *Counseling Psychologist, 5*, 2-10.
- Sackett, P. R., & Mullen, E. J. (1993). Beyond formal experimental design: Towards an expanded view of the training evaluation process. *Personnel Psychology, 46*, 613-627.
- Schneider, B., Wheeler, J. K., & Cox, J. F. (1992). A passion for service: Using content analysis to explicate service climate themes. *Journal of Applied Psychology, 77*, 705-716.
- Searle, J. (1979). *Expression and meaning: Studies in the theory of speech acts*. Cambridge: Cambridge University Press.
- Sells, J. M., Goodyear, R. K., Lichtenberg, J. W., & Polkinghorne, D. E. (1997). Relationship of supervisor and trainee gender to in-session verbal behavior and ratings of trainee skills. *Journal of Counseling Psychology, 44*, 406-412.
- Shadish, W. R., & Ragsdale, K. (1996). Random versus nonrandom assignment in controlled experiments: Do you get the same answer? *Journal of Consulting and Clinical Psychology, 64*, 1290-1305.
- Simard, L., Taylor, D. M., & Giles, H. (1976). Attribution processes and interpersonal accommodation in a bilingual setting. *Language and Speech, 19*, 374-387.
- Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton-Century-Crofts.
- Steckler, N. A., & Rosenthal, R. (1985). Sex differences in nonverbal and verbal communication with bosses, peers, and subordinates. *Journal of Applied Psychology, 70*, 157-163.
- Sundararajan, L. (1995). Echoes after Carl Rogers: Reflective listening revisited. *Humanistic Psychologist, 23*, 259-271.
- Sykes, J. B. (Ed.). (1982). *The concise Oxford dictionary of current English* (7th ed.). New York: Oxford University Press.
- Tannenbaum, S. I., & Woods, S. B. (1992). Determining a strategy for evaluating training: Operating within organizational constraints. *Human Resource Planning, 15*, 63-81.
- Tyler, L. E. (1992). Counseling psychology: Why? *Professional Psychology: Research and Practice, 23*, 342-344.
- Watkins, C. E. Jr., Savickas, M. L., Brizzi, J., & Manus, M. (1990). Effects of counselor response behavior on clients' impressions during vocational counseling. *Journal of Counseling Psychology, 37*, 138-142.
- Weber, R. P. (1990). *Basic content analysis* (2nd ed.). Newbury Park, CA: Sage.
- Wickman, S. A., & Campbell, C. (2003). An analysis of how Carl Rogers enacted client-centered conversation with Gloria. *Journal of Counseling and Development, 81*, 178-184.
- Wolf, M. M. (1978). Social validity: The case for subjective measurement. *Journal of Applied Behavior Analysis, 11*, 203-214.

- Wood, C. D., & Davidson, J. A. (1987). PET: An outcome study. *Australian Journal of Sex, Marriage and Family*, 8, 134-141.
- Zajonc, R. B. (1989). Styles of explanation in social psychology. *European Journal of Social Psychology*, 19, 345-368.

Acta Universitatis Upsaliensis

*Ego rtgj gpukxg'Uwo o ctkgu'qh'Wrrucx'Fkuagt vvkqpu
Itqo 'yj g'Hcewnf'qh'Uqekri'Uekgpegu*

Editor: The Dean of the Faculty of Social Sciences

A doctoral dissertation from the Faculty of Social Sciences, Uppsala University, is either a monograph or, as in this case, a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series *Ego rtgj gpukxg'Uwo o ctkgu'qh'Wrrucx'Fkuagt vvkqpu'ltqo 'yj g'Hcewnf'qh'Uqekri'Uekgpegu* (Prior to July 1985, the series was published under the title “Abstracts of Uppsala Dissertations from the Faculty of Social Sciences”.)

Distributor:
Uppsala University Library,
Box 510, SE-751 20 Uppsala, Sweden
www.uu.se, acta@ub.uu.se

ISSN 0282-7492
ISBN 91-554-6036-4