



drieharingstraat 6
postbus 1568
3500 bn utrecht
T 030 2 729 614/615
F 030 2 729 729

DOCTOR-PATIENT COMMUNICATION AND THE QUALITY OF CARE

Methods and Measures; Process and Outcome

report of an Invitational Workshop

Eds. S.S.L. Mol, J.M. Bensing

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Doctor-patient

Doctor-patient communication and the quality of care : methods and measures; process and outcome : report of an invitational workshop / eds. S.S.L. Mol, J.M. Bensing. - Utrecht : Netherlands Institute of Primary Health Care (NIVEL)

With ref.

ISBN 90-6905-222-9

Subject headings: doctor-patient communication / medical education.

Wordprocessing

Bernadette Kamphuys

Cover

Mieke Cornelius

CONTENTS

	page
PREFACE	5
INTRODUCTION	
1. DL Roter The significance of talk in the medical encounter	7
METHODS AND MEASURES	
2. L Meeuwesen Dotor-patient communication and continuity of care: an empirical plea for longitudinal research	17
3. MM Kuyvenhoven, RM Pieters Simulation versus real life patients	27
4. RM Pieters Utrecht Consultation Assessment Method (UCAM): reliability and validity of actor simulation in assessing consultation skills of trainees in the vocational training.	39
5. J van Thiel, H Kraan, CPM van der Vleuten Reproducibility of scores with the revised MAAS History-taking and Advice Checklist (MAAS-R)	47
6. HF Kraan, A Crijnen, J van Thiel, J van Dalen, C van der Vleuten The patient satisfaction with communication checklist	53
7. LHC Tan, CHGM Foolen, CPM van der Vleuten LACONTO: Development of a national instrument to assess consultation skills of future general practitioners	59
8. JM Bensing, LHC Tan, SA de Boer, SSL Mol LACONTO in everyday general practice: a multisystem comparison	65

PROCESS AND OUTCOME

- | | |
|--|----|
| 9. VCL Tielens | |
| Consultation length and the patient's view:
aspects of quality | 77 |
| 10. H Mookink | |
| Doctor-patient communication, medical performance
and patient characteristics | 87 |
| 11. I Heszen-Niejodek | |
| Polish research on doctor-patient communication | 95 |

LITERATURE 105

PREFACE

In 1991 a small but stimulating workshop took place on doctor-patient communication and the quality of care. It was organized by the Dutch College of General Practice (NHG) and the Netherlands Institute for Primary Health Care (NIVEL) who both stress the importance of doctor-patient communication as probably the most powerful tool of the General Practitioner. The workshop was attended by nearly all Dutch researchers on doctor-patient communication and some outstanding researchers from other countries. Several of them presented papers with original material from their own research; all of them participated in the discussions which were very lively and full of ideas. Although, of course, it will not be possible to transmit the atmosphere of the conference in this way, we have decided to publish the presented papers to give an idea about the main problems that were addressed during the workshop and about the main challenges for doctor-patient communication research in the next future.

Debra Roter took the responsibility for the Introduction to the workshop (and accordingly for the Introduction of this book) with a paper about 'the significance of talk in the medical encounter'. Two methodological problems were addressed by **Ludwien Meeuwesen** and **Marijke Kuyvenhoven**, respectively. The first author gives a plea for more longitudinal research in doctor-patient communication, because in her opinion the GP's continuity of care is not sufficiently reflected in cross-sectional research. The second author explores the methodological problem of choosing real patients, simulated patients or written vignets, which is often a problem of choosing between standardizing the situation and losing validity, or ensuring validity but creating a fuzzy situation with many disturbing variables. Next follows a section about measures for doctor-patient communication. In several Dutch University Departments of General Practice, observation systems have been developed for the assessment of communication skills among GP trainees. Examples are presented by **Ron Pieters** (Utrecht) and **Jacques van Thiel** (Maastricht). Herro Kraan (also from Maastricht) presents the results of a patient satisfaction scale that has been developed in conjunction with the MAAS-observation system. **Lisa Tan presents the first reliability- and validity tests of a national instrument** to assess consultation skills of future general practitioners, that is currently being developed by the joint University Departments of General Practice in the Netherlands (LACONTO). **Jozien Bensing** tested this instrument in everyday General Practice and compares its performance with two other observation systems that were applied to the same videotaped consultations.

In the last section some research studies are presented on the process and outcome of doctor-patient communication. **Victor Tielens** considers the value of consultation length and patient satisfaction as parameters for the quality of care, based on the results of a large study by the University of Nijmegen. **Henk Mookink**, from the same University, presents a GP-typology, which he relates to production scores like referrals and prescription of medicines; he also compares Dutch GPs with English and Belgium GPs, thereby opening the discussion on cross-cultural differences in doctor-patient communication. The last contribution is made by **Irena Heszen-Niejodek** who, on the same line, compares the results of her own Polish research on doctor-patient communication with the results of American research, also suggesting cross-cultural differences between the two countries, Polish doctors being warmer, but more authoritative than their American colleagues who gave their patients more influence, but displayed less affective behaviour.

The idea of more systematic cross-cultural comparisons proved to be an exciting idea to many attenders to the workshop, although many realized the conceptual, methodological and measurement problems involved. Nevertheless some plans were made, knowing that without starting one never gets far. Doctor-patient communication is worth the trouble.

Jozien Bensing

1. THE SIGNIFICANCE OF TALK IN THE MEDICAL ENCOUNTER

Debra L. Roter, Dr.P.H.

Associate Professor

The Johns Hopkins University School of Hygiene and Public Health

Department of Health Policy and Management

Baltimore, Maryland 21205

United States of America

A hot air balloonist was sailing over an open stretch of land and became confused as to her location. She called down to several people below and asked where she was. The people yelled back "Your up in the sky"... No, no, the balloonist called back, be more specific-where in the sky am I? "In a balloon" they said. I was hoping, the balloonist said, to hear something other than what I already know and which is totally useless.

If we as researchers aspire to make significant changes in how medicine is practiced, and to contribute more than "what everyone knows and is totally useless", we have a special duty to address the challenges of change that current knowledge about doctor-patient communication presents in a way which is clinically meaningful.

The goals of this paper are threefold and relate to these challenges. First of all, the themes in doctor-patient communication with the greatest "challenge to change" potential shall be discussed. These are not, as one would expect themes for which there are groundbreaking single studies that dazzle with their unique contribution and insight, but rather the themes that share an uncommon consensus about their importance from often disparate sources. Doctors and teachers of medicine, patients and in their less vulnerable state, consumers and clients and finally both qualitative and quantitative researchers have all addressed these themes in one way or another in their work. These themes have to do with both relationship and dialogue as requirements for curing and caring and relate to the fundamental need "to know and understand and to feel known and understood" (Engel, 1988).

Six themes of this kind will be discussed, with potential to transform the way doctor-patient communication is thought about, in a positive and meaningful way. Second, leaving the specific content-themes, the more generic challenges to the field from the theoretical, methodological, clinical, and training perspectives will be discussed. Finally, recommendations for future directions in

research shall be derived from these perspectives.

Figure 1 Six themes in doctor-patient communication with a great "challenge to change" potential

- 1.The importance of telling the story
- 2.Patients as experts
- 3.Beyond biomedicine to patient experience of life
- 4.Physician as expert
- 5.Information carries emotional content
- 6.Reciprocity

The importance of telling the story

The first communication-transforming principle to be addressed is that: *Communication should serve patients' need to tell the story of their illness and doctors' need to hear it. The function of the story is twofold: it is the scientific method by which the meaning of the illness and the meaning of the disease are integrated and interpreted.*

The emotional content and context of the medical encounter make this relationship different from others; people expect the doctor to know them in a fundamental and intimate way. In this realm it is important for patients to feel that their doctor takes a personal interest in them as individuals, that the doctor likes them, is concerned and committed to their welfare, and will consequently take pains to do a good job. Telling one's story fulfils this function. But that is not all. Telling one's story can be therapeutic in its own right in providing a cathartic release and the opportunity for insight and perspective.

The arithmetic of each physician having an average of 2500 patients does not equal the experience of each of those individual patients with his or her doctor. Each patient expects that treatment will be uniquely suited to their needs. But expression of these needs must be made within the constraints of short appointment blocks spread throughout the year. It is in this context that patients search for the opportunity to tell their story and to experience the feeling that their story is heard.

The telling is not so easy, however. Medical practice is widely characterized as taking place in a context of potential death, disability, trauma, pain, and uncertainty. And it is this characterization that often colors doctor-patient relations. Patients defend themselves against feelings of overwhelming complexity, demoralization, and helplessness by recourse to idealization or

denigration of the physician. Physicians, on the other hand, also need to protect themselves from feelings of grief and helplessness made worse by an overriding sense of their ever-present potential for making a fatal mistake (Hilfiker, 1985). In this context of heightened emotions and defensive responses, the stereotypic encounter is characterized as a retreat away from rational (logical) communication on the part of the patient, and a retrenchment to pure rationalism (to "science" of medicine) on the part of the physician. It is the patient story that can break this ever-escalating non-communication. The story can act to restore the dialogue and provide a basis for true talk.

Patients as experts

This leads to a second communication-transforming principle: *Communication should reflect patients' insight and expertise regarding their physical state, functional status, and quality of life. Patients should be considered "expert" in their own right and as such have a unique perspective and valuable insight into their illness condition.* In studies of over 23,000 people (Idler and Kasl, 1991) investigators have concluded that a simple self-evaluation of health - the answer to the question, "At the present time, how would you rate your health?" - provided strong clues to patient survival over periods as long as seventeen years. Furthermore, people's self-evaluations of health were stronger predictors of death than "objective" indicators of health, including detailed physical exams and batteries of tests.

Why would self-evaluations of present health predict future mortality so well, and better than physician assessments? No one really knows. It is possible that there is a direct effect of health optimism or pessimism on subsequent health. An individual who believes himself healthy then acts healthy, and indeed becomes more healthy over time; someone who believes himself sick, acts sick and indeed becomes sick. The second possibility is that self-evaluations reflect a calculation of life expectancy based on a broad range of information known only to the individual himself. This might include such "objective" things as his or her family's chronic disease history, longevity of parents and grandparents, or more "subjective" feelings - energy and well being.

Another explanation for these results has focused on judgments of quality of life. Physicians tend to associate general health ratings with the number and severity of chronic conditions, irrespective of the effect of these conditions on the patient's quality of life. However, the impact of disability associated with health conditions is highly variable and contingent not only on individuals themselves, but on their social network and resources.

Indeed, it has been found that patients' and physicians' understandings of

"overall health" differ in basic meaning, with physicians relying on narrowly defined medical data while patients draw on a more complex and multifaceted notion.

Patients' ratings of health are related to their functional abilities (such as the ability to walk without a cane or dress oneself), their emotional distress, and the number of different medical diagnoses appearing in their medical charts. They consider overall health to be a broad notion that covers distinct ways of being "healthy." In contrast, physicians are one-dimensional in their view; their biomedical blinders cut off all but one vision of patient health.

Communication then, should reflect patients' insight and expertise regarding all aspects of their health and illness experience - so that doctors' can indeed heed the words of Osler "Listen to the patient, he is telling you the diagnosis" (Osler, 1904).

Beyond biomedicine to patient experience of life

The importance of mental state to the patient's experience of health has relevance for the third communication-transforming principle: *Communication should facilitate recognition of the link between patients' mental state and their experience of illness- it should go beyond biomedicine to reflect and respect the patient's experience of life.*

By some estimates, doctors are thought to be able to diagnose and treat about 10% of conditions, and can diagnose but cannot treat another 10%, while the remaining 80% cannot even be diagnosed. Many of these conditions cannot be diagnosed because they are not symptoms of disease in the medical sense but are reactions to life. Headaches, rashes, dizziness, fatigue, stomach disorders, aches, chronic constipation or diarrhea, and weight fluctuations may very well portend problems of living rather than underlying disease. For some patients the distress they experience is expressed in genuine physical symptoms such as those just mentioned. People experiencing life stresses are very high users of health services for physical complaints which are never linked to any organic problem. In fact it is these patients which take the lion's share of physicians' time.

It is often difficult for patients experiencing physical symptoms related to life stresses to communicate these problems to their physician. While patients may be describing their experience from the perspective of their experience physicians are interpreting symptoms from their technological perspective. A patient may feel discouraged from talking about "non-medical" things such as stresses because this may not seem appropriate for the medical encounter, or because of a perception that the doctor can't do anything about it anyway. Both patients and physicians need to know this is not the case. The talk of the

medical visit should facilitate recognition of the link between patient's mental state and experience of illness. Together the doctor and patient can often come up with effective ways to meaningfully explore and address the physical manifestation, as well as the underlying cause of the patient's distress.

Physician as expert

A singularly consistent finding in studies of doctors and patients conducted over the past 25 years has been that patients want as much information as possible from their physicians. And this is the fourth communication-enhancing principle: *Physicians have the duty to share their medical expertise with patients in such a way that this information is clear, relevant, and useful to patients.*

The first definition for doctor in the Webster Dictionary is teacher. The word teacher implies helping, but this help is not limited to the usual clinical sense of providing correct diagnosis and treatment, or empathy and reassurance. A teacher helps by equipping learners (patients) with what they need to know to help themselves; this includes not just information but also confidence in the value of their own contributions. The educator model is thus more egalitarian and collaborative than the traditional doctor-patient model. In contrast to this collaborative model of adult education, pedagogist Paulo Freire (1970) describes the traditional model of education in which teachers deposit knowledge and directives for living into passive recipients as though they were empty vessels or bank accounts; indeed, he calls this the "banking" method of education. Freire advocates instead a relationship of teacher to learner that acknowledges what learners can impart to teachers. Thus, his model is one in which learning goes both ways and the teacher-learner distinction is considerably blurred.

Applying this to the doctor-patient relationship, one can imagine greater mutual recognition of the unique store of knowledge (not to speak of personal values) possessed by the patient, which can be as crucial for a positive treatment outcome as the physician's biomedical knowledge is. As the English researcher Tuckett noted: The medical visit is truly a "meeting between experts" (Tuckett et al., 1985). Even if one prefers the more traditional conception of the doctor alone as the expert, the doctor still has responsibility to educate patients. Most doctors do recognize this up to a point, but a critical difference between a true educator and the traditional doctor model is that the educator, by considering the patients' expertise, can share knowledge in a way that is most meaningful to the patient. While this is not meant to suggest that physicians should try to share the most technical aspects of their knowledge with patients, it is in fact

the case that the knowledge most pertinent to most patients' conditions is easily conveyed and readily understood. The imparting of this information creates a spirit of collegiality that enhances patients' ability and willingness to make informed decisions and meaningful commitments to treatment.

Information carries emotional content

Though patients certainly want as much information as possible from their physicians, information is not all that patients want. Physicians are not simply expert consultants, although they are that; they are also someone to whom people go when they are particularly vulnerable. People depend on this relationship to be in their best interest, trusting that the doctor is indeed "on their side." The fifth communication-transforming principle is: *Communication should acknowledge and attend to its emotional content. The exchange of information carries both cognitive and emotional significance.* Patients need factual information as a basis for "rational" behavior (for example, they can't take drugs intelligently if they don't know about the dosage and side effects). But they also need information for the reduction of uncertainty and anxiety. Further, physicians who provide adequate information are likely to be interpreted as being competent and caring, and these interpretations may be of major influence on the relationship and the entire course of illness, more important even than the information itself.

Most talk of the medical visit is information giving or question asking, with relatively little communication explicitly identified as the expression of emotions and feelings. This, of course, is not the full picture of the medical exchange, for the nonverbal channels of facial expression, voice quality, posture, touch, gaze, and so forth constantly communicate feelings even when the spoken words may not seem to have any particular emotional content. Feelings get introduced in another way as well. Patients draw conclusions about the doctor-what kind of a person the doctor is, and the nature of the doctor's attitudes and intentions toward the patient-from the totality of doctor behavior. Thus, even a doctor who is bland in expression and who spends all his or her time on medical business may still inspire good feelings in a patient because the patient values being taken seriously. The patient may infer that the doctor cares a lot because the doctor went to the trouble to be thorough, informative, and accurate.

In fact, all exchanges between doctors and patients carry cues about feelings and attitudes. People emit cues that are given meaning whether they want to or not. Thus, a distressed patient who tries to conceal her agitation may "leak" such cues anyway. A physician who equates professionalism with a neutral

demeanor and emotional distance will most likely react in spontaneous expressive ways in spite of these intentions - for example, by revealing some irritation with a particularly troublesome patient. And even if a physician is able to be consistently unexpressive, a patient may well interpret that lack of expression not as "neutrality" but as aloofness, disinterest, or even as a sign that important information is being withheld. Thus, doctor-patient interactions can be as emotionally laden as any other relationship, and more so than some because of the anxiety that is often brought on by the illness experience, and the patient's hypersensitivity to cues given off by the physician.

Reciprocity

The last theme to be addressed is one relating to the notion of reciprocity. Negotiation and bargaining are implicit in any relationship and thus also for the doctor and patient. The notion of reciprocity is a powerful motivating force in the exchange of the medical visit. The common notion of reciprocity centers on things people can do for, or give to, each other in a spirit of exchange-often linked to assessments of fairness, justice and good effort. This exchange includes both feeling and calculation. In medical encounters, (as in other relationships) there is a trade-off between good feelings and coolly calculated obligation. When feelings of gratitude, love, or esteem are strong, a sense of obligation is scarcely felt, a deed is done because one wants to. But when positive feelings are not so strong, a sense of obligation emerges, and reciprocity becomes a duty, a "should." Under these circumstances, awareness of calculation is likely to be quite high, and calculation of what is due can be exceedingly detailed. The sixth communication principle is based on this notion of reciprocity: *Communication should foster open negotiation in the fulfillment of mutual needs and expectations as doctors and patients continually evaluate the adequacy of each other's performance to their values and expectations. Each responds in a way that they feel somehow matches with, or is deserved by, the other's behavior.*

In the medical relationship a special case exists with regard to reciprocity. The doctor can do things for the patient and the patient can do things both for the doctor and for himself or herself. For example, the doctor who has worked hard to set a leg may find that the patient not only expresses positive feelings toward the doctor but may follow the doctor's advice with special conscientiousness - for example saying "the doctor worked so hard to set my leg, the least I can do is stay off of it, for him" or the doctor thinking to himself "why bother trying to get this patient's blood sugar under control when he will only be back in the same condition in a few weeks".

Patients can choose between doctor-directed responses and responses on their own behalf. If a physician fails to provide the drug regimen the patient prefers, the patient may reciprocate disappointment by noncompliance with the drug that was prescribed, or by dropping out of care altogether.

There are many ways in which a patient may try to please or "repay" a doctor. The patient can simply be friendly and positive and warm. In addition, payment can be quite literal: payment of bills is expected. But many patients go beyond the expected payment by providing their doctors with additional things, particularly at Christmas, on special occasions, and after a serious illness (Drew, Stoeckle, & Billings, 1983). Retaliation is also a form of reciprocity. Patients who feel they did not receive the kind of care they expected may withhold payment of bills, boycott further services, discourage friends from going to this doctor, or even file malpractice suits. The extent of retaliation is modulated by the reciprocal calculation; minor disappointments might result in partial noncompliance but larger disappointments might result in more drastic action.

More subtle reactions as well as moment-to-moment acts within the medical visit can also be seen in this light. Thus, positive statements or facial expressions by the doctor are likely to produce the same behavior in the patient. A warm greeting pleases the patient, and the patient returns good feeling (satisfaction) to the doctor. If one participant behaves in a reserved manner, chances are the other will behave coolly too. Behaviors that are more task-oriented in both doctor and patient can also show reciprocity; for example, more information giving by the doctor and greater expertise produce better understanding and better patient compliance. Though some of these benefits are fostered directly by the information, an equally important force is the motive to repay that is aroused by the patient's perception of good doctor performance. It is also likely that the reciprocation of task-oriented behaviors by other task-oriented behaviors is fostered when a positive emotional climate exists between the doctor and patient.

Reciprocity can also occur when the doctor's and patient's behaviors look quite different on the surface. For example, a patient may respond with liking (an emotional reaction) to a doctor who is highly competent (a technical skill). Such a reaction would be reciprocity if, as mentioned above, the patient interpreted such a doctor to be showing his own liking for the patient by being so conscientious. What does not happen much is the reverse - a patient making a competence attribution based on the doctor's friendly personal style. Just this

point was made through an experiment by Willson and McNamara (1982) which showed that people viewing taped vignettes of doctor-patient interaction interpreted a competent doctor to be courteous, but did not interpret a courteous one to be competent unless actual competence was portrayed. Courtesy and friendliness was not enough to convince patients of the doctor's expertise nor sufficient to guarantee commitment to the recommended regimen. Communication should foster open negotiation in the fulfilment of mutual needs and expectations as doctors and patients continually evaluate the adequacy of each other's performance to their values and expectations. Doctors would do well to consider the impact of what they do and say on their patients.

Challenges to the field: future directions

What recommendations can be made on the basis of the aforementioned?

Research should be more theory-driven and to be able to do this more should be known about the black box that is the talk of the medical visit. The black box is large, complex, and sensitive to inputs from many quarters - the biopsychosocial perspective being an important element in this. Like the blind men inspecting different parts of the elephants anatomy, fair detail about many individual pieces of this puzzle is known - what is not yet known is what type of animal it is. And also, is this animal in the process of metamorphosis?

The research should not only be done more critically, it should also be done with greater depth and breadth.

To this end, fuller collaboration among qualitative and quantitative researchers is necessary in the future and the boundaries and discipline-based loyalties must give way to a more insightful and richer approach.

It will be important to key this work to issues of quality of care - not only as a defined outcome measure but as a way to assure the usefulness and significance of it's contribution to medicine.

The author also hopes the scope of research will be broadened to include longitudinal studies - the current knowledge in this field being extremely limited. Do patients with longstanding relationships with their doctors, as is so common in the Netherlands, differ in fundamental ways from patients who lack substantial continuity as is often the case in the United States?

More thought has to be put in the training of young doctors. Not only in terms of content but also in the way they learn and to what end. Medical education is designed not only to teach biochemistry and physiology but also to teach how one acts as a doctor. Unfortunately, the current training process is all too often dehumanizing; young doctors are forced to give up any semblance of a normal home life and typically work every other night at the hospital (Klass, 1987). This

cannot help but create a callous view of patients: patients are but one more obstacle to seeing their family or getting some sleep. Even after training, many doctors feel their lives are unreasonably burdened by the stress of on-call responsibilities and hospital pressures. These experiences undoubtedly carry over into the medical visit in negative ways (Hilfiker, 1985).

Medical school needs to do a better job of helping young doctors define for themselves what is truly important about being a doctor and what are effective, and humane, doctor-patient roles. Our society must figure out how to influence their attitudes so that they come to value certain aspects of patient care differently. If physicians saw themselves more as patient educators, medical education would be different, and the profession would engage in a different kind of self-scrutiny. More attention would be paid to the process of teaching which will translate into more sensitive involvement of doctors in the process of healing.

Finally, the clinical implications of the work in the field of doctor-patient communication must be worked out. This calls for nothing less than a transformed clinical model of medicine which is patient centered and biopsychosocial in nature.

References

- DREW, J., J.D. STOECKLE, J.A. BILLINGS. Tips, status and sacrifice: Gift giving in the doctor-patient relationship. *Social Science & Medicine*; 17, 1983, p. 399-404.
- ENGEL, G. How much longer must medicine's science be bound by a seventeenth century world view? In: WHITE, K.L. *The task of medicine: Dialogue at Wickenburg*. California: Menlo Park, The Henry J. Kaiser Family Foundation, 1988
- FREIRE, P. *Pedagogy of the oppressed*. New York: Seabury Press, 1970
- HILFIKER, D. *Healing the wounds: A physician looks at his work*. New York: Pantheon, 1985
- IDLER, E.L., S.V. KASL. Health perceptions and survival: Do global evaluations of health status really predict mortality? *Journal of Gerontology*; 46, 1991, S55-65
- KLASS, P. *A not entirely benign procedure: Four years as a medical student*. New York: Signet, 1987
- OSLER, W. The master-word in medicine. In: *Aequanimitas with other addresses to medical students, nurses, and practitioners of medicine*. Philadelphia: Blakiston, 1904
- TUCKETT, D., M. BOULTON, C. OLSON, A. WILLIAMS. *Meetings between experts*. New York: Tavistock Publications, 1985
- WILLSON, P., J.R. MCNAMARA. How perceptions of a simulated physician-patient interaction influence intended satisfaction and compliance. *Social Science & Medicine*; 16, 1982, p. 1699-1703

2. DOCTOR-PATIENT COMMUNICATION AND CONTINUITY OF CARE: AN EMPIRICAL PLEA FOR LONGITUDINAL RESEARCH

By Ludwien Meeuwesen
University of Utrecht
Dept. of General Social Sciences
P.O. Box 80140
3508 TC UTRECHT
The Netherlands

Introduction

"About eight years ago, whilst conducting research on doctor-patient communication based on audiotapes, I had the following experience. The research material consisted of 600 tapes of about 100 patients, so there was more than one interview per patient. The first interviews of every patient were listened to. When, after this, random tapes of second or later consultations were listened one consultation seemed very familiar. At the same time I was also 100% sure that this was the first time I was listening to this tape. What was going on? Had a mess been made of the research material or was the administration not secure enough? Further investigation revealed that it was the fourth visit of one and the same patient to the physician. The conversation of this fourth interview was quite the same as the first one in this series; the same introductory, the same jokes, nearly a copy of the first. It reminded me of playing a record that has got stuck, hearing the same sound over and over."

This anecdote is not representative for doctor-patient communication, but it tells us something about the consistent patterns of doctor-patient communication, especially in case of somatic fixation. It confirms the importance of doing research in a more longitudinal way.

The subject of this study is an element that has been completely neglected up to now, it refers to doing **longitudinal research** on doctor-patient communication. First of all, I shall motivate the need for this type of research. The link with somatisation disorders will be illustrated and an example of longitudinal research will be shown, followed by some results of a discourse analysis of a series of five interviews of one and the same patient.

Why longitudinal research?

Most research on physician-patient interaction consists of observational studies of the association between communication behaviors and a variety of health care outcomes. A quick inventarisation of review studies of mainstream research learns that most of the research on doctor-patient communication is designed as large-scale quantitative observational studies which relate factors inside the consultation (patient and physician behaviors) to factors outside the consultation (satisfaction, compliance, etc.). Excellent reviews (Pendleton, 1983; Inui & Carter, 1985; Hall et al., 1988; Bensing, 1991) of the results across these studies have shown that patient satisfaction, compliance, and problem resolution improve when physicians establish rapport, encourage patients to express their concerns and point of view, provide information and use good interpersonal skills. All these studies are cross-sectional in nature, meaning that at one moment one interview of an x-number of patients is collected. Also the qualitative case-studies done by psycho- or sociolinguists using discourse analysis or conversation analysis, describing the inner dynamics of conversation more in detail, always refer to studies done at one point in time (e.g. Waitzkin & Stoekle, 1972; West, 1984; Waitzkin, 1985; Meeuwesen, 1988) . Although all these studies are of great importance indeed, there are several reasons for doing more longitudinal research.

One important reason has to do with the continuity of care: the contact of patients with their physician, GP, is longlasting and of a changing intensity. This is especially true in the Dutch situation, where most of the general practitioners know their patients quite well. In general one third of the complaints a patient presents are familiar to his or her GP. So one can ask oneself how the quality of care should be continued over a longer period of time and what the consequences for the doctor-patient relationship are.

Over the past decade, the nature of the complaints presented by patients have changed. Several infectious diseases disappeared; the same holds true for purely somatic diseases. And frequently, somatic complaints are psychosocial in nature. With a health care system becoming more and more complex and the patient's medical consumption rising, the GP has become more of a manager who coordinates the different sources of health care for a specific patient.

Two specific groups should be mentioned:

1. the increasing number of patients with chronic diseases, especially among the elderly. For this group continuity of care is extra important.
2. patients with somatisation disorders (approximately 15 to 20% of the patients). One of the characteristics is their high medical consumption.

The somatisation disorder in the context of the general practitioner's surgery will

be elaborated on. This will be done by describing a case study of longitudinal research on doctor-patient communication. It is meant as an illustration of the many ways in which one can do longitudinal research.

Somatisation

What is somatisation? And what do we know about doctor-patient communication in case of somatisation? There is a real and well documented danger that patients get stuck in and become dependent upon the medical system as a result of a one-sided and continuous emphasis on the somatic aspects of their complaints. This process is called somatisation or somatic fixation (Grol et al., 1983).

Several factors have been formulated, that contribute to the process of somatic fixation. These are individual, and environmental factors of the patients, and factors in the doctor-patient relationship. As far as the last factor is concerned, in case of somatic fixation it is assumed that the perspectives of physician and patient diverge and that because of this there will be a lot of negotiation.

Based upon a review of research and on the author's own research project the conclusion can be drawn that in case of somatising patients medical interviews take more time. The structure of the interview is more complicated. Somatising patients try to exert more control over the conversation, and show more signs of psychosocial problems (Meeuwesen, 1988).

In summary, the results of cross-sectional studies suggest that at the end of the interview somatic patients share the doctor's frame of reference more than the group of somatising patients do. In other words they both bring different perspectives to the interactions. In case of a doctor who is illness-centred this will be quite clear. But what happens when a GP also focuses on the psychosocial aspects of the patient's complaints? The next step is: How does one discuss these most delicate matters with the patient? A difference in perspective will often result in negotiation about the cause of the complaints.

Studying these issues by means of a cross sectional analysis may give some insight, but at the same time will leave many questions unanswered. For example, it does not inform us about the conditions under which sequences of negotiations occur, nor about the nature of the negotiation and the way it may change. What are the features of the rigid communication pattern, often characteristic for doctor-patient communication, in case of a somatisation disorder? Longitudinal research can offer more insight in the dynamics of interactional communication, causes of rigid patterns of communication. It might also give suggestions on how to change them in a more productive pattern of communication.

So the aim of the longitudinal approach will be to identify some factors in

doctor-patient communication that are associated with somatic fixation and that can possibly be used to prevent such processes. An example of longitudinal research in case of somatisation will be given now.

Method

The dataset for this project consisted of 600 audiotapes of some 100 patients seen by 10 physicians partly collected by the Nijmegen University Institute of General Practice, and partly collected by the author. It is considered a unique data-set: a group of about one hundred selected patients was followed during a 12 month period, and every visit to their GP was audiotaped. Of most patients more than one tape was available. The mean number of visits in that year was about 6, the maximum was 22.

The results given here are those of one and the same patient, who visited her doctor 5 times during that year, and who presented the same vague complaints over and over again.

Case description

The case refers to a married women, 56 years old. She frequently visits the doctor, (now) with belly pain as the most important complaint. She has also visited another GP. Both of the GPs discuss the psychosocial aspects of her complaints, (her isolated position in the neighbourhood, the relation with her husband and so on). Their styles differ a lot, one using far more time than the other (see Figure 1).

Figure 1 Five interviews of one patient

Time-lag	Interview	Physician	Duration	Complaint cq diagnosis
	First interview	(A)	5 minutes	Belly pain
After 2 months:	Second interview	(A)	4 minutes	Skin irritation, sleeping problems
After 4 months:	Third interview	(B)	13 minutes	Muscle pain, malaise, belly pain
After 1 week:	Fourth interview	(A)	15 minutes	Muscle pain, belly pain, sleeping problems
After 4 months:	Fifth interview	(B)	22 minutes	Throatache, belly pain

The method used in data collection is called the triangular method: several kinds of information about the same process are gathered in different ways. Emphasis is put on qualitative aspects of communication, to grasp the inner dynamics of the doctor-patient communication.

The three aspects observed were:

1. **Relational aspects of communication:** Verbal Response Mode System by Stiles: the focus is on the intent of communication, and not on the content (Stiles, 1978).
2. **Formal aspects of communication:** here, turntaking and division of turns is concentrated on, as well as the way in which communication is organized (how to get and hold the floor) (Sacks et al., 1974).
3. **Content aspects of communication:** attention is given to discussion about facts and discussion about values/norms (physical vs psychosocial aspects) or the way negotiation is expressed (Springorum, 1981).

One of the methods and results will be shown extensively; namely the first one: the Verbal Response Mode System by Stiles. This Verbal Response Mode system has been used frequently for analysis of medical interviews in the United States (e.g. Stiles et al., 1979; Stiles et al., 1982). The VRM system does not focus on the content of utterances but on the intent or relational aspect that an utterance has in communication. It consists of the following eight mutually exclusive and exhaustive categories, which are presented in Figure 2.

The classification of an utterance is determined by answering three two-choice questions:

1. does it refer to the experience of the speaker or the listener?
2. does it refer to the speaker's frame of reference or to the listener's?
3. does the speaker presuppose specific knowledge about the listener?

These three dichotomous classification principles are labelled:

- source of experience
- frame of reference
- focus.

The principles are dichotomous; each can have the value speaker or other, as Figure 3 shows.

Figure 2 The VRM (Verbal Response Mode) codes each utterance on linguistic form and intersubjective intent

Disclosure	(D)	Declarative, first person singular (I)	Reveals thoughts, feelings perceptions, intentions
Question	(Q)	Interrogative; inverted subject-verb order	Requests information or guidance
Edification	(E)	Declarative; third person (e.g., "he", "she", "it")	States objective information
Acknowledgement	(K)	Non-lexical or contentless utterances; terms of address and salutation	Conveys receipts of or receptiveness to communication
Advisement	(A)	Imperative or second person with verb of permission, prohibition, or obligation	Attempts to guide behaviour; suggestions, commands, advise
Interpretation	(I)	Second person ("you"); verb implies an attribute or ability	Explains or labels the other; judgements, evaluations of other
Confirmation	(C)	First person plural ("we") where referent includes other	Compares speaker's experience with other's; agreement, disagreement
Reflection	(R)	Second person ("you"); verb implies internal experience or conscious action	Puts other's experience into words; repetitions, restatements, clarifications

Note: Both the form and the intent of each utterance are coded; the intent symbol is written in parentheses. For example, "Would you roll up your sleeve?" is coded Q(A), which means question form and advisement intent.

Figure 3 Taxonomy of verbal response modes

Focus	Frame of reference	Source of experience	
		Speaker	Other
Speaker	Speaker	Disclosure (D)	Question (Q)
	Other	Edification (E)	Acknowledgement (K)
Other	Speaker	Advisement (A)	Interpretation (I)
	Other	Confirmation (C)	Reflection (R)

Results

The figures 4 through 6 show the results of the VRM categories, comparing five interviews of two GPs. They are based on an extensive analysis done by Schröder (1985).

These results give a global impression of the conversational structure of the medical interviews. As far as the physician's behavior is concerned, the two GPs are behaving in a quite different way. Although, as has been already mentioned, the conversations in case of doctor B took more time, doctor A gave relatively more utterances in the interview. Most of them were acknowledgements, and some other categories like edifications, disclosures and interpretations (see below on 'content'). The modes of doctor B were far more varied than those of doctor A: more questions were asked to the patient. This can explain the fact that the patient gives more information when confronted with doctor B.

The greatest differences in behavior of the patient refer to the fact that in case of the interviews with doctor B, the patient gave more acknowledgements and asked more questions.

The VRM-analysis gives a first overview of the conversational structure, and does not pretend to be complete, while "..... other kinds of information are needed to put flesh on the exchange structure skeleton" (Stiles et al, 1982).

As far as the analysis of turntaking was concerned, doctor A used more utterances to structure the conversation. Doctor B used more ways to get the floor or to hold the floor. It is suggested that doctor B had the intention to have more of his own contribution in the conversation.

In case of doctor A, the patient used varied means to hold the floor: such as like selfselection, attention getters, ending the turn of the conversation partner.

As far as the content was concerned, doctor A tried to clarify rather boldly a relation between the stress of the patient and the complaints she had (the patient wants to have an X-ray taken, but the GP does not believe that something is really wrong with her belly). The conversation was characterized by extremes: talking about somatic aspects, talking about stress. It was too confronting for the patient: she felt her doctor did not take her complaints seriously.

Doctor B tried to discuss and ask information about the problems the patient experienced with her complaints and with their psychological and sociological consequences. This turned out to be a more subtle way of handling the problems of the patient.

In summary, it can be stated that doctor B managed to establish - in psychotherapeutic terms - rapport: a relation based on trust, the patient feeling

recognized as a patient, with her pain. She was then able to express matters that were really worrying her, which in fact she did.

Figure 4 Utterances of GP A and GP B of 5 medical interviews

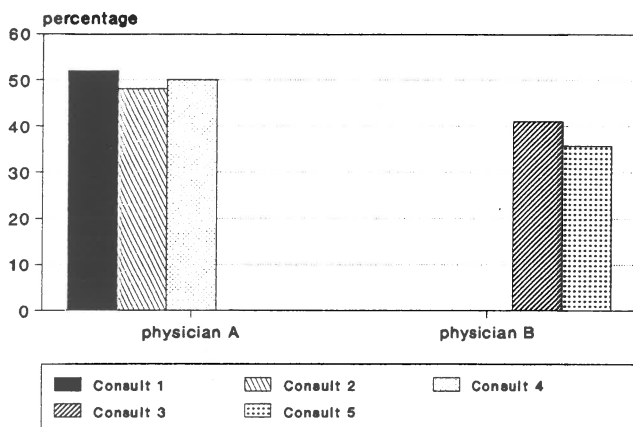


Figure 5 VRM-profile patient (5 interviews)

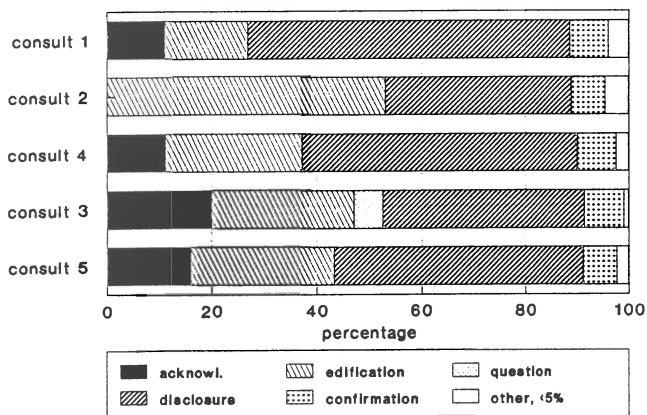
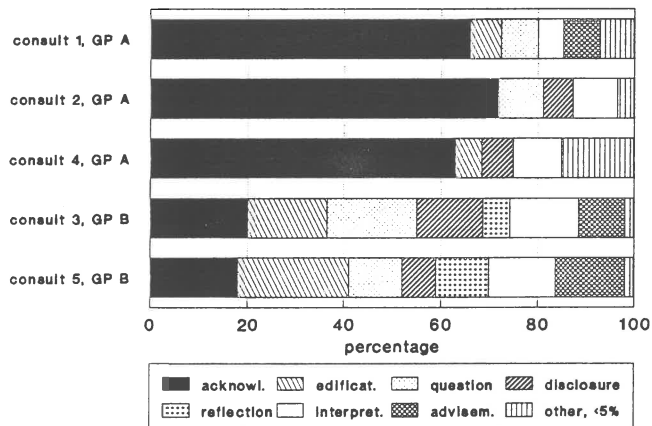


Figure 6 VRM-profile GP A and B (3 and 2 interviews)



Conclusion

In this specific example the conclusion would be that when the rigid pattern of communication between doctor A and his patient turned out to be fruitless, a fresh approach by another colleague gave the solution. The patient felt taken seriously by doctor B. Although one is never sure whether it is a matter of disappearing defense mechanisms on the side of the patient, it is the author's strong impression that the way the physician handled the problems greatly influenced the behaviour of the patient. The lesson to be learned from this example would be: do not discuss the genuineness of the complaints.

In the case study presented here, one and the same patient was followed during a certain period of time; a patient with a somatisation disorder. This is one way of doing longitudinal research. Quite another way would be to compare audio or video tapes from the sixties or seventies with tapes from the nineties. Subjects of interest could be the way a physician handles requests for information, or the attitude of physicians towards children.

A third suggestion, of a somewhat different nature: one could follow a patient who is a high consumer, on his or her various visits to several health services during a certain period. One could then focus on the cooperation between health care workers.

The benefits of doing longitudinal research are clear: instead of looking at a snapshot one is looking at a movie. This gives one insight in the way the process of communication can change over a longer period in time. It also

teaches about different aspects of the continuity of care and can be used as an instrument of evaluation. Hopefully this study will stimulate the use of longitudinal research in studying doctor-patient communication.

References

- BENSING, J. Doctor-patient communication and the quality of care: An observational study into affective and instrumental behavior in general practice. (Dissertation). Utrecht: NIVEL, 1991
- GROL, R.P.T.M. et al. Huisarts en somatische fixatie: Theorie en praktijk van de preventie van somatische fixatie (General practitioner and somatic fixation: Theory and practice of the prevention of somatic fixation). Utrecht/Antwerpen: Bohn, Scheltema & Holkema, 1983
- HALL, J.A., D.L. ROTER, N.R. KATZ. Meta-analysis of correlates of provider behavior in medical encounters. *Medical Care*; 26, 1988, p. 657-675
- INUI, T.S., W.B. CARTER. Problems and prospects for health services research on provider-patient communication. *Medical care*; 23, 1985, p. 521-538
- MEEUWESSEN, L. Spreekuur of zwijguur: Somatische fixatie en sekse-asymmetrie tijdens het medisch consult (Speaking or being silent: Somatic fixation and gender asymmetry during the medical interview.) (Dissertation). Nijmegen: University of Nijmegen, 1988
- PENDLETON, D. Doctor-patient communication: A review. In D. PENDLETON & J. HASLER (Eds.), *Doctor-patient communication*, p. 5-53. London: Academic Press, 1983
- SACKS, H.E., E.A. SCHLEGLOFF, A. JEFFERSON. A simplest systematics for the organisation of turntaking for conversation. *Language*; 50, 1974, 4, p. 696-735
- SCHRÖER, C. Geleidelijk onderhandelen (Gradual negotiation) (Unpublished doctoral thesis). Nijmegen: University, Department of Clinical Psychology, 1985
- SPRINGORUM, D. Spreken in gesprekken (Speaking in conversations). Groningen: Wolters-Noordhoff, B.V., 1981
- STILES, W.B. Verbal response modes and dimensions of interpersonal roles: A method of discourse analysis. *Journal of Personality and Social Psychology*; 36, 1978, p. 693-703
- STILES, W.B., S.M. PUTNAM, M.H. WOLF, A.J. SHERMAN. Interaction exchange structure and patient satisfaction with medical interviews. *Medical Care*; 17, 1979, p. 667-681
- STILES, W.B., S.M. PUTNAM, M.C. JACOB. Verbal exchange structure of initial medical interviews. *Health Psychology I*; 4, 1982, p. 315-336
- WAITZKIN, H., J.D. STOECKLE. The communication of information about illness: Clinical, sociological and methodological considerations. *Advances in psychosomatic Medicine*; 8, 1972, p. 180-215
- WEST, C. Routine complications: Troubles with talk between doctors and patients. Bloomington: Indiana University Press, 1984

3. SIMULATION VERSUS REAL LIFE PATIENTS

M.M. Kuyvenhoven and R.M. Pieters
Department of General Practice/Family Medicine
University of Utrecht
Bijlhouwerstraat 6
3511 ZC Utrecht
The Netherlands

Introduction

The point of this paper relates to the question of "standardized patients versus real life patients". In general, real life patients are preferred above standardized or simulated patients in assessing the quality of physicians' work. How does the validity of standardized patients compare to that of real life patients in assessment of the quality of general practitioners' performance - especially their communication skills.

Performance of physicians in real life may be recorded by video or audio cassettes. These approaches have a high face validity. Because there is a high intra-individual variation and case-specificity in coping with patients' problems, one needs several - and often many - rather similar cases, to make a reliable assessment of physicians' performance for summative evaluation and research. In "real life patient-doctor encounters" recorded by audio or video, however, the medical history and the history of the patient-doctor relationship is often unknown. Yet these may strongly influence the physicians' performance.

Another factor that makes comparison of physicians difficult is the fact that motivational and situational variables e.g. the stress of a busy day, influence physicians when they are treating patients. The stimulus is not objective.

A simulation with standardized patients is an alternative. It may compress processes which normally occur over a longer period of time: This is called telescoping of time. Physicians can be confronted with identical patients and situations, so there is a more objective stimulus. Thus comparison of physicians is more legitimate.

Standardized patients in a standardized setting are suitable to assess physicians' competence rather than performance. Competence is defined by Senior as "what a physician is capable of doing" (Senior, 1976). Performance as "what a physician actually does in his day to day practice" (Senior, 1976). It is common reasoning that competence is a good predictor of performance.

At first glance one has the tendency to stress the importance of physicians' performance above competence. But is there such a thing as general performance in, for example upper respiratory tract infections? It depends not only of the physicians' skills and competence, but also on the setting and the patient [performance = f (competence + setting + case)]. This explains the frequently rather low correlation between measurements of competence and performance. One has to consider the fact that there is a relatively high intra-individual variation over cases and a high interaction among physicians and patients even in one well defined domain of complaints (MCQuire, Page, 1973; Elstein, Schulman, Sprafka, 1978; Norman, Tugwell, Feightner, Murrin, Jacoby, 1985). In using one or two cases in assessing quality of performance there is a high risk of at random misclassification which, in general reduces the correlation with other measurements.

There are several reasons for the interest in competence rather than in performance. The medical curriculum and the Vocational Training for General Practice primarily guarantees a certain level of competence. Furthermore, the health care policy institutions are interested in competence and the variables which are associated with it.

Design of the paper

In this paper the following points will be discussed.

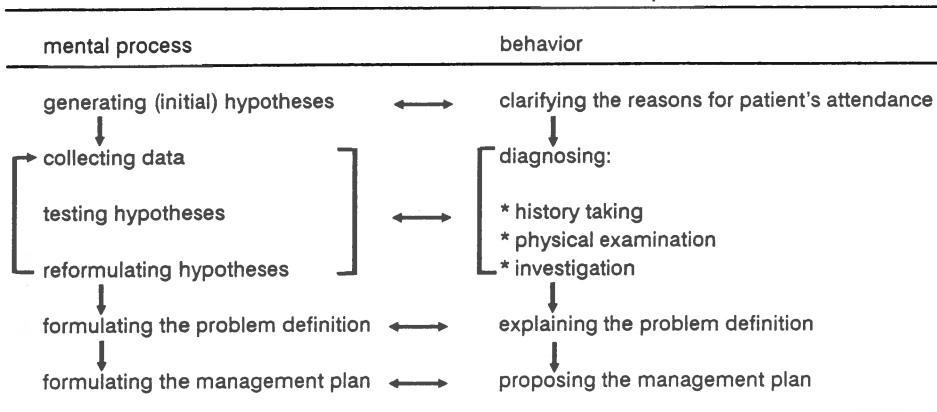
First of all the problem-solving process will be described. The next point is the difference between written and actor simulation and their validity in comparison with real patients. The validity of each method will be described by a case; one by Marijke Kuyvenhoven relating to written standardized patients and one by Ron Pieters, relating to actor simulation (Pieters, 1992). The aim of both examples is to make clear which aspects of competence are assessed and what is the validity of the method in comparison with reality. Which part of patient-doctor communication can be assessed per case and what the relevance is in view of the quality of care will be elucidated.

The problem-solving process

Ideally the problem-solving process of physicians may be delineated as follows. Firstly, the doctor clarifies the reasons for patient's attendance (request for

help). Then he takes the history and if necessary, does a physical examination and decides on further investigations. Based on the findings in the diagnostic phase the problems are defined. Finally, he formulates a management plan and discusses it with the patient (see Figure 1) (Kuyvenhoven, 1988).

Figure 1 The diagnostic-therapeutic process of physicians: the observable behavior and the non-observable mental component



Two components may be distinguished in the problem-solving process: a visible component, the performance and an invisible one, the mental component. The mental problem-solving process of physicians seems to have a general form; if a patient presents a complaint, physicians generate two to six hypotheses within the first few minutes. These hypotheses are called initial hypotheses to distinguish them from those hypotheses which are formulated after the diagnostic phase as the problem definition.

It is not possible to measure medical knowledge or the mental, cognitive competent independently from the observable behaviour (performance). Competence, which is the capability of doing something, consists of both mental/cognitive components and observable behaviour. Measurements implicitly assess knowledge and the mental problem-solving process.

Standardized patients; actor or written

There are two main designs of simulation: (a) actor simulation and (b) written simulation. Actor simulation, where an actor or actress plays the role of the patient is obviously more realistic than written simulations.

Written simulation refers to extended written standardized patients. Requirements are the availability of many topics or categories with questions, no cues

for a specific management and an opportunity to formulate a diagnosis and to propose management as in real life.

The written simulation is a reliable method to ascertain the way a physician integrates medical knowledge and diagnostic and therapeutic skills in his performance. Because of the relatively high reliability this method can be used for research and selective evaluations. In general written simulation assesses fewer aspects of competence than actor simulation. Written standardized patients are suitable for measuring somatic competence, patient orientation (the instrumental aspect of doctor-patient communication) and the quality of the management. These aspects can also be assessed by actor-simulated patients. On top of that "exploring reasons for encounter", "communication skills" and "structuring" the interview can be assessed in the latter manner of simulation.

Another component in the measurement of competence is to define what is meant by good care. However, this is also a problem when using real patients. It is possible to define more or less explicit guidelines with regard to somatic competence, history taking, physical examination etc. Relating to the other aspects of competence there are less explicit criteria. That is why the ratings of experienced doctors are often used in assessing these aspects. In that case the reliability of scores of raters must be analysed, even when the instructions are very explicit.

Figure 2 Standardized patient; actor or written?

Aspects of competence	registration by means of standardized patients	
	actor	written
- exploring reasons for encounter	+	-
- structuring the interview	+	-
- somatic competence	+	+
- patient orientation	+	+
- communication skills	+	-
- quality of management	+	+

- + suitable
- not suitable

At last a few words about vignettes. These are short written cases with structured or open answer categories. They are not suitable to assess a GP's skills in history taking, physical examinations and further investigations, but are aimed at describing differences in management due to known differences between cases or due to respondent's characteristics. They should contain no cues for a

specific management. The costs of this method are low in relation to the relatively great number of respondents one can reach. This method was used, for example, for studying the rough differences in antibiotics prescription and in patients with a sore throat, acute otitis media or sinusitis (Melker, Kuyvenhoven, 1991). By means of a secondary analysis of data collected in the NIVEL's National Study of Morbidity and Intervention in General Practice, a study is presently being done, to see if the differences found are realistic.

Case: written simulations of patient-doctor encounters

The aim of the construction of this simulation of written patient-doctor encounters was to validate a research instrument to measure differences in clinical competence of general practitioners (GPs) when dealing with patients with non-specific complaints (Kuyvenhoven, 1988).

The simulation consists of consultations with five "written patients" with non-specific complaints such as low backpain, tiredness etc. These patients are represented by a system of questions and answers. These questions, presented in the form of a "menu card" are arranged under different headings or topics, such as "history" or "self-medication". The GP selects a heading and chooses the question he wants. After each question it is possible to change categories.

At any moment the GP is allowed to conclude the encounter. He is asked to say what he would have told the patient in the way of a "strategy". This is recorded on tape. He may write a referral letter or a prescription. No cues for specific management are given. The GP has two encounters with every patient, an initial and a repeat encounter.

Three aspects of competence are independently assessed: somatic competence, patient orientation and the management's potential risk for unnecessary harm. The ratings are scored on a five-point scale. Somatic competence relates to the quality of attention paid to somatic aspects or causes of the complaints (raters: 3 GPs). Patient orientation relates to the quality of the physician's understanding of the patients' reaction to his complaints (raters: one GP and one psychologist). This aspect will be described further on. The third aspect is the management's potential risk for unnecessary harm. Each medical decision can induce unwanted results such as anxiety in patients, sick role fulfilment, side effects of drugs, lesions caused by invasive clinical examinations. The expected advantages do not always outweigh the possible disadvantages.

Patient-orientation

This concept is similar to what is called the instrumental or task-oriented aspect of doctor-patient communication. When confronted with non-specific complaints the GP should explore how the patient defines his own complaints and how they trouble or worry him. Appropriate questions are: "what do you think about your complaints" or "are you concerned you have a serious disease"? Thus the GP investigates the presence of a vicious circle of anxiety and complaints.

After the diagnostic phase the GP should interpret the somatic aspects of the symptoms to the patient providing insight and reassurance. The general practitioner can add that such symptoms often occur in people under stress and ask the patient if this might be a possibility in his or her case.

When a patient is anxious about his symptoms the general practitioner should explore the coincidence of anxiety-provoking events and continuation or exacerbation of the complaint. Awareness of this vicious circle may help the patient to cope with it. Doctor-patient agreement about the nature of the problem has been related to a relatively early resolution of health problems (Bass, Buck, Turner et al., 1986).

The doctor's understanding of the patient's reaction to his symptoms (patient-orientation) also requires awareness of the psychic functioning of the patient, e.g. does the patient suffer from a depressive mood or some psychic disorder? Depressive feelings may be connected with the way the patient experiences the complaints. If the patient seems to be depressed, nervous or anxious, the general practitioner should explore this feeling and the relationship with life-events. When the patient presents non-specific symptoms the general practitioner should limit himself to expressing empathy and relating the patient's mood to his physical complaints.

Finally, patient-orientation also means that the physician must be aware of the consequences of the symptoms for the patient and his family.

Summarized the quality of patient-orientation means the quality of the physician's understanding of the patient's reaction to his or her complaints. This is the mental or cognitive base of communication skills.

It is possible to assess the quality of a physician's patient orientation with written simulations under two conditions: there must be enough questions related to the just mentioned components of patient-orientation and, secondly, the physician must be enabled to conclude the encounter without any cues.

Reliability and validity

The validation of the simulation in a pilot study with 19 GPs consisted of three parts. Firstly the face-validity was investigated. A semi-structured interview was held with the GPs after the simulated encounters to elicit their experiences in the

simulation. In the simulation the GPs recognized many of the feelings they experience in real life situations. As for example one GP said: "To my mind, work in normal practice is the same as here, except that it is more laborious here". "I clearly recognize the same feelings of certainty and uncertainty".

The practitioners formed a vivid mental picture of the simulated patients, as illustrated by the following remark with regard to the first patient: "She somatizes and manoeuvres the physician into thinking that he has to solve the problem. She is rather reticent, not frank. She seems rather dependent".

Simulated history-taking, physical examination and laboratory work were sometimes more extensive than in reality, a finding which is in agreement with the reports of others. The availability of many questions of this type as well as the relatively low pressure on time stimulated them to ask such questions. By asking more questions, the physicians were seeking a firmer basis for a diagnosis already made.

In the simulation the GPs were entirely free to conclude the encounter whenever they wished and could record what they would say to the patient in reality, without being offered policy alternatives. Thus there was no cueing problem of importance. After listening to the recorded conclusions of each encounter, the physicians confirmed that their therapeutic procedures in the simulation closely resembled those they used in normal practice.

In general, the face validity of the three aspects is rather good. The next validation step was the reliability of the ratings.

"Somatic competence" and "potential risk for harm" were reliably rated (intra-class correlation coefficient $r = .90$ respectively $.80$) (Kuyvenhoven, 1988). Patient-orientation was less reliably assessed, mainly due to differences in judgment between the GP and the psychologist. The differences were not related to what they noticed, but to what they interpreted as important feelings of the patient. These differences would also exist in rating real life doctor-patient encounters.

A third parameter of validity was the comparison of the performance in the simulation with that in reality. Having no valid counterpart this comparison had to be made indirectly by resorting to prescription and referral data of the Sickness Funds.

A potential risk for unnecessary harm to the patient is, in general, induced by a GP who prescribes drugs where the indication is either arbitrary or undetermined and which may induce side effects or adverse effects. This risk often applies to analgesics, antirheumatics, hypnotics, sedatives, tranquilizers, neuroleptics and antidepressant drugs.

Referral to a specialist can also induce a potential risk for unnecessary harm.

Many complaints - such as gastric and abdominal pain and dyspnoea - have no readily definable pathological origin and a first examination by a specialist may reveal no definite causes. In this situation specialists often continue the investigation with more invasive methods, which may entail both inconvenience and potential risk for harm to the patient.

The results showed that the GPs who in reality were consistently frugal (or generous) in prescribing the earlier mentioned drugs and also frugal (or generous) in referring patients to medical specialists, indeed induced little (or much) potential risk for unnecessary harm in the simulation (Kuyvenhoven, Jacobs, Touw-Otten, Van Es, 1984).

So we conclude that the extended written standardized research technique developed is a reliable and valid method for discriminating between GPs who are more likely to cause potential risk for harm to their patients and those who are less likely to do so.

Degree of peer consultation

The simulation has also been used in another study which investigated the question whether there is a relationship between degree of interaction and consultation among peers - from now on called: "degree of peer consultation" - and practice setting on the one hand and the ratings on the 3 dimensions of clinical competence on the other hand (Kuyvenhoven, Pieters, Jacobs, Touw-Otten, Spreeuwenberg, 1990).

Figure 3 Correlations (Kendall's tau) between the three aspects of clinical competence (n=49) (Pieters, 1992)

	1	2	3
1. somatic competence	-		
2. patient-orientation	.34*	-	
3. potential risk for harm	-.16	-.54*	-

* $p \leq .05$

The rating of "somatic competence" was weakly associated with the rating of "patient-orientation" ($\tau=.34$) and not at all to "potential risk for harm". The rating of "patient-orientation" was moderately correlated with the "potential risk for unnecessary harm" ($\tau=.54$), indicating that general practitioners who had a high rating of patient-orientation induced less potential risk for unnecessary

harm than those who had a low rating of patient-orientation. A high rating of potential risk for unnecessary harm means unnecessary referrals or further investigations.

A causal model of the relationship between degree of consultation and practice setting on the one hand and the three aspects of competence on the other hand was tested by means of elaboration and practical correlation analysis (Kuyvenhoven, 1988).

There was a direct relationship between degree of peer consultations and the rating of somatic competence. (.24). GPs who do not consult peers at all had a lower rating of somatic competence than general practitioners who consult peers anyway. Practice setting was only related to potential risk for unnecessary harm (-.35). Those working in health centres induced less potential risk for unnecessary harm than those working in groups or solo-practices.

Nor the degree of peer consultations, neither the practice setting were associated with patient-orientation. This aspect, however was related to the GPs' opinions of the degree to which several parts of the vocational training (case discussions, training in communicative skills and discussing topics in general practice) contributed to their preparation for general practice (Figure 4). These opinions are also personal characteristics.

Figure 4 Correlations (pearson correlation coefficient r) with regard to the opinion of the degree to which several parts of the vocational training contributed to their preparation for general practice on the one hand and the three aspects of competence, peer consultation and practice setting on the other hand (n=49).

	Opinion of the degree to which several parts of the vocational training contributed to their preparation for general practice		
	case-discussions	communication skills	GP-topics
somatic aspects	-.14	.10	.36*
patient-orientation	.34*	.37*	.27*
risk of harm	-.29*	.13	.13
peer consultation	-.03	-.07	.10
practice setting	-.11	-.05	.07

Summary of results and conclusions

Conclusion regarding written simulations

- * requirements:
 - availability of many headings and questions
 - no cues for specific management
 - opportunity to formulate diagnosis/propose management as in reality
- * suitable:
 - somatic competence
 - quality of management
 - patient-orientation (cognitive base of communication skills)
- * not suitable:
 - exploring reasons for encounter
 - communication skills
 - structuring the interview
- * ratings of patient orientation are moderately reliable:
 - complexity of this aspect of competence
 - independently of 'real life' or 'standardized patients'

Conclusions of the study

- * somatic competence, patient orientation and potential risk for harm are loosely associated
- * degree of peer consultations and practice setting are associated with somatic competence and potential risk for harm; patient-orientation is a more individual, physician characteristic
- * case-discussions, training in communicative skills and discussing topics during the Vocational Training seem to contribute to the quality of patient-orientation of GPs

General conclusions

- * quality of primary care consists of various aspects regarding competence and performance
- * there is a high intra-individual variability in performance; many cases are necessary to assess performance reliably and validly
- * the aim of assessments defines suitability of parameters and measurement methods

References

- BASS, M.J., G. BUCK, L. TURNER ET AL. The physician's actions and the outcome of illness in family practice. *J Fam Pract.*; 1986, 23, p. 43-47
- ELSTEIN, A.S., L.S. SCHULMAN, S.A. SPRAFKA. *Medical Problem Solving: an analysis of clinical reasoning.* Cambridge: Harvard University Press, 1978
- KUYVENHOVEN, M.M. Quality of performance of general practitioners confronted with patients' non-specific complaints. Dissertation University of Utrecht 1988
- KUYVENHOVEN, M.M., H.M. JACOBS, F.W.M.M. TOUW-OTTEN, J.C. VAN ES. Written simulations of patient-doctor encounters 1. Comparison of the performance in the simulation with prescription and referral data in reality. *Family Practice*; 1984, 1, p. 25-29
- KUYVENHOVEN, M.M., H.M. PIETERS, H.M. JACOBS, F.W.M.M. TOUW-OTTEN, C. SPREEUWENBERG. The association between peer consultations and three aspects of clinical competence. *Family Practice*; 1990, 7, p. 163-167
- MCQUIRE, C.H., G. PAGE. The assessment of clinical performance by written and oral simulation. Report to the faculty 1972-3. Chicago/Illinois: Centre for Educational Development, University of Illinois College of Medicine, 1973
- MELKER, R.A. DE, M.M. KUYVENHOVEN. Management of upper respiratory tract infections in Dutch family practice, *British Journal of General Practice*; 1991, 41, p. 504-507
- NORMAN, G.R., P. TUGWELL, J.W. FEIGHTNER, L.J. MURRIN, L.L. JACOBY. Knowledge and clinical problem solving. *Med. Educ.*; 1985, 19, p. 344-356
- PIETERS, H.M. Utrecht consultation Assessment Method. Workshop Doctor-patient communication, June 1992
- SENIOR, J.R. Towards the measurement of competence of medicine. Philadelphia: National Board of Medical examiners, 1976

4. UTRECHT CONSULTATION ASSESSMENT METHOD (UCAM)

Reliability and validity of actor simulation in assessing consultation skills of trainees in the vocational training. Construction and validation of a measurement instrument.

Ron Pieters
Department of General Practice/Family Medicine
University of Utrecht
Bijlhouwerstraat 6
3511 ZC Utrecht
The Netherlands

Introduction

Over the past decade a major development, in terms of a process of professionalisation, has taken place in general practice. This process has been related not only to the content of general practice but also to training methods involving general practice skills. Consultation skills have been described more thoroughly. Various theories, based on systematic approach, have thrown fresh light on working methods and on prevention of somatic fixation. These studies and theories have opened the way to the formulation of various approaches to patients' complaints. It has become possible to draft educational objectives for vocational training as a prerequisite for the construction of a curriculum. A curriculum based on educational objectives requires assessment and evaluation. In the early eighties a method for the assessment of consultation skills of trainees in the vocational training course was launched at the Department of General Practice in Utrecht, the Netherlands. The Utrecht Consultation Assessment Method (UCAM) was developed to provide a profile of the trainees consultation skills. It reveals strong and weak points of the trainees skills.

The formal form of this instrument UCAM went through a long development phase. It was necessary to operationalize the educational objectives into observable and assessable items. The next task was to establish when and where assessment should occur and who would be the assessor. It was decided that clinical competence was to be assessed at the end of the third month of the one year vocational training course, followed by a second check six months later. The assessment took place at the department of General Practice with the help of simulated patients. One can choose between written

standardized patients and actor simulation (M. Kuyvenhoven, in these proceedings). For the assessment of communication skills actor simulation seems more suitable.

The simulated cases were patients with complex, multiconditional problems and complaints such as low back pain, abdominal pain or headache. These problems are not easy for physicians to interpret and they offer the general practitioner many options for diagnosis and therapy. Psychosocial and somatic aspects may play a role in these cases.

These complaints require sophisticated clinical performance on the part of a general practitioner. In order to prevent the general approach from focusing on either only the psychosocial or only the somatic aspect, a systematic approach is required. Therefore an effective management of the doctor-patient relationship and an adequate psychosocial approach are required.

The various skills a general practitioner should be able to exhibit in the communication with the patient were translated into observational categories with a rating scale. The observational instrument consisted of 24 items. An explanatory dictionary was used for the items and the way of scoring. For each item the observer rated on a 3-point scale (adequate/ uncertain/ inadequate).

The research questions of the study concern the validity, reliability and feasibility of the UCAM.

This paper focuses on the external validity only.

Method

In 1987 and 1988 data was gathered at the Departments of General Practice in Utrecht and Groningen. The Utrecht study group consisted of 98 trainees , covering 8 training groups; Groningen participated with 1 training group of 12 trainees. The trainees performed a consultation with a simulated patient at the Department of General Practice on two separate occasions. Video recordings were made of consultation activities in the trainees practice at the same time. The ratings of these consultations was done by 6 staff members of the Utrecht Department and 2 staff members of the Groningen Department.

Results

The study of the content validity revealed that UCAM covered a high percentage of the educational objectives. Only a few tasks, carrying little weight in consultation with patients with multiconditional complaints, were excluded.

The external validity, in terms of a comparison between consultation behaviour measured by simulation and in day-to-day practice, showed divergent results.

The recognition of behaviour in practice by using a simulated session is given in

the Figures 1 and 2. Probabilities were calculated of trainees with an adequate/inadequate rating in practice in relation to their ratings when seeing a simulated patient.

Trainees with an adequate rating in practice were well recognised by simulation, for all aspects, with exception of 'doctor-patient interaction' (Figure 1) at the first moment.

Figure 1 Sensitivity of the simulation test for an adequate rating

Skills	t1		t2	
	p(p+)	p(s+/p+)	p(p+)	p(s+/p+)
Systematic approach	.06	1.00	.02	1.00
Non somatic approach	.03	1.00	.02	1.00
Communication skills	.12	1.00	.08	.80
Doctor-patient interaction	.18	.33	.09	.83
Information giving	.27	.56	.14	.67

- p(p+) = proportion of trainees with an adequate rating in practice
- p(s+/p+) = proportion of trainees with an adequate rating in simulation given an adequate rating in practice
- t1 = after three months vocational training
- t2 = after nine months vocational training

Trainees with an inadequate rating in practice were moderately to fairly well recognized by simulation (Figure 2).

Figure 2 Sensitivity of the simulation test for an inadequate rating

Skills	t1		t2	
	p(p-)	p(s-/p-)	p(p-)	p(s-/p-)
Systematic approach	.74	.64	.83	.74
Non somatic approach	.92	.94	.86	.79
Communication skills	.55	.78	.72	.41
Doctor-patient interaction	.42	.86	.58	.84
Information giving	.29	.80	.47	.73

- p(p-) = proportion of trainees with an inadequate rating in practice
- p(s-/p-) = proportion of trainees with an inadequate rating in simulating given an inadequate rating in practice

The rating 'adequate' for skills applied in the simulated session was not a good predictor of the same skills in day-to-day practice.

There is a great proportion of 'inadequate ratings' in practice. The 'inadequate' ratings however, were good predictors.

In the vocational training program it is important to know whether the assessment of the trainee in a simulated session at the department is a good predictor of the quality of the trainee in real practice. The special interest of this study was in extreme differences: could a trainee with an adequate rating in practice get an inadequate rating in the simulation? The proportion of adequate ratings to be seen in the first column (Figure 3) is low so the chance of an adequate rating in practice is low.

Figure 3 Predictive value of an inadequate rating in the simulation

Skills	t1		t2	
	p(p+)	p(p+/s-)	p(p+)	p(p+/s-)
Systematic approach	.06	.00	.02	.00
Non somatic approach	.03	.00	.02	.00
Communication skills	.12	.00	.08	.13
Doctor-patient interaction	.18	.17	.09	.04
Information giving	.27	.33	.14	.11

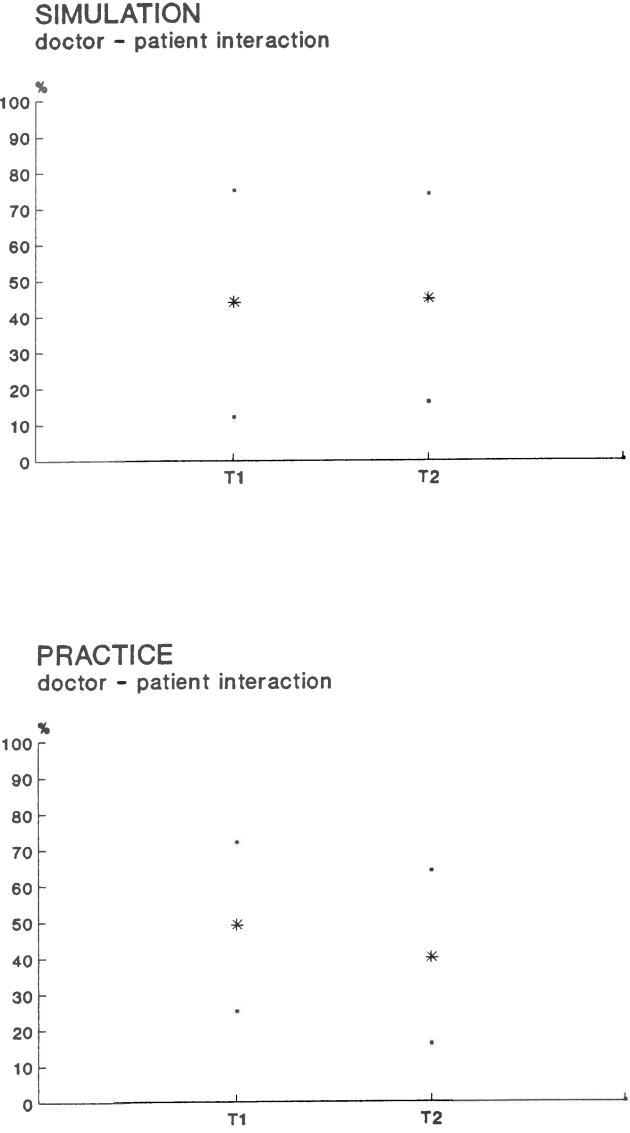
p(p+) = proportion of trainees with an adequate rating in practice

p(p+/s-) = proportion of trainees with an adequate rating in practice given an inadequate rating in simulation

The results of the external validity are related to the fact that the observed general skills were more adequate in simulated sessions than in practice. The results of all trainees for 'systematic approach' and 'doctor-patient interaction' are given in Figures 4 and 5 respectively. The results of the other aspects were the same as for the 'systematic approach'.

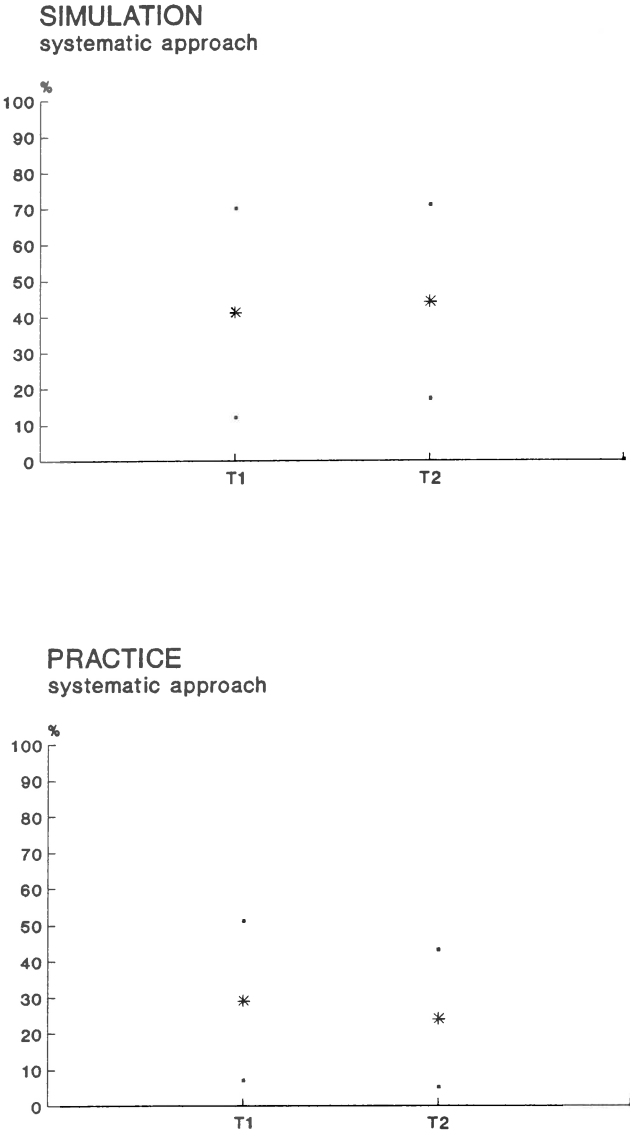
Lower scoring in practice was also found in each training group for all aspects except 'doctor-patient interaction'.

Figure 4 Mean percentage of the maximal score with standard deviation of all trainees (n=98)



t1 = after three months
t2 = after nine months vocational training

Figure 5 Mean percentage of the maximal score with standard deviation of all trainees (n=98)



t1 = after three months
t2 = after nine months vocational training

Discussion

The findings may be due to the difference between competence and performance (see the paper by Kuyvenhoven and Pieters in these proceedings). The simulation method is probably not a suitable method for measuring performance. During the simulation session the trainee is aware of the video registration and observation conditions and as a result will spare no effort. This so called audience or reactivity effect may be less in practice as the registration there lasts a whole week.

5. REPRODUCIBILITY OF SCORES WITH THE REVISED MAASTRICHT HISTORY-TAKING AND ADVICE CHECKLIST (MAAS-R)*

J. van Thiel, General Practitioner
H. Kraan, Psychiatrist
C. van der Vleuten, Educational Psychologist
University of Limburg
Department of General Practice
P.O. Box 616
6200 MD Maastricht
The Netherlands

Introduction

At the University of Limburg in Maastricht research is being done on the assessment of medical interviewing skills. The general goal in this research is the development of a relatively uncomplicated procedure for measuring medical interviewing skills of students at the medical faculty. The scores obtained with this instrument should have a high correlation coefficient of at least .80 between the scores of a group of students at 2 separate moments under the same conditions.

Medical interviewing skills are necessary to establish effective communication. Communication between doctor and patient is defined to be effective when both doctor and patient try to attune their mutual goals and each tries to understand the meaning the other attaches to the information exchanged. In this definition the emphasis is on the importance of effective communication. Key features are mutuality or reciprocity and awareness of the goal to reach.

In Maastricht an instrument has been developed for the measurement of medical interviewing skills by Kraan and Crijnen. They called it the Maastricht history-taking and advice checklist, shortly the MAAS. Based on their research and recommendations we revised the MAAS. Its revision was also based on four years of experience with the MAAS in our medical school.

* This paper is a short modification of research published in Medical Education; 1991, 25, p. 224-229

The revised version consists of a list of one hundred and twenty nine concrete behavioral items, covering an extensive domain of medical interviewing skills. The scoring is done checklist-wise: yes or no.

The list is structured in sections according to the phases reflecting the structure and chronology of a consultation. Before history-taking there are the sections entry, overall orientation, exploration of the reasons for encounter and diagnostic plan. Physical examination is excluded from the list because it is time-consuming and not necessary for measuring interviewing skills. The section history-taking is followed by the sections evaluation/diagnosis, management plan and evaluation of the consultation. A last section consists of skills relevant to all the other sections.

On the MAAS-Revised the scores were calculated on 2 scales. The first scale called Basic Skills contains items about the exploration of the reasons for encounter, items about structuring the consultation by planning, announcements and explanations. Behaviour that promotes reciprocity, such as inviting to react to the doctor's statements, exploring and reflecting emotions and giving proper summaries. Finally, items about the quality of giving information.

The second scale is on the level of the Medical Content. It includes items concerning medical questioning and medical information supply in the sections history-taking, evaluation/diagnosis and management plan.

The assumption underlying the 2 different scales is that behaviour concerning Medical Content is predominantly a reflection of knowledge or problem-solving skills and differs from 'Basic Skills' behaviour. Basic Skills are the essential skills to achieve effective communication. They form the frame work within which the Medical Content can take place. Strictly speaking they are the medical interviewing skills as described in our definition.

The research questions dealt with in this paper are:

1. How much testing time, or which number of cases and raters, is required to achieve a reproducibility of scores of at least 0.80 (generalizability coefficient) with the scale Basic Skills?
2. Are Basic Skills and Medical Content different domains of medical competence to be measured by using the correlation of scores of the two scales?

Method

The research material was based on consulting sessions done by twenty four doctors with eight standardized patients each. The doctors were trainees in general practice. Each patient portrayed one single case. From each consultation audio-recordings were obtained. These were scored by three raters; general practitioners trained in using the MAAS-Revised.

Item-analysis was performed with descriptive statistics and generalizability analysis. This resulted in removing items with no discrimination and items with poor reproducibility according to very high rater variance. The item-analysis reduced the amount of items of the MAAS-Revised from 129 to 74.

Final analysis was performed on sumscores for each case and the 2 scales, expressed in a percentage score. The mean score across cases determined the total test score for each doctor. Reproducibility of scores for each scale was assessed by calculating generalizability coefficients as a function of number of cases and raters used. Correlation between the total test score on Basic Skills and Medical Content was calculated by means of the Pearson formula.

Results

Figure 1 shows the results of the final analysis with regard to reproducibility of scores on the scales Basic Skills and Medical Content of the MAAS-Revised. Generalizability coefficients are given as a function of testing time and the number of cases and raters. The reproducibility of scores on Basic Skills is .79 with single ratings and 8 cases or a testing time of 2 hours. With 2 raters about 6 or 7 cases are needed for a reproducibility of .80. For the scale Medical Content the results are markedly lower. Nearly 7 hours of testing time is needed with 28 cases using a single rater per case, while five hours are needed when using two raters per case. This is mainly caused by higher rater variance and lower doctor variance.

Figure 1 Reproducibility of scores for examination purposes.

For each doctor each case is scored by a different rater.

Testing time in hours	Number of cases	Basic Skills		Medical Content	
		Number of raters		Number of raters	
		1	2	1	2
1	4	0.65	0.72	0.39	0.46
2	8	0.79	0.84	0.56	0.63
3	12	0.85	0.88	0.65	0.72
4	16	0.88	0.91	0.72	0.77
7	28	0.93	0.95	0.81	0.85

The correlation between scores on the scale Basic Skills and the scale Medical Content, is a low correlation of .31 and not significant ($p = .07$).

Discussion

The answer to the first research question is that measuring Basic Skills needs a testing time of 2 hours with 8 cases and one single rater to obtain a reproducibility of about .80. With 2 raters a testing time of one hour and a half and about 6 or 7 cases are needed. This result is encouraging. As far as the authors know, it is the best result up to now, with instruments measuring medical interviewing skills.

The second question can be answered 'Yes'; the low correlation supports the view that Basic Skills and Medical Content are different skills. Furthermore, the long testing time needed for Medical Content, in contrast to Basic Skills, suggests that a doctor is much less consistent in behaviour concerning Medical Content than in behaviour concerning Basic Skills. This also supports the idea that both types of behaviour are essentially different. The conclusion could be that it is sufficient to measure Basic Skills when measuring someone's medical interviewing skills in the strict sense of the used definition. In that case reproducibility could be improved upon, because raters would not have to rate Medical Content.

The lessons and results of this research were encouraging, and further improvement of the MAAS-Revised has resulted in the MAAS-Revised 2. However it still has to be proved that it is an improvement. That is why the MAAS-Revised 2 has been recently used in an experiment with 40 doctors, 4 cases and 2 raters per case. Again both Medical Content and Basic Skills have been included in the MAAS-Revised 2. If the results show a similar trend, Medical Content could be omitted in the list, when measuring medical interviewing skills in the strict sense of the used definition.

References

- DIMATTEO, M., D. DINICOLA. Achieving patient compliance. New York: Pergamon Press Inc., 1982
- EISENTHAL, S., A. LAZARE. Expression of patient's request in the initial interview. *Psychological Reports*; 40, 1977, p. 131-138
- ELSTEIN, A., L. SHULMAN, S. SPRAFKA. Medical problem solving. Cambridge: Harvard University Press, 1978
- FRASER, C. An analysis of face-to-face communication. In: *Communication between doctors and patients* (ed. A. Bennett) p. 5-29. Oxford: Oxford University Press, 1976

- KORSCH, B., V. NEGRETE. Doctor-patient communication. *Scientific American*; 227, 1972, p. 66-74
- KRAAN, H., A. CRIJNEN. The Maastricht History Taking and Advice Checklist: Studies of instrumental Utility. Amsterdam: Lundbeck, 1987
- PUTMAN, S., W. STILES, M. JACOB, S. JAMES. Patient exposition and physician explanation in initial medical interviews and outcomes of clinic visits. *Medical Care*; 23, 1985, p. 74-83
- SCHOUTEN, J. *Anamnese en advies*. Alphen a/d Rijn, Brussel: Stafleu, 1982
- SPRECHER, P., E. THOMAS, L. HUEBNER, B. NORFLEET, K. JACOBY. Effects of increased physician-patient communication on patient anxiety. *Professional Psychology: Research and Practice*; 14, 1983, p. 251-255
- STILLMAN P., P., M. REGAN, D. SWANSON. A diagnostic fourth year performance assessment. *Archives of Internal Medicine*; 147, 1987, p. 1981-1985
- SWANSON, D. A measurement framework for performance-based-test. In: *Further Developments in Assessing Clinical Competence* (Ed. I. Hart & R. Harden) p. 13-36. Montreal: Heal Publications, 1987
- TUCKETT, D. *Meetings between experts: An approach to sharing ideas in medical consultations*. London: Tavistock, 1985
- VLEUTEN, C. VAN DER, D. SWANSON. Assessment of clinical skills with standardized patients: state of the art. *Teaching and Learning in Medicine*; 2, 1990, p. 58-76

6. THE PATIENT SATISFACTION WITH COMMUNICATION CHECKLIST

By H.F. Kraan, MD, PhD
A. Crijnen, MD, PhD
J. van Thiel MD
J. van Dalen PhD
C. van der Vleuten, PhD
University of Limburg
Department of General Practice
P.O. Box 616
6200 MD Maastricht
The Netherlands

Introduction

In the past ten years extensive research has been done in evaluating medical interviewing skills at the university of Limburg (Maastricht).

Two instruments in this area were developed: the first is the MAAS, "the medical history taking and advice checklist", an observation instrument for initial medical interviews (Kraan, Crijnen, 1987). It was devised for expert evaluation, selfevaluation and as a summative test for medical interviewing skills. The second is the PSCC, "the patient satisfaction with communication checklist", which is an instrument for patients to evaluate the medical interviewing skills of their physicians. Both checklists have been revised since, by van Thiel, Kraan, van der Vleuten and van Dalen (Van Thiel, Kraan, Van der Vleuten, 1991).

The subject of this paper is the PSCC. A short motivation for the development of a new instrument for the patient's satisfaction with the communication will be followed by the development and testing by this PSCC.

There are several psychometric problems with the existing methods measuring patient satisfaction. These are the low variability due to the tendency of patients to consistingly give their physician a relatively high score in medical interviewing skills. The case specificity is also an impediment to the reliability of the instrument as it diminishes the stability of competence ratings. The most important problem is the lack of divergent validity, the dimensions measured by such instruments being rather non-specific. In general, they reflect general traits of likingness toward the physicians (for an overview of these problems, see (Crijnen, Kraan, 1987)).

The importance of a patient satisfaction checklist as an immediate outcome

measure of health care was one reason to try and develop an instrument that could surmount the above named short-comings. In this study we applied more advanced methodologies to the patient satisfaction with communication checklist. In this way we attempted to answer the question whether patients can reliably evaluate the medical interviewing skills of their physicians.

Method/construction of the PSCC

The first step in the development of the instrument was a literature search to find out which dimensions of the physicians interviewing style have a documented impact on the patient's satisfaction (Crijnen, Kraan, 1987). The most important ones will be named below:

- the ability of the physician to facilitate the patient in the telling of his or her story
- providing adequate information regarding diagnosis, prognosis etc.; giving it in small portions and in comprehensible terms
- fostering the patient's own insight in the problem
- enhancing patient compliance, bargaining with the patient, checking whether everything has been understood
- manifesting cognitive understanding i.e. the physician showing that he understands the patient's problem on a cognitive level
- showing empathy
- disrupting the communication by using jargon, making negative remarks, interrupting
- directivity, the degree to which the physician structures the interview

From these dimensions a 54 item checklist was constructed. 40 residents saw 4 patients each, resulting in 160 simulated patient contacts. On top of this, 7 practitioners were observed during a total of 117 real life first episode consultations. Of the total of 277 PSCC checklists 246 were available for analysis. Factor analysis yielded five interpretable dimensions related to the dimensions derived from the literature review. These were: facilitation, insight (the physician's endeavour to foster the insight of the patient), disruptive communication, communication intended to enhance the patient's compliance, and directivity. 65% of the total variance could be explained by these 5 dimensions.

As factor analyses give rather unstable data, additional analysis was done with the Rasch model to confirm these dimensions.

The Rasch model is a probabilistic scale model. It specifies a mathematical relationship (ogive-shaped logarithmic function) between scores on a psychometric scale (the PSCC) and an unobservable dimension (the physicians interviewing ability), the Rasch model requires, that items making up a pertinent

scale have equal discriminativity power (a similar shaped logarithmic function) and only differ in terms of intensity with respect to the underlying dimension (the physicians interviewing ability). When the scores on the items (of the raw version of the PSCC) fit the Rasch model, then these items constitute one single underlying dimension. These included items will only differ in the level of measuring this underlying dimension.

So the original question can be restated as follows; when PSCC-items (after selection) fit in the Rasch model, patients reliably can evaluate the underlying dimension i.a. their physicians interviewing skills. Several items per dimension (obtained from factor analysis) fitted in the Rasch model resulting in a total of 19 items left over from the original checklist of 54 (Addendum).

The first results were promising, given the fact that Rasch analysis is a demanding model yielding internal consistent (evaluation) scales, that also provides a grading in difficulty of the items. However, further analysis revealed less optimistic results especially after investigating how the specificity of the case characteristics and the way of presentation by the patients influence the interviewing skills of the doctor, and consequently its evaluation by patients. A generalizability theory (Mitchell, 1979; Thorndike, 1982) assumes that test scores are the results of a number of variance sources, like subjects characteristics, item domains, raters reliability and test conditions, the scores of variance make up the 'universe' to which test scores may be generalized. By means of analysis of variance estimations are made about the contribution of each source to the overall variance and the size of the different variance component can be established. According to this theory analysis was done to study the influence of the case presentation and the content of the case on the evaluation of the doctor-patient communication, i.e. the case specificity. In general, the scores of the PSCC will result from a number of sources such as the doctors interviewing ability, patient attributes, and the characteristics of the medical problem (case). The scores on the PSCC also depend on the patients way of using item and scaling format. With analysis of variance of a completely crossed design (each physician interviewing the same simulated patients) the contribution of these sources to the final score can be estimated.

In the presented generalization analysis the physician interviewing ability and the case influence were selected as the sources of variation. The influence of the combination of these sources is expressed in the generalizability coefficient which can be taken as a composite measure of reliability (Thorndike, 1982).

Results/discussion

In Figure 1 the results of this analysis and the generalizability coefficients are shown.

Figure 1 Results of the generalizability analysis of the PSCC

effects	variance components (%)	
	original PSCC	revised PSCC
physician's interviewing ability	0%	3%
cases	14%	31%
ability cases (+ error)	86%	66%
generalizability coefficient:	4 cases: 0.00	4 cases : 0.14 8 cases : 0.25 16 cases : 0.40 32 cases : 0.57

The first column (left side) shows the variance components as a percentage of the total variance, as found for the original version of the PSCC. The physician variance is zero, indicating that according to the patients scores there is no variance in the physician's interviewing style. The influence of the cases is 14% and the interaction of the physician and the cases is 86% including an error term. These results suggest a strong influence of the case (characteristics and as presentation by the patient) on the evaluation of interviewing skills by patients. No generalizability coefficient could be obtained. After the good results on Rasch analysis it was very disappointing to find the unacceptable low generalizability coefficient.

The PSCC was revised a few years later. The patient instructions were also slightly changed. Only simulated patients (N=8) were used. Factor and Rasch analysis showed almost the same results as the first time. In factor analysis the same dimensions were found, which in Rasch analysis were confirmed. The generalizability analysis (Figure, column on the right side) showed a marginally higher doctor's ability variance of 3% instead of 0%. The generalizability coefficient was very low for 4 and 8 cases, even with 32 cases it reached .57 which is still unacceptable. These results show, that for real patients 32 consultations are not sufficient to rule out the effect of case-specificity. As to test-situations with simulation patients, more than 32 of them are needed to provide a reliable evaluation.

These results are paradoxical: on the one hand the fit of some evaluative dimensions of doctors interviewing ability in the Rasch model and on the other hand the effects of case specificity leading to low reliability. This paradox points to a validity issue: Rasch analysis and generalizability analysis may focus on different aspects of the complex phenomenon, where patients evaluate the interviewing skills of their doctors (that are influenced by themselves!). Rasch analysis may focus on more stable, general, trait-like dimensions of the

Physician's interviewing skills, like facilitation patient-centredness etc. We, exclusively, show the effect of case-specificity in relation to the physician's more concrete and detailed interviewing skills (as indicated in the PSCC items). Application of these skills and their evaluation by patients are too strongly dependent of the problems presented by the patients themselves.

Conclusions

The patient's evaluation of the facilitative, directive and disruptive dimensions in their physician interviewing style look rather stable. To a lesser extent patients reliably indicate if they will be a compliant patient with the physician (intention to comply) and if they have been given sufficient insight in their problems (insight). However, generalizability analysis proved, the reliability of these dimensions is low because of case specificity. To overcome the case specificity more than 30 consultations would be necessary to entail a reliable evaluation of physician's interviewing skills. The results with the Rasch analysis suggest, that the PSCC evaluates more stable, general characteristics of the physician, like being facilitative, directive and (non-)disruptive, reflecting his/her patient-centredness. However, reliable evaluation of these characteristics, and in particular of the giving of insight and of the patient's intention to comply is hampered by case-specificity. Validity studies are needed to assess which dimensions of the physician-patient relationship the PSCC really measures. Altogether the construction and use of such instruments as the PSCC for measuring the physician's medical interviewing skills as seen through the eyes of the patients turns out to be a rather disappointing endeavour and is most probably a dead end street, because only general, trait-like physician's characteristics are evaluated instead of the more detailed interviewing skills.

References

- CRIJNEN, A.A.M., H.F. KRAAN. Measuring patient satisfaction with communication. In: Kraan H.F., A.A.M. Crijnen. *The Maastricht History-taking and Advice Checklist*. Amsterdam: Lundbeck fond., 1987
- HAMBLETON, R.L., L.L. COOK. Latent trait models and their use in the analysis of educational test. *Journal Educat. Measur.*; 1977, 14, p. 95-96
- KRAAN, H.F., A.A.M. CRIJNEN. *The Maastricht history-taking and Advice Checklist (MAAS)*. Amsterdam: Lundbeck fond., 1987
- MITCHELL, S.K. Interobserver agreement, reliability and generalizability of data collected in observational studies. *Psychol. Bull.*; 1979, 86, p. 376-390
- THORNDIKE, R.L. *Applied psychometrics*. Boston: Houghton Mufflin cie, 1982

VAN THIEL, J., H.F. KRAAN, C.P.M. VAN DER VLEUTEN. Reliability and feasibility of measuring medical interviewing skills: the revised Maastricht History-taking and Advice checklist. *Med. Educ.*; 1991, 25, p. 224-229

WRIGHT, B.D. Solving measurement problems with the Rasch model. *Journal Educat. Measur.*; 1977, 14, p. 97-116

Addendum

Patient satisfaction with communication checklist Original version (7-point likert scale)

Facilitation

1. The physician gave me the opportunity to tell my own story.
2. The physician enabled me to talk friendly.
3. I could ask anything I wanted to.

Insight

4. I know the pros and cons of my treatment.
5. The physician clarified the meaning of my conduct.
6. The physician explained side-effects of my medication.
7. The physician gave me new insight into my problem, enabling me to cope with it.
8. The impact of my problem on people close to me became clear.

Intention to comply

9. I'm able to recall the agreements which I reached with the physician.
10. I will certainly cooperate with the proposed treatment.
11. The physician checked whether I had understood the advice and further information.
12. The physician gave me much responsibility in the choice of my treatment.

Disruption of communication

13. The physician used incomprehensible jargon.
14. The physician switched too fast from one subject to another.
15. I was made anxious and uneasy by the questions the physician asked.
16. Sometimes the physician reexpressed my statements differently from my original meaning.

Directivity

17. Sometimes the physician invited me to talk about subjects which I would rather not have discussed at that moment.
18. Sometimes the physician would have liked to take a decision in my place.
19. On some aspects of my life, the physician wanted me to adopt a view that was different from my own.

Revised Version (7-point likert scale)

In this version the items 7 and 8 of the original one are removed, they are replaced by the two following items, that constitute the subscale 'reassurance' in combination with item 15 of the original version:

- My anxiety decreased after the consultation.
- The physician reassured me.

7. LACONTO: DEVELOPMENT OF A NATIONAL INSTRUMENT TO ASSESS CONSULTATION SKILLS OF FUTURE GENERAL PRACTITIONERS

Dr. L. Tan
C.H.G.M. Foolen
Dr. C.P.M. van der Vleuten
Department of General Practice
University of Utrecht
Bijlhouwerstraat 6
3511 ZC Utrecht
The Netherlands

Introduction

In the Netherlands the education of general practitioners consists of six years of medical school (including clerkships) followed by two years of vocational training. The latter part is based on the Basic Job Description for the general practitioner which has been formulated by the profession.

Vocational training is organized by the departments of General Practice at the eight medical schools and is made up of two components. During four days of the week the trainees work either in the practice of a general practitioner trainer or in a clinical setting. One day a week teaching takes place in groups of twelve trainees at the respective departments of General Practice.

Since 1987 the eight departments of General Practice in the Netherlands co-operate in the evaluation of their vocational training. This co-operation is supported by the administrative office in Utrecht, where the author works.

The national evaluation system in vocational training for general practitioners comprises the evaluation of teaching activities and the assessment of knowledge and skills of the trainees. By skills both technical skills and consultation skills are meant.

National assessment extends to subject matter on which consensus has been reached between the various parties involved. Furthermore the content of the instruments used in national assessment procedures is independent of any specific educational programme of the eight departments. It may be said that the process of developing a national evaluation system stimulates the creation of consensus on the core of General Practice as a specialty.

The construction of the national instrument

For the actual construction of the national instrument a list of available instrument measuring consultation skills in medical education and in vocational training was made. These were:

- PREVARA-SETH made in Nijmegen to which amongst others Vic Tielens and Henk Mokka contributed considerably;
- the UCEM made in Utrecht with Ron Pieters as researcher;
- the MAAS-R made in Maastricht on which amongst others Jacques van Thiel and Herro Kraan work as researchers.

In the Basic Job Description for the general practitioner consultation skills are subdivided into four dimensions:

- 1 systematic behavior;
- 2 communicative behavior;
- 3 case specific medical performance in the somatic domain;
- 4 case specific medical performance in the psycho-social domain.

All three available instruments agree on the first two dimensions 'systematic behavior' and 'communicative behavior'. However, they disagree on the dimension regarding case specific medical performance: the Maas-R does not aim at medical content, the UCEM deals slightly with it, only the PREVARA-SETH was designed to assess all four dimensions through one instrument.

As the national evaluation system was to cover subject matter on which consensus had been reached, the two dimensions chosen were systematic behavior in General Practice and communicative behavior. In concurrence with the international literature in this field, it was decided that testing of case specific medical performance should be done by means of a separate assessment tool.

The first draft of the national instrument was composed of the elements on which the three locally developed instruments agree, it thus heavily leant on the work of others. This draft was critically commented on by 27 experts from all the eight departments of General Practice. The experts were general practitioners and behavioural scientists in a fifty/fifty ratio. Their comments were given in two consensus rounds using an extensive questionnaire. The response was 96% in the first round and 67% in the second round. The comments given resulted in radical changes in the first drafts of the instrument.

The ultimate format of the experimental instrument is a one page itemlist of thirty items with a separate booklet of explanatory notes. The thirty items are to be scored on a six-point scale which is a combination of a checklist and a rating scale. The checklist is used to indicate whether an item 'was' or 'was not' performed. The rating scale is used to rate the items performed and goes from 'very badly' to 'very well performed'.

The itemlist concerns the two dimensions 'systematic behavior' and 'communicative behavior'. The thirty items are divided into the three phases which can be distinguished in the consultation process: intake, diagnosis and management. The last two items are meant for giving an overall impression on the two dimensions. The list is shown in Figure 1.

Items printed in italics refer to the dimension on communicative behavior. The rest of the items concern the systematic behavior.

Figure 1 Itemlist structured working and method doctor-patient communication

	no	very bad				very good
Intake	0	1	2	3	4	5
1. greets, opens consultation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. <i>has an inviting attitude</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. lists complaints/problems	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. clarifies reasons for visit	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. <i>applies appropriate communication technique</i>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. explores treatment till now (incl. selfhelp)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Pilot study

Method

The goals of the experimental testing of the instrument were the following: to find out about the reliability, validity and feasibility of the newly constructed instrument.

Of the forty general practitioner trainees that took place in the testing, twenty had followed four months of vocational training. They were called the 'beginners'. The rest of them had followed sixteen months of training: 'advanced trainees'. The trainees were asked to do four consultations lasting ten minutes each at the most. All patients were simulated patients. The consultations were video-recorded.

The scoring on the two dimensions was done by eight general practitioners and eight behavioural scientists. This scoring took place after the consultations using the video-taped version of the consultations.

The scoring design for the national instrument was a 'raters-within-cases' design in which each of the four cases was scored by two pairs of raters of the same discipline (general practitioners or behavioural scientists). For instance: case 1 was scored by the raters A1 and A2, both behavioural scientists, and by the raters B1 and B2, both general practitioners. The first pair scored the beginners and the second pair rated the advanced trainees. They did the rating independently and they were not informed about the level of the trainees: beginner or advanced.

The trainees got feedback about their own score on the two dimensions. They were also informed about the mean scores of their peers: their training group as well as their year group.

The simulated patients filled out a six-item form immediately after the consultation. This concerned the way they felt treated by the trainee. This feedback was given to the trainees individually.

Results

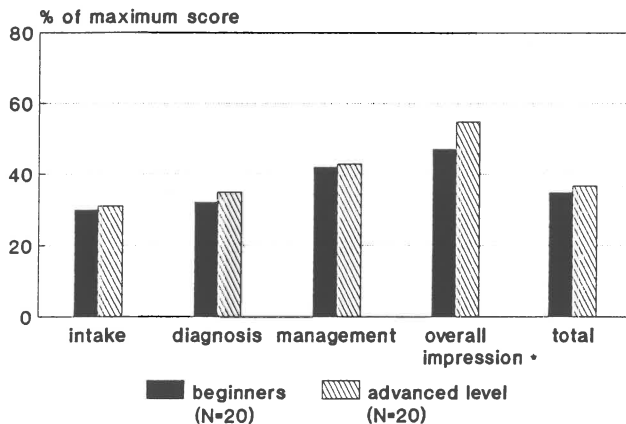
Figure 2 shows the mean scores on the various phases of the consultation. The black bars represent the beginners, the shaded ones the advanced trainees.

Excluding the overall impression, the mean scores never exceed 45% of the maximum score. This applies as well to trainees at the beginners level as trainees at the advanced level. The mean differences between the beginners and the advanced trainees point in the expected direction (the advanced level being higher than the beginners level). However, in this first experimental testing only the differences for the overall impression were found to be statistically significant.

The inter-rater reliability was as follows: .29 for general practitioner raters and .44 for behavioural scientists. The inter-rater reliability for the total test was .36.

By means of the generalizability analysis it was estimated that a reliability of .80 could be reached using two raters and 16 cases per trainee. In a single rater design, a general practitioner rater would need 24 cases per trainee to reach a sufficient reliability whereas a behavioural scientist would need 12 cases per trainee.

Figure 2 Mean scores per year group (N=40)



* $p < .01$

Conclusions

Compared to other testing in the field, the *inter-rater reliability* between general practitioners is rather 'low'. However, the reliability among the behavioural scientists approximates the values usually found in testing consultation skills.

To reach a *total test reliability* of .80 using one behavioural scientist rater in a test situation one needs 12 cases per candidate.

The consensus method used in constructing the national instrument produced a high *content validity*. As to the *construct validity*: the statistical testing produced only significant differences between beginners and advanced trainees in the case of the overall impression. Although the empirical differences point in the expected direction: the advanced level is higher than the beginners level.

The *feasibility* of the construction and testing of the instrument turned out to be quite good. The eight departments of General Practice not only co-operated in discussing the subject matter thoroughly but also co-operated in supplying raters and trainees needed for the testing. Furthermore the organization of the centralized data processing and feedback of test results was feasible.

A point to be mentioned in this context is the opinion of the trainees about the testing. All trainees felt that the simulated cases were relevant to general practitioner practice. Seventy five percent of the trainees stated that they would support regular testing in consultation skills.

The problems met during the course of the project were the following. The organization of a centralized training for sixteen raters coming from all over the country was rather painstaking, even though the training took as little as four hours.

Another problem was set by the case specific content of the testing. To join the important process of setting professional standards relevant to general practice, the cases were based on four standards published recently by the profession. The trainees knew about this and could prepare themselves. Yet in the test situation they seemed to focus rather on medical content than on the process of consultation. This was obviously not optimal for research on the quality of an instrument focusing on the process of consultation.

The third problem arising was that of 'simultaneous testing and marketing'. Two of the conditions for the development and implementation of an evaluation system were: high quality instruments and readiness amongst the subjects to be tested. To establish this readiness, in other words for marketing reasons, the instrument was tested at two departments of general practice without a tradition in instructing and testing trainees by means of an observation system. This fact may have contributed to the low discriminatory power which the instrument showed in this first experimental testing.

Recommendations

The itemlist and the explanatory notes should be corrected on the basis of item analysis and the critical comments given during the testing by raters and trainees.

Both trainees and raters should have the opportunity to get experienced with the observation system.

The raters' training should be intensified and extended. Training should be limited to a relatively small central board of raters.

Furthermore testing with the nationally developed instrument should be organized on local scale.

Finally, further research is necessary to guarantee a high quality of the national assessment instrument on consultation skills.

8. LACONTO IN EVERYDAY GENERAL PRACTICE - a multisystem comparison

J.M. Bensing
L.H.C. Tan
S.A. de Boer
S.S.L. Mol
Netherlands Institute of Primary Health Care
P.O. Box 1568
3500 BN UTRECHT
The Netherlands

Introduction

In the Netherlands much attention is paid by the University Departments of General Practice to the development of instruments to measure consultation skills among general practice trainees. Examples, described in this book are the Utrecht Consultation Assessment Method (UCAM) of the Utrecht University (Chapter 4), the Maastricht History-taking and Advice Checklist (MAAS) of the Limburg University (Chapter 5) and PREVARA, developed by the Nijmegen University (Chapter 8). Besides these locally developed instruments a national assessment instrument on consultation skills is being developed by the joint Dutch University Departments of General Practice: LACONTO (Chapter 6). This instrument is based on the locally developed instruments and is aimed at standardizing the assessment of GP-trainees all over the Netherlands.

In this paper LACONTO will be tested in every-day general practice. The testing consists of a reliability-test and a validity-test. Besides LACONTO's performance will be compared with the performance of some other observation systems that are commonly used on real-life consultations on the same set of consultations: NIVEL's observation system and Roter's Interaction Analysis System (RIAS). Some preliminary conclusions will be drawn at the end.

Methods

This study's sample consists of 103 real-life consultations with hypertensive patients. The doctors are all experienced general practitioners. There are some important differences between this sample and the sample used to test LACONTO (see Chapter ...). In the latter test-sample GP-trainees (beginners and advanced trainees) were assessed while working - in a laboratory-like

situation - with patients simulating four different types of complaints. When discussing the results of both studies these differences must be kept in mind. In the past, the 103 hypertension consultations had been assessed in 3 different ways. First of all the quality of the consultations was assessed by 12 experienced general practitioners independently. The quality assessment was given for three different domains of the GP's health care (technical-medical care, psychosocial care and doctor-patient relationship); the format was a 10-point rating-scale. The two other observation systems used were the NIVEL-system and Debra Roter's Interaction Analysis System (RIAS). Now the same video consultations were scored using the Dutch National Assessment System on consultation skills (LACONTO). With the latter the scoring was done by two observers, one general practitioner and one psychologist. This paper will concentrate on the last system. Information about the other two systems can be found elsewhere (Bensing, 1991).

Results

LACONTO's reliability

As can be read in Lisa Tan's paper, LACONTO had some reliability problems, especially among the general practitioner observers. In this study the observers represented two different disciplines, which might be an additional source of error with detrimental effects on the reliability.

Before starting the reliability-studies some preliminary work was done. Three items were removed from the observation system as they were not filled out in any of the 103 consultations. Whatever is learned in medical education, Dutch general practitioners do not announce to their patients that they want to discuss the plan of management; they don't ask the patient's permission to do so, and they don't end the consultation with an evaluation. This finding is in itself worth a discussion about the necessity of these items in the instrument.

The remaining items were grouped on the analogy of LACONTO's structure: **two types** of behavior: systematic working and communicative behavior in the **three phases** of the consultation: the intake phase, the diagnostic process and the management phase. This makes 6 components.

The interobserver-reliability figures of the individual items of the observation-system show a wide variability. Four items have a zero or non-significant correlation (Pearson's Product-Moment correlation): diagnostic 14, clarification of physical examination; diagnostic 17, patient centered presentation of findings and conclusions; management 22, discussion of possible outcome and feasibility of the management plan; management 25, patient-centered presenta-

tion of management plan. Generally the general practitioner was more positive about the delivered care than the psychologist. One should note that all these items point to the way the doctor's words are attuned to the patient's level of understanding. Perhaps it is more difficult for a general practitioner to step into the patient's shoes than for a psychologist, who is not burdened by the general practitioner's medical knowledge.

Figure 1 shows the mean and ranges of the inter-observer-reliability-figures for the individual items within each component of the observation-system. Most item-reliabilities are too low to enable us to use the individual items in the analysis.

Figure 1 Inter-observer reliability on item- and component level (Pearson's P.M. correlation)*

			Reliability on item-level			Reliability sumscore
			Lowest	highest	mean	
o Intake	{	s	.27	.76	.52	.63
		c	.31	.57	.48	.62
o Diagnosis	{	s	.21	.61	.38	.53
		c	.05	.59	.30	.56
o Management	{	s	.21	.60	.40	.62
		c	.04	.30	.17	.26

* s = systematic behavior
c = communicative behavior

Though the reliability might seem disappointing the components' reliabilities shown in the last column being much higher than the item-reliabilities, is certainly acceptable for further research. Communicative behavior during the management phase is a notable exception. The reliability-figures of the sum-scores for systematic working, respectively doctor-patient communication as well as the Grand Total (see figure 2) are still even higher, and reach levels to be proud of. The reliability of the overall assessment of the two types of behavior are a bit lower, again. Yet, because of the reasonable intra-observer-reliability-figures on a subset of 13 consultations (figures in brackets) these overall assessments are also included in the analyses.

Figure 2 Inter-observer reliability on consultation level (Pearson's P.M. correlation)

I	Sumscores		
*	systematic behavior	.73	
*	communicative behavior	.70	
*	total consultation	.75	
II	Overall impression		
*	systematic behavior	.32	(.45)
*	communicative behavior	.49	(.74)

(between brackets: intra-observer reliability - test-retest)

LACONTO's construct validity

As is shown in Lisa Tan's paper, the construct validity of the NAS-cs was initially tested by comparing the results of the beginners and the advanced trainees. The differences all pointed in the same direction but they were quite small. Only the overall assessment proved to be higher among the advanced trainees, as compared to the beginners.

In this study the construct validity is tested by relating the results on LACONTO to panel-assessed quality-ratings on three different dimensions of the GP's health care: technical-medical care, psychosocial care and doctor-patient relationship. Consultations which were assessed as "insufficient" by the panel of experienced general practitioners (these were consultations that had received a rating under 6) were compared with the consultations that had received a "good" panel-rating (a rating over 7). To enable a comparison with the previous study, the scores were computed in exactly the same way.

Figure 3 shows that there are significant and large differences between consultations which were assessed to be of insufficient medical quality and those with a high medical quality. The same is true for the quality of psychosocial care (figure 4). With regard to the quality of the doctor-patient relationship (figure 5), the differences between the two groups are smaller, and not significant for the management phase.

Figure 3 Technical-medical quality; mean scores per consultation phase

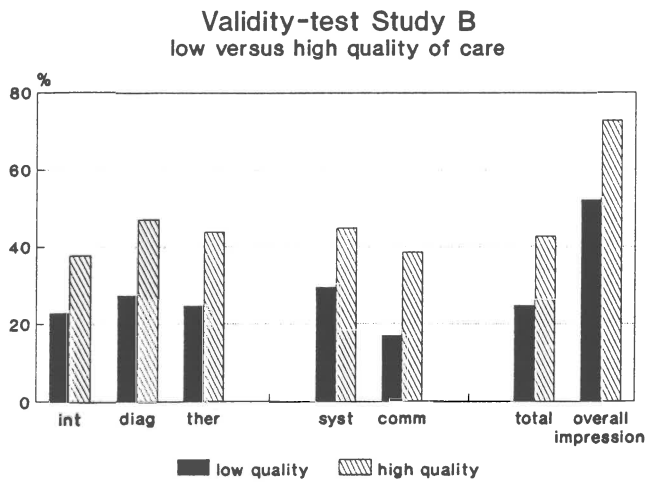


Figure 4 Psychosocial quality; mean scores per consultation phase

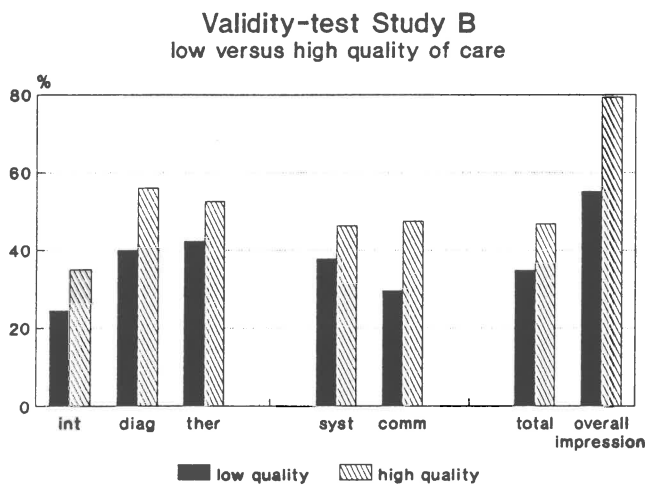
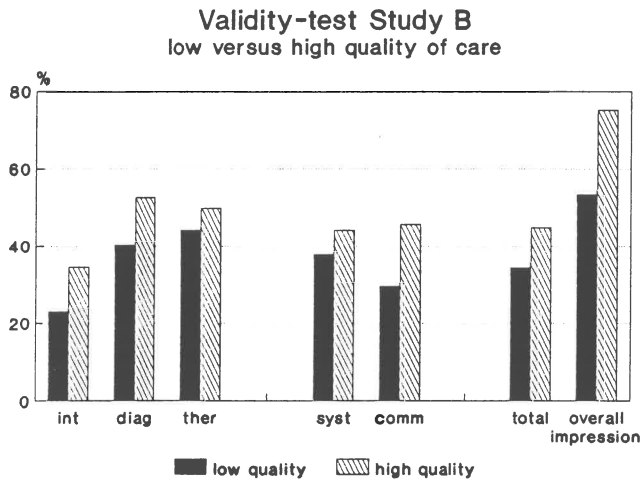


Figure 5 Doctor-patient relationship; mean scores per consultation phase



In summary: with the exception of systematic behavior during the intake phase all components of LACONTO do show differences between consultations with good quality of technical-medical care and consultations with insufficient quality of medical care (figure 6). For the psychosocial domain, the same result is found; In the domain of the doctor-patient relationship, only three of the components are able to distinguish between consultations of good, respectively insufficient quality of care.

Moreover, we see that systematic behavior is especially important in the quality of the diagnostic phase of the consultation, and furthermore especially in determining the technical-medical quality of care, while communicative behavior seems important in all three domains of GP's health care, least in the management phase. Here one should keep in mind that this was the only component that did not reach an acceptable reliability level.

An overall conclusion can be that in this study, the construct validity of the NAS-cs proves to be rather high, in fact higher than in Lisa Tan's original test-sample. A possible explanation is the fact that the trainees have acquired most of their communication skills in the first period of their vocational training, before the first test took place (see also Pieter's paper). Another possibility is that the differences can be ascribed to situational factors: an artificial test-situation with simulated patients against every day general practice and real life patients.

Figure 6 Summary: discriminative power of the 6 components of LACONTO

		technical- medical quality	psycho- social quality	quality doctor-patient relationship
i n t a k e	systematic	n.s.	n.s.	n.s.
	communicative	***	***	***
d i a g n o s i s	systematic	***	***	***
	communicative	**	***	**
m a n a g e m e n t	systematic	***	*	n.s.
	communicative	***	***	n.s.

n.s. = no difference between low and high quality

* = p < .05 between low and high quality

** = p < .01 between low and high quality

*** = p < .001 between low and high quality

Comparison with other observation systems

How do the results of this study relate to previous studies, in which the NIVEL observation system and Roter's Interaction Analysis System are applied to the same 103 consultations?

It is possible to get an idea of the explanatory power of an observation system by applying a multiple regression analysis with the system's elements or components as independent variables and a relevant criterium as a dependent variable. Figure 7 summarizes the results of these analyses for the three observation-systems and the three panel-assessed quality ratings. LACONTO

performs somewhat better than the NIVEL-system with regard to the technical-medical quality; the NIVEL-system performs a little better in the psychosocial domain. Roter's system is superior to both, but that can mainly be ascribed to the high explanatory power of its four global affect-ratings. Altogether, notwithstanding small differences all three systems are performing rather well, which means that they represent relevant types of doctor's behavior.

Figure 7 % Explained Variance

	Medical quality	Psychosocial quality	Doctor-patient relationship
LACONTO	39%	39%	35%
NIVEL	34%	58%	40%
ROTER	60%	70%	59%

The most important differences between the various observation systems are:

- Roter's system is comprehensive: each and every verbal unit of speech is scored in one of the mutually exclusive categories. LACONTO is selective: only those behaviors that are thought relevant for a good consultation are rated. It is based on the "official" Task-Description of the Dutch general practitioner, so it is assumed that each behavior that is thought necessary for a good consultation is rated within LACONTO. The NIVEL-system is also a selective system, but it is a rather opportunistic system; it does not try to cover all behavior, but tries to find the really relevant behaviors. It ignores most of what the general practitioner does.
- In Roter's system just as in the NIVEL-system, instrumental behavior is measured in its verbal form only; this behavior is counted, rather than rated. In LACONTO instrumental behavior is rated from an undifferentiated mix of verbal and nonverbal cues. In this system, affective behavior is measured in the same way. Roter's as well as the NIVEL-system do have different measures for the verbal and nonverbal aspects of affective behavior, which are partly based on counting and partly on rating.

Counting or rating, verbal or nonverbal comprehensive or selective; the observation-systems are very different from each other in several respects. Yet, the results show some striking similarities. Affective behavior seems more important than instrumental behavior, except perhaps in some parts of the technical-medical domain. And global ratings seem to explain more of the consultation's quality assessment than detailed observations, especially in the affective

domain of doctor-patient communication.

The importance of global assessments in explaining the variance in the consultation's quality of care can be seen when all global assessments of all three observation systems are put together in one multiple regression analysis. (Figure 8)

NAS-cs: the overall assessment of the general practitioner's systematic behavior and communicative behavior;

Roter's: global affect ratings of the general practitioner's interest in the patient, his warmth, his irritation and nervousness;

NIVEL: as assessment of the general practitioner's patient-centredness in the diagnostic and therapeutic phase of the consultation; both from the NIVEL-system.

Together these 8 global assessments explain an enormous amount of the variance in panel-assessed quality of care, 51% of the variance in technical-medical quality; 72% of the variance in psychosocial quality; and 67% of the variance in the quality of the doctor-patient relationship.

Figure 8 Relevant behaviors in determining the quality of care

	medical quality	psychosocial quality	doctor-patient relationship
LACONTO	* intake-c * diagnosis-c * diagnosis-s * management-c * management-s	* intake-c * diagnosis-c * diagnosis-s	* intake-c * diagnosis-c * diagnosis-s
RIAS	* medical info * verbal attentiveness * directions	* medical info * verbal attentiveness * psychosocial questions	* medical info * verbal attentiveness * social behavior
NIVEL	* % eye-contact	* % eye-contact * verbal affective beh.	* % eye-contact * verbal affective beh.

Discussion

First, it seems that the results of observation research are less determined by the choice of the observation system as such, than by the study's methodology: the sample of subjects, the use of simulation patients versus real life patients, the choice of the particular health problem studies, and - perhaps - the use of audio versus video. After having observed the 103 hypertension consultations according to three different observation-systems, it can be concluded that in these consultations affective behavior is more important than instrumental behavior, not exclusively in the non-medical domains of care. The right amount of medical information, forwarded in a patient-centered way is an important additional element of high-quality care. Furthermore, psychosocial care needs psychotherapeutic forms of behavior like Roter's verbal attentiveness and global affect-ratings, NIVEL's verbal and nonverbal affective behavior, and LACONTO's communicative behavior. It is not to say that these results will be found in other studies too, studies with different subjects, different health problems, or a different methodology. This could be tested in further research. The second point to be discussed is the cost-effectivity of observation research. In this study a very high proportion of the consultation's quality of care can be explained by only a few global assessments on relevant elements of general practitioner's behavior: affective behavior, patient-centredness, systematic behavior and communicative behavior. These global assessments can be made within minutes after viewing the videotaped consultation. It costs a lot more time to use the detailed observation systems, the LACONTO being the least time-consuming one. The profit for making more detailed observations is rather low: 7% more explained variance in the technical-medical quality of the consultation, 10% additional explained variance in the psychosocial quality, and only 1% of the variance in the quality of the doctor-patient relationship. Is this difference worth the effort?

One possible answer is that knowing these global assessments can be of little help in medical education: it does not show which types of behavior a particular trainee should learn or unlearn. A second answer has to do with the need for more knowledge about the ingredients of those global assessments: when do we call a doctor 'interested', or 'patient-centered'? Using detailed observation systems can teach more about these complex but highly relevant questions. Fundamental research with detailed observation systems is necessary to provide these answers. The conclusion is that different purposes can and must lead to the use of different observation systems, which can differ widely in the 'globality' of measures. One should be aware of the cost-effectiveness of the observation-system one chooses.

References

- BENSING, J.M. Doctor patient communication and the quality of care. *Social Science & Medicine*; 32, 1991, No. 11, p. 1301-1310
- BENSING, J.M. Doctor-Patient Communication and the Quality of Care. An observation study into affective and instrumental behavior in general practice (thesis). Utrecht: NIVEL, 1991
- BENSING, J.M., J. DRONKERS. Instrumental and affective aspects of physician behavior. *Medical Care*; 30, 1992, No. 4, p. 283-298
- PIETERS, R.M. Utrecht Consultation Assessment Method (UCAM). Reliability and validity of actor simulation in assessing consultation skills of trainees in the vocational training. Construction and validation of a measurement instrument. Utrecht: Invitational workshop on Doctor-patient communication and the quality of care, 1991
- ROTER, D.L. The Roter Method of Interaction Process Analysis. Baltimore: The Johns Hopkins University, 1989
- TAN, L.H.V., C.A.G.M. FOOLEN, G. VAN GELDROEP. Towards a national assessment of skills in the vocational training for general practice. In: Scherpbier, A.J.J.A., W. Bender, R. Zwierstra (eds.) Proceedings 3e Ottawa Conference on assessing clinical competence - Groningen, Boekwerk 1989
- THIEL, J., H.F. KRAAN, C.P.M. VAN DER VLEUTEN. Reliability and feasibility in measuring medical interviewing skills with the revised Maastricht history-taking and advice checklist (MAAS-R). *Medical Education*; 1991, 25, 224-229
- THIEL, VAN J., H. KRAAN, C. VAN DER VLEUTEN. Reproducibility of scores with the revised Maastricht history-taking and advice checklist (MAAS-R). Utrecht: Invitational workshop on Doctor-patient communication and the quality of care, 1991
- TIELENS, V.C.L. Consultation length and the patient's view: aspects of quality. Utrecht: Invitational workshop on Doctor-patient communication and the quality of care, 1991

9. CONSULTATION LENGTH AND THE PATIENT'S VIEW: ASPECTS OF QUALITY

V.C.L. Tielens
Nijmeegs Universitair Huisartsen Instituut
P.O. Box 9101
6500 HB NIJMEGEN
The Netherlands

Introduction

The length of the general practitioners' consultation is often associated with the quality of the consultation. From rather considerable research into the action of general practitioners recommendations emerge that the general practitioner must devote more time and attention to a multiplicity of subjects that are concerned with consultation practice. For their general practitioners often respond to this by saying: that is all very well, but I haven't got the time; if I had to do all that, my surgery would be never-ending. Through this discrepancy many general practitioners think that they are remiss and they feel guilty. Certainly not the worst general practitioners are rather troubled by this. They also often think that patients expect must more time from them than they can give.

To get a better idea about the complex connection between consultation length, quality of general practitioners' action and the expectations and satisfaction of patients, a study was set up at the Nijmegen University Institute of General Practice (Schellekens, Mokkink and Tielens). In that study the length of the consultation was studied in relation to some characteristics of the doctor, the patient, the complaint and the practice. The two questions formulated in the study were:

- what is the connection between the length of the consultation and the quality of the general practitioners' action?
- what is the connection between the length of the consultation and the patient's satisfaction?

Method

For this investigation use was made of data collected in an extensive observation study of surgeries of general practitioners, performed by the Nijmegen University Institute of General Practice. The general practitioners had to have been working as general practitioners for at least five years and to have at least

1000 insured persons in the health insurance fund. Out of the 181 general practitioners who satisfied these criteria a sample of 75 general practitioners was taken, stratified by degree of urbanization (urban versus rural areas) and referrals (much versus little).

A questionnaire was submitted to these general practitioners, audio-recordings were made of an arbitrary surgery and observations were performed by specially trained observers. Action that was not audible on tape, such as physical examination, was recorded on registration forms. The general practitioner's action was assessed by means of the PREVARA observation instrument. This instrument consisting of a medico-technical and a behavioral-science part, which makes it possible to assess the four basic skills of the general practitioner: systematic behavior, managing the doctor-patient relationship, adequate somatic action and adequate psycho-social action. Per general practitioner at least 15 consultations were observed.

The consultation length was regarded the dependent variable. The consultation length was rounded to the nearest half-minute. The values for the average length were expressed in one decimal.

The patient data was obtained by giving all Dutch speaking patients older than 15 years a questionnaire to take along. This enquired after such matters as general state of health, highest educational level attained, the frequency of contact during the preceding two months and the reason for coming, whereby the patient could choose one or more of the following nine reasons:

- worries about the complaints
- trouble and hindrance from the complaints
- wants to know what is going on
- wants to talk about problems or worries
- check-up on complaint, result and examination
- wants to be thoroughly examined
- wants explanation of/advice
- wants medicine
- wants to be referred.

Regarding their satisfaction to the consultation, the patients could answer the following four questions on a five-point scale:

- did you feel at ease during this visit to the doctor?
- did the doctor give you the chance to tell and ask about what was on your mind?
- was it clear to you what the doctor told you?
- did the doctor take enough time for you?

The patients were further asked to indicate to what extent the patients found that the doctor had devoted sufficient attention to the underlying reason for their visit.

From the observation data obtained by means of PREVARA eight indicators for action by general practitioners were constructed. These may be regarded as a measure of the quality of action at consultation level.

The two somatic indicators originate from the medico-technical part of PREVARA. The first of these is the percentage of obligatory acts performed as a criterion of adequate somatic action. The second indicator, inadequate action, consists of the absolute number of superfluous acts and acts performed under unknown conditions.

The five behavioral-science indicators were derived from a factor analysis on the behavioral-science items of PREVARA. These indicators are:

- openness to the patient's contribution
- elucidating the patient's request for help
- giving an adequate explanation
- structuring the consultation
- devoting attention to the continuity of the care.

The indicator with regard to psychosocial action were derived from a factor analysis on the items that were scored when there was reason for psychosocial action.

Results

First of all some results of the full study will be given.

In 1049 of the 1087 patient consultation, the consultation length could be reliably determined. Of these, 541 persons completed the questionnaire, the total having been reduced partly by exclusion on account of selection criteria, partly by non-response. Regarding sex these persons corresponded to the total research population. Their average age was somewhat higher. The consultation length varied from 1 to 32 minutes, with an average of 8.0 minutes and a standard deviation of 4.3 minutes. To give a better idea of the scatter, the total group was divided into four quartiles, for which the average length was calculated (Figure 1). Half of the consultation were of shorter duration than average; a quarter of the consultations had an average length of 13.9 minutes which is well above the average.

The consultation length increases up to the patient age of 54 years, declining thereafter.

Figure 1 The average consultation length (in minutes per quartile)

Quartiles	Length average	min.	max.
I	3.6	1.0	4.5
II	5.7	5.0	6.5
III	8.4	7.0	10.0
IV	13.9	10.5	32.0

The patient's sex, the frequency of contact during the preceding two months and the practice size did not result in any significant differences in the consultation length.

People who considered their health to be very good on average had shorter consultation than those who regarded their state of health as less good.

Patients with higher vocational or university education and patients with only primary education had longer consultations than people with an average level of education.

The consultation length increased with patients presenting more complaints. The presence of psychosocial signals caused the consultation length to increase by an average 2 1/2 minutes. The nature of the complaints proved to have a fairly strong correlation with the consultation length.

The ear, nose and throat complaints and those with respect to coughing for instance took less time than average.

When looking at the five complaints with the longest consultation length, it was striking that they were almost all complaints with a longer differential diagnosis, which required more physical examination and in which psychosocial factors played a greater part.

Of the nine reasons for coming, only three resulted in a significant difference in consultation length. If a patient came to be examined thoroughly for once, the consultation lasted more than a minute longer. Coming for explanation or advice also led to a longer consultation. If the patient came for medicine, the consultation lasted over a minute less on average.

The next part will deal with the relationship between the consultation length and the indicators that were used to assess quality of the general practitioners' action.

It is striking that all indicators describing the quality of the general practitioners' action at consultation level displayed a similar picture.

The consultation length increased as the doctor performed more acts, obligatory or superfluous (Figures 2 and 3).

Figure 2 Consultation length and obligatory medical acts

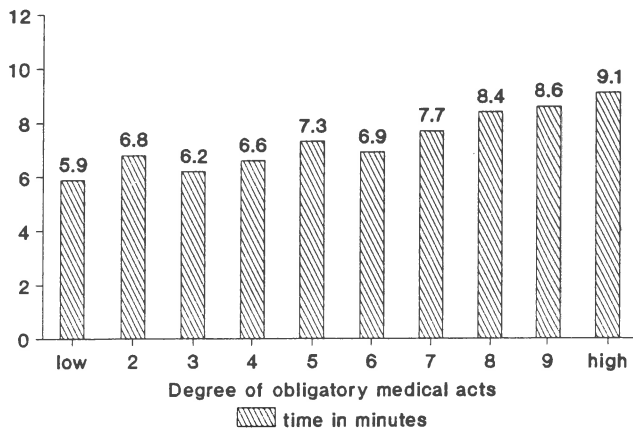
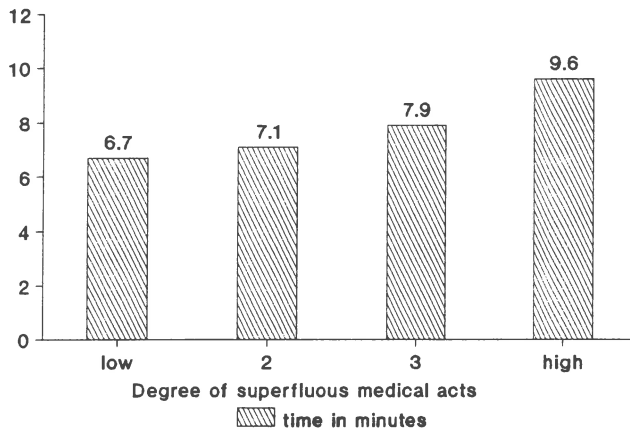


Figure 3 Consultation length and superfluous medical acts



The average consultation length per general practitioner varied strongly between the practitioners. The maximum was 13.5 and the minimum 4.7 minutes. The general practitioner proved to have a considerable effect on the length of the consultation ($\eta = .47$) and the differences were highly significant. The degree of the general practitioners' openness to the patient's ideas, his giving explanation, structuring the consultation etc. was proportionate to the length of the consultation (not shown in the figure).

For the quality of the action at general practitioner level the style of working of the doctor was plotted against the average consultation length (Figure 4). It is striking that on average the consultation of the "doer" lasted over a minute longer and that of the general practitioner with "minimum style of working" lasted over a minute less than that of the general practitioner with the "integrated style of working". The "middle group" occupied an intermediate position between the "doer" and "integrated general practitioner".

Figure 4 The average consultation length divided into working-style

Working-style	N	%	average length
Integrated	10	13.3	8.1
Doer	16	21.3	9.4
Minimum	25	33.3	7.0
Middle group	24	32.0	8.3

ETA = .42
P = .0028

The second question to which more detailed attention shall now be given is that for the relationship between consultation length and patient satisfaction.

Consonant with other literature, the patients in this study were generally highly satisfied with the consultation. For the various reasons for coming the percentages for sufficient attention from the doctor varied between 85 and 96%.

If we now investigate what the actual average length of the consultation in question was for the various reasons for coming (Figure 5), we are struck by a few interesting points, although the small absolute numbers make it difficult to draw general conclusions.

With a number of reasons for coming, the few patients who are thought of the doctor devoted too little attention to that reason, proved in fact to have had longer consultations than the large group who were satisfied with the attention received. That applied to talking about problems, coming for a check-up, the desire for explanation and advice. It applied very strongly to the four cases in which the patient came for a prescription and was dissatisfied with the attention paid. These four consultations were over twice as long as average.

Conversely, it was striking that an admittedly small number of patients were of the opinion that the doctor devoted too much attention to the reason for coming, whereas the consultations were of clearly shorter duration than average. When the opportunity presents itself we shall listen to these consultations again to grasp what was exactly going on.

Figure 5 The average actual length of the consultations distributed by the satisfaction of the patients with the attention devoted by the doctor to the applicable reason for coming. The relevant n appears in parentheses

reason for coming	attention devoted		
	too much	exactly enough	too little
worries about complaint	-	8.6	7.5
much trouble and hindrance	7.8	8.4	8.2
wanted to know what is going on	6.0	8.5	8.4
wanted to talk about problems	-	8.6	12.8 (3)
came for check-up	14.0 (1)	7.9	9.9 (4)
wanted to be thoroughly examined*	3.5 (1)	9.2	8.2 (8)
wanted explanation/advice*	4.5 (1)	9.1	9.7 (23)
wanted medicine*	5.5	7.5	16.2 (4)
wanted to be referred	7.0 (1)	8.3	7.9 (12)

* the differences are significant $p < 01$

Most of the answers to the questions about the patient's satisfaction were insufficiently scattered for further analysis. The only exception was the question concerning feeling at ease with the doctor. In that connection it is striking that the consultations in which patients did not feel at ease, or only partly so ($n = 74$), lasted significantly longer on average.

Conclusions

With some caution a number of conclusions can be drawn.

There are great differences in the length of general practitioners' consultation in this field. This fact emerges from most investigations. This also applies to the great differences between general practitioners in their average consultation length; there are fast doctors and slower ones. It is also striking that the consultation of one doctor can vary considerably in length.

Of all the variables, the general practitioner proved to be by far the most important one. He or she largely determines the consultation length. Probably personal disposition and management aspects play an important part in this. In addition, the general practitioner seems to be well able to adjust the consultation length to the circumstances of the consultation, such as:

- the nature of the complaint: simple complaints are dealt with more quickly than more complex problems;
- the presence of psychosocial signals: when these are present, more time is

- allotted;
- reason for coming: if one comes for a prescription the consultation is of shorter duration than if one comes for an explanation or a detailed examination;
- state of health: the doctor devotes more time to people feeling less healthy than to people who in general feel health;
- the number of complaints: when more complaints are presented, more time is devoted to the consultation.

In general a picture emerges of the general practitioner who deals very subtly with allocating time for the consultation. That not all general practitioners manage to this in an equally efficient manner is demonstrated by the great differences between the average consultation length of the doctors and the relation between the consultation length of the doctors and the relation between the consultation length and the style of working of the general practitioner. The doctors with an "integrated style of working" manage to combine this with efficient time management. Conversely, the "doers" prove to need more time for a consultation, while this is not accompanied by an increase in quality. The general practitioners with a minimum style of working do, it is true, manage to keep the consultations short, but fail to perform necessary acts.

An important conclusion from this may be that creating more time, for instance by reducing practice size, does not automatically lead to an increase in quality, but also that an improvement in quality does not necessarily lead to consultations of an unacceptable length.

With regard to the quality of action at consultation level, the eight indicators of general practitioners' action all display a similar picture. As the doctor does more, the consultations last longer. The indicators with respect to openness and psychosocial action are found to give the best explanation of the differences in consultation length. It proves that the "better" consultations take longer. But the indicator concerning superfluous action demonstrates that longer does not mean better by definition. This tallies with the results at general practitioners' level.

The patient's satisfaction, as measured here, is probably not a good criterion of the quality of general practitioner action, because patients are in general way satisfied or do not easily give vent to their dissatisfaction. In answer to an explicit question, 85 to 95% of the patients say that they are satisfied with the attention that the general practitioner pays to their reasons for coming and with the time devoted to them by the doctor.

If we return to the beginning of my story, we may note that the general practitioner need have less fear that the patient expects much more time from him

than the can give. Patients are evidently very realistic in their expectations. On the other hand, we may also say that the general practitioner must not a priori dismiss any correct advice to improve his consultation practice with the claim that this would cost him too much extra time. It has proved that good general practitioner consultations need not last longer than eight minutes.

10. DOCTOR-PATIENT COMMUNICATION, MEDICAL PERFORMANCE AND PATIENTS CHARACTERISTICS

By Henk Mokkink Ph.D.
University of Nijmegen
Department of General Practice and Family Medicine
P.O. Box 9101
6500 HB Nijmegen
The Netherlands

Introduction

Observing GPs at work in their practices it is interesting to note that each GP behaves quite uniformly from one consultation to the next and that in contrast to this, different GPs have very different working styles. Would it be possible to distinguish a limited set of practice styles?

In this paper the following issues will be discussed.

- the relationship between the way the GP communicates with patients and his or her medical performance; the practice style
- the relationship between the GP's practice style and his or her attitudes
- the relationship between the practice style and patient characteristics
- patient-centered attitudes: a comparison between GPs in Belgium, Britain and the Netherlands.

Part I

For the first three topics results from two major research projects done at the department of general practice in Nijmegen, were used; For the last topic an international study was done with J. de Maeseneer from Belgium and M. Whitfield from England.

1. Doctor-patient communication and GP's practice style

Doctor-patient communication has been an issue of interest for many years. In evaluation studies on medical education, it is often a topic isolated from the rest of the medical performance: how can we teach the student to be friendly to the patient, to smile now and then, not to interrupt the patient and so on.

The object of research was to determine the relationship between these communication skills and the GPs medical performance.

Methods

Two studies were performed.

In the first study (1980-'85) 57 GPs in Nijmegen and surroundings were observed. Sickness-fund data on prescription and referral rates of the observed GPs were available. The aspects of attitude considered relevant were measured by means of a questionnaire.

In the second study (1986-'90) another 75 GPs in Arnhem and surroundings were observed. The same method of observation and the same sickness-fund data were used. Furthermore a random sample of 20 female patients (aged 50-64 years) from each practice were studied.

Both studies took place in the east of the Netherlands. Both regions comprise a rural area and a city of about 200.000 inhabitants.

GP's performance

To measure GP's consultation behaviour two sources of data were available: data from the sickness-fund and observation data.

Data from the sickness-fund were used to construct two indicators:

- a. non-specific prescription behaviour: i.e. prescription of analgetics, antibiotics, corticosteroids, antitussives, tranquilizers, sedatives. These drugs cover a wide range of indications and are mostly used in case of minor ailments.
- b. The annual referral rate

The second source of data consisted of the observations of the GPs during at least two surgery hours in their own practice room. The observers were GPs who had just finished their vocational training.

The consultations were tape-recorded and all non-verbal activities like physical examination were registered.

Three indicators were constructed:

1. patient- and goal orientedness:
 - a. giving information to the patient: is the stream of information from the GP towards the patient to the point and comprehensive?
 - b. listening to the patient: does the GP stimulate the patient to talk about his or her thoughts, his feelings, his fears etc.?
 - c. is the consultation structured in a logical manner?
2. adequate diagnosis
3. superfluous diagnosis

The two latter indicators are measured for medical performance.

To judge whether a diagnostic action was to be considered adequate or superfluous, 23 different protocols were used. Each protocol indicates which medical actions are required (obligatory) and which medical actions are superfluous for an indicated reason for encounter. Thus the number of obligatory and superfluous actions performed by each GP can be calculated. For this purpose 600 recorded actions were used per GP.

Results

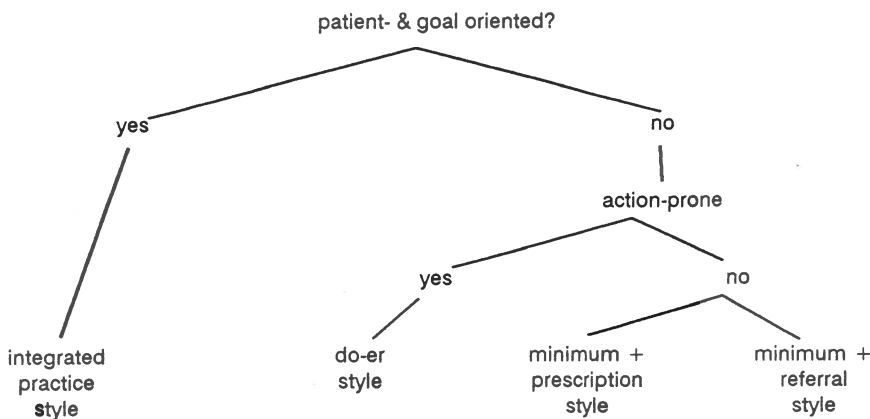
The results of the data from the sickness-fund were the following:

- a. No correlation was found between the non-specific prescription and the age distribution of the practices. This might indicate that differences between GPs are not due to differences in morbidity in the population. In other words: GPs who prescribe twice or even three or four times as much of these non-specific drugs, don't have a double or triple number of ill persons in their practices. They just have different prescription routines.
- b. Differences between GPs in referral rates could hardly be explained by differences in age-distribution in the population.

The data gathered by observing the GPs led to a typology of practice styles.

It has the form of a decision-tree (Figure 1).

Figure 1 Scheme of practice styles



The basic question is whether the GP acts in a patient- and goal oriented fashion. If so, he generally performs the most obligatory actions and hardly any superfluous ones. Besides that, he has a moderate referral rate and moderate non-specific prescription behaviour. This combination of characteristics was called the integrated practice style.

When the GP does not act in a patient- and goal oriented manner, the next question is whether or not he is "action-prone". In other words: does the GP always feel the need to do something?

A practice style that is "action-prone" in combination with not being patient- and goal oriented was called the do-er style. This do-er goes hand in hand with the performance of many obligatory actions as well as many superfluous actions. Besides this, the do-er style GP prescribes a lot of non-specific drugs and has a rather high referral rate.

Finally two so called minimum-styles were discussed. According to these styles the GP is neither patient- and goal oriented, nor action-prone: he performs few obligatory as well as few superfluous actions. What he does do however is prescribe a lot of non-specific drugs (minimum + prescriptive style) or refer often (minimum + referral style).

A combination of the two minimum-styles was also encountered.

For each GP a score on each style was calculated. A high score indicated that the GP had much in common with the style in question.

Practice styles and attitudes

Figure 2 shows the correlations between GP's practice styles and their attitudes.

Figure 2 GP's practice styles and attitudes

	integrated style	do-er style	min + presc style	min + referral style
responsibility for med. tasks	.33*	-.10	-.20	-.41**
risk taking in med. decisions	.38**	-.25*	-.15	-.39**
external control on gen. practice	.32*	-.45**	-.30*	-.14
doctor-patient relationship	.36**	-.45**	-.17	-.18
communication med. specialist	.29*	-.14	-.27*	-.39**

pearson: * p < .05, ** p < .01

There is a clear contrast between the positive correlations of the integrated practice style and the negative correlations of the other styles.

In general these correlations confirm the typology of practice styles. When looking, for instance at the style of the do-er, the negative correlations with "risk taking", "external control" and "doctor-patient-relationship" are striking. This indicates that the do-er is not in favour of external control or an open and equivalent relationship with the patient. He is afraid of being criticised by external control or by the patient for not having done enough.

The GP with the "minimum + referral style" feels a restricted responsibility for medical tasks and prefers to be rather safe than sorry. Although he refers often, he does not ask the specialist a clear question when doing so.

GPs with the integrated style differ clearly in their attitudes from the colleagues with other practice styles.

One can conclude that the correlations between practice styles and attitudes confirm the validity of the constructed styles!

Practice styles and patient-outcome

Figure 3 shows that in 15 out of 18 diseases no correlation was found between the practice style of the GP and the percentage of patients with untreated diseases. Yet, in three cases significant correlations were found. They indicate that in practices of GPs with an integrated style, there is relatively less untreated hyperglycaemia, chronic bronchitis and heart failure than in other practices.

These findings do not support the hypothesis that low referral and low prescription rates are at the cost of the patient's health.

Figure 3 GP's practice styles and patients' health

NO CORRELATION in 15 out of 18 diseases / risk factors

	integrated style	do-er style	minimum style
hyperglycaemia	-.19	.32**	.06
chronic bronchitis	-.25	.35**	.16
heart failure	-.28*	.19	.23*

spearman: * p < .05, ** p < .01 (n=75)

Figure 4 shows some interesting and significant correlations. There is a pronounced contrast between the integrated style and the other styles. This indicates that patients of GPs with an integrated practice style feel more healthy, have less complaints, are more in favour of self care in minor ailments and visit their general practitioner less often.

Figure 4 GP's practice styles and patients' attitudes

	integrated style	do-er style	minimum style
feeling healthy	.24*	-.34**	-.16
number of complaints	-.38**	.35**	.28*
notions about selfcare in minor ailment	.30**	-.25*	-.28*
number of visits to GP	-.41**	.35**	.32**

pearson: * p < .05, ** p < .01 (n=75)

In contrast, the do-er, who is very active, tries to do his best for his patients, prescribes a lot of drugs, has patients who feel less healthy and have more complaints. These patients are not in favour of self care and they visit their GP more often.

Thus, the GP's practice style - including communication skills - is related to patients expectations and illness behaviours.

Part II

Some of the questionnaires, mentioned in part I, especially those measuring patient-centered attitudes were also completed by a sample of 90 Belgian GPs and (after translation into the English language) by 470 English GPs. They came from both urban and rural areas.

There were two important drawbacks in the collection of data in this study. The translation of the questionnaire from Dutch to English posed some problems. The samples were not quite representative for the three countries.

Results

Dutch and British general practitioners are more in favour of an equivalent doctor-patient relationship than Belgian GPs. Similar differences were found in

attitudes towards the taking of risks in medical decision making: Belgian GPs prefer to take little risk when taking medical decisions while Dutch and British doctors tend to wait and see.

Figure 5 GP's attitudes in Belgium, UK and the Netherlands

	Belgium	UK	The Netherlands
Patient-oriented attitude	44%	64%	70%
Prefer not to take risks when making medical decisions	60%	54%	24%

Conclusions

How can the difference between the countries be explained? Belgian GPs have no registered patients and are in a situation of competition with their colleagues and with medical specialists. Furthermore, Belgian patients can consult whichever doctor they want to. This results in patients frequently changing GPs and it explains the GP's defensive working styles.

The difference in vocational training between the 3 countries could also be part of the explanation of the difference in attitudes between the general practitioners. Belgium has no compulsory vocational training, the British work in hospital in 2 years and in GP during one year. In Holland two thirds of vocational training is spent in general practice.

Conclusion

- Two types of outcome research in general practice can be distinguished:
 - * the evaluation of the outcomes of a certain diagnostic or therapeutic procedure or the evaluation of a drug.
 - * the evaluation of the outcomes of the GP as the user of medical procedures and as a individual person with his or her own feelings, opinions, ways of decision making, daily routines and so on.

This study shows that the latter type is important in outcome research.

- Besides outcome criteria such as mortality and morbidity, two criteria can be added:
 - * patient's subjective health
 - * patient's coping with health-problems.

Every now and then a person doesn't feel well. Some people are inclined to complain and go to a doctor while others do not. GP's practice style seems to affect the way patients experience their own health. Many GPs educate their patients in how to cope with their health problems.

- As the reader undoubtedly will have noticed, the design of this study is cross-sectional. This is a weakness which makes it difficult to state what is the cause and what is the effect or outcome. It might be possible for instance, that patients with certain attitudes choose a GP with a certain practice style. Two remarks must be made here:
 - * the presented analyses were done separately for urban and for rural practices. In urban areas people can easily choose another GP when they are dissatisfied. In rural areas people in general do not have this choice. They can choose the doctor in their own village or go to a GP in an other town or village at least 10 kilometres away (which is quite far in Holland). Yet in both groups we found similar results.
 - * Maybe the question of cause and effect is the question of "which was first: the chicken or the egg", in other words, it is an irrelevant question. The correlations found, indicate a process that is going on in the interaction between the GP and his/her patients. Like in all other human interactions there is two way traffic. The question of real concern is: do we want these processes, and if not, how can they be changed?

- The last conclusion is that the ability of GPs to take care of an effective communication between doctor and patient, is a very basic skill in general practice. This communication however is not an isolated happening; it gets its relevance from the context. The context of doctor-patient communication is the problem solving process. That is why it is important to be not only patient-oriented but also to be goal-oriented.

11. POLISH RESEARCH ON DOCTOR-PATIENT COMMUNICATION

Irena Heszen - Niejodek
Department of clinical psychology
Faculty of education and psychology
Silesian University
Tyski 53, 40-126 Katowice
POLAND

Introduction

The majority of the Polish medical staff is biologically-oriented and the problem of doctor-patient communication is relatively neglected. About ten years ago some psychologists of the Warsaw Academy of Medicine started a line of research on doctor-patient communication from the conviction that in the straitjacket of financial limitation a possible means of bettering the quality of medical care was through improvement of doctor-patient communication. These studies were continued at the Silesian University. A great deal of empirical data has been gathered, of which this paper is only a small part.

Method

At the beginning of the project, two methods designed for analysis of verbal interaction were available in Poland: those developed by Bales (1950) and by Stiles (1978). Attempts undertaken at the department to adapt each of these showed discouraging results. The Balesian system is too general and does not take into account the peculiar nature of doctor-patient communication, and moreover some of the categories overlap. Stiles' taxonomy does not distinguish emotional exchange between doctor and patient which seems to be an important lacuna, the categories were found to be unreliable.

Hence, the decision of the research group was to develop a new taxonomy. It was designed to study the intent of verbal responses, not their content or form. First, the verbal units of interaction were distinguished, the smallest ones which could be separately evaluated. Then, each of them was categorized. Congruently with Bales' concepts, the assumption was that every act by doctor or patient during their encounter was "interactive", that is, its intent was for the speaker to influence his partner or himself. From the psychological point of view several spheres of this influence are possible, therefore two main classification principles were used:

- who is the object of interaction: the speaker himself or his partner
 - what is the sphere of intended influence: cognition, emotion or behavior.
 Another fact that was taken into account was that the intent of affecting someone's emotional state could be improving it or making it worse. A complex intent was also possible, for example affecting emotions and behavior at the same time. The taxonomy was based on these principles (Figure 1).

Figure 1 The system of verbal categories

Sphere of intended influence		Addressee of intended influence	
		Speaker*	Partner*
Cognition		1. S : C*	12. P : C*
	Improving	2. S : E+*	13. P : E+
Emotion	Making worse	3. S : E-	14. P : E-
Behavior		4. S : B	15. P : B*
Cognition and Emotion	Improving	5. S : C/E+	16. P : C/E=+*
	Making worse	6. S : C/E-	17. P : C/E-*
Behavior and Cognition		7. S : B/C	18. P : B/C*
Behavior and Emotion	Improving	8. S : B/E+	19. P : B/E+*
	Making worse	9. S : B/E-	20. P : B/E-
Behavior	Improving	10. S : B/C/E+	21. P : B/C/E+*
Cognition and Emotions	Making worse	11. S : B/C/E-	22. P : B/C/E-

* The categories used in at least 50% of interviews.

Twenty two different categories could be distinguished. In a set of tape recorded interviews it was found that only nine of these were used by doctors or patients in at least 50% of interviews. Only these nine were taken into account in further analysis. These were:

- (1) S : C. In this utterance the intent is to gain emotionally neutral information, for example: "How long have you been feeling that?".
- (12) P : C. The intent is to give the partner emotionally neutral information, for example: "I am 34".
- (13) P : E+. The intent is to improve the partner's emotional state by giving him reassurance, emotional support, encouragement and so on, for example: "don't worry".
- (15) P : B. The intent is to influence the partner's behavior by giving him advice, instructions, directions, for example: "Open your

- mouth".
- (16) P : C/E+. The intent is to give the partner information which will improve his emotional state, for example: "This illness is completely curable".
- (17) P : C/E-. The intent is to give the partner information which will worsen his emotional state, for example: "The results of the tests are alarming".
- (18) P : B/C. The intent is to influence the partner's behavior by not only giving him instruction on what to do, but also how to do it, for example: "Take these pills twice a day".
- (19) P : B/E+. The intent is to influence the partner's behavior not only by giving him instructions, but also by encouraging him to follow them, for example: "Be sure to take this medicine".
- (21) P : B/C/E+. The intent is to influence the partner's behavior by giving him instructions on what to do and how to do it and by encouraging him to do it as well, for example: "Be sure to take these pills twice a day".

With this taxonomy, similarly to those constructed by other authors, only the structure of the doctor-patient exchange can be discovered. Both the process of the interaction in its dynamics and the general character of the encounter are outside our considerations.

To describe the general type of doctor-patient relationship, each entire interview was evaluated on the following dimensions:

- (1) Doctor's directiveness: the number of orders given to the patient.
- (2) Doctor's activity: the proportion of his utterances.
- (3) Patient's activity: the proportion of his utterances.
- (4) Doctor's emotional attitude to the patient:
 - unfavorable, coded 1,
 - changing, coded 2,
 - favorable, coded 3.
- (5) Patient's partnership status:
 - patient does not initiate activity of his own, coded 1,
 - patient tries to initiate, but doctor ignores him, coded 2,
 - patient does act as a partner, coded 3.

Results and discussion

The structure of doctor-patient interaction in American and Polish studies

Polish-American differences in the structure of doctor-patient interaction were

investigated in the first study group, consisting of 62 patients and their 11 doctors. Patients were selected suffering from three illnesses, each representing a different degree of threat to health. All of them were "new" patients at specialist out-patients clinics in different districts of Warsaw. Two complete verbal interactions between a patient and his physician were tape-recorded: the first visit and a subsequent one, terminating the first stage of treatment. Hence, the material for analysis was a set of 124 complete interviews between doctor and patient.

In the available literature three studies were found with data on the structure of doctor-patient exchange comparable to the presented study. These were works by American psychologists: Davis (1971), Freeman with her team (1971) and Stiles with his team (1979). Each of them applied a different method to analyse doctor-patient interaction: Davis used Balesian categories, Freeman used a modified Bales' system and Stiles applied a taxonomy of his own. To make the comparison between their results and the results of the present study possible, "rough" categories had to be employed, which resulted in the loss of much of the detailed information. Nevertheless, the differences found seem to be meaningful.

Figure 2 The structure of doctor-patient interaction

Interaction characteristics	Source			
	Davis (1971)	Freemon (1971)	Stiles (1979)	Present study
Doctor's behavior				
Asking questions	14,4	20,3	28,5	25,3
Giving information	46,6	35,5	-	28,2
Giving instructions	9,7	8,0	10,1	16,0
Positive emotions	3,8	5,6	-	19,6
Patient's behavior				
Asking questions	6,8	2,9	2,6	4,3
Giving information	54,0	46,9	-	66,8
Positive emotions	1,4	2,9	-	21,2
Ratio of verbal output				
Doctor : Patient	51 : 49	56 : 44	55 : 45	64 : 36

The table presents accordingly mean percentages of doctors' and patients responses.

Figures add up to less than 100% due to omission of rare categories and absence of some categories in Stiles's taxonomy.

Figure 2 shows that Polish doctors were more active, more directive and much more warmer towards their patients. On the other hand, American doctors were more informative. The general image of an American doctor emerging from these data may be defined as "the cold expert", giving a lot of information but little emotional support. A Polish doctor seems to be a "warm autocrat" - friendly, protective but perhaps too dominant. The positive emotional attitude is mutual; Polish patients are also much warmer towards their doctors than Americans. Complementary to the doctors' behavior, they are also more informative but less active when compared to Americans.

Doctor-patient interaction and patient compliance in American and Polish studies

The scale of patient noncompliance is alarming. In many studies the percentage of persons who follow the regimen prescribed is far below 100. The studies on this issue performed by the author and her team included more than 650 outpatients suffering from chronic illnesses. The results showed that the numbers taking the doctors' advice varied from 42% to 93%, depending mainly on the kind of treatment recommended. Only about 50% of patients complied with doctors' advice concerning lifestyle changes (Heszen-Klemens, 1987).

In these same studies the influence of doctor-patient interaction on patient compliance was evaluated on the basis of a longitudinal approach. The entire exchange during the patients' first visit to the doctor was recorded. The doctor's instructions were then transcribed from these recordings. The patient was interviewed 7 - 10 days after the first visit (i.e. after the time needed for carrying out the doctor's instructions) in order to estimate patient compliance. This procedure was repeated at the end of the first stage of treatment, thus the relation between the doctor-patient interaction and patient compliance was evaluated twice: at the beginning of treatment and at the end of its first phase.

Figure 3 Doctor-patient interaction and patient compliance

Interaction characteristics	Patient compliance	
	At the beginning of the treatment	At the end of the first phase
Doctors' directiveness	- 0,39**	- 0,26
Patients' questions /S: C/	0,33*	0,34*
Patients' activity	0,41**	0,02

The table presents Spearman Rank Correlation Coefficients

- * significant at 0,05 level
- ** significant at 0,01 level

As can be seen in Figure 3, very few significant connections between doctor-patient interaction and patients' subsequent compliance were found. Much stronger relations were found between the interaction and patients' own health activity, but these are beyond the present concern. The table shows that patient compliance at the beginning of treatment deteriorated when the amount of doctors' orders increased and improved when patients asked more questions and were generally more active during the interview.

Only one significant correlation was found at the end of the first phase of treatment. It may be concluded that the doctor-patient interaction, as a factor influencing patient compliance, loses in importance as time goes by. Again patients were more inclined to follow doctors' instructions when they asked more questions.

The influence of doctor-patient interaction on patients' compliance has been studied among others by the American psychologists: Davis (1968, 1971) and also Francis et al. (1969).

The results of the present study failed to confirm their findings. Unlike in these authors' studies, no significant correlations were found between the characteristics of interaction in the social-emotional area and patients' compliance. Contrary to Davis's previous study, the present study's findings proved that doctors' directiveness decreased patients' compliance, while patients' greater activity improves it. How may these divergencies be explained?

A possible reason may be the differences between the United States and Poland in the structure of doctor-patient interaction, as is shown in Figure 4.

Figure 4 Doctor-patient interaction and patient compliance in American and Polish studies

Interaction characteristics	American compared with Polish	Influence on compliance in American studies	Influence on compliance in Polish studies
Emotional exchange	Less intense	Positive	None
Doctors' directiveness	Less directive	Positive	Negative
Patients' activity	More active	Negative	Positive

It was already stated that emotional exchange between American doctors and patients is less intense, doctors are less directive and patients more active in comparison to their Polish counterparts. Taking into account these differences it may be concluded that the relations between certain characteristics of doctor-patient interaction and subsequent patient compliance, as many others in

psychology, may be curvilinear. Thus expression of mutually positive feelings can improve compliance to some degree (as was found in American studies), then becomes unimportant (as found in our study) or even, when the feelings are too intense, compliance may decrease. Similarly, the doctor's directiveness has a certain positive influence on the patient's compliance (as proved by Davis's data), but when the doctor is too directive (as the doctors in our study probably were), compliance decreases. Finally, the relationship between patient's activity during the medical interview and his compliance can be of the same kind; i.e. increase in patient's activity to some extent leads to better compliance (our study gives an example of such relationship), but when the patient becomes too active, compliance gets worse (as in Davis's study). Broadly speaking, an excessively strict role performance by doctor and patient, just as refusal to accept role requirements, seems to have a negative influence on patient's compliance.

There are many possible reasons for poor patient compliance which cannot be considered here in detail. One is the lack of sufficient medical information on the patient's side, resulting from distortion in doctor-patient communication.

Empirical studies aimed at improving doctor-patient communication

The last example from our research is a continuation of American and English investigations on increasing patient participation in the medical interview in order to improve the effects of communication between doctor and patient. It has already been shown that the number of patients' questions was very small both in American and Polish studies. Although patients' questions are not always answered by their doctors, undoubtedly these patients who ask questions, have a chance of gaining more medical information and of better understanding doctors instructions.

Roter was the first to show that it is possible to increase the number of patients' questions by purposeful intervention, i.e. preparing questions together with the patient and stressing the importance of asking questions during medical interview.

However, in the experimental group doctor-patient interaction involved more anger, less sympathy and more anxiety from the patient's side (Roter, 1979). This research line was continued by the English psychologists Robinson and Whitfield, who were also successful in encouraging patients to produce more questions and comments (1987).

The present study aimed not only at confirmation of these results in Polish patient, but at examination of the relation between patients' questions and the amount of medical information they receive, and also at learning about doctors' attitudes to this form of patient activity.

The subjects were 80 out-patients and their 20 doctors. Half of the patients were randomly assigned to the experimental group. Immediately before a visit to the doctor they were encouraged to ask questions and instructed how to do so. In both groups the entire doctor-patient interview was recorded and analysed by two independent judges. The Kendall Coefficient of Concordance between them was very high, i.e. .99.

Figure 5 The number of persons asking questions

The group	Asking questions		Total
	Yes	No	
Experimental	39	1	40
Control	7	33	40
Total	46	34	80

Chi² = 52.378; df = 1 p < .001

In Figure 5 one can see that the intervention proved to be very effective in increasing the number of patients asking questions. In the experimental group all but one person questioned their doctors, while in the control group only seven persons asked questions.

Figure 6 Asking questions and information received

Asking questions	Information received		
	\bar{x}	s	t
Yes	32,28	20,63	4,7
No	16,02	9,80	p < .001

As shown in Figure 6, comparison of the amount of medical information given to the patients questioning their doctors and that given to those who did not ask any questions, showed that the former received twice as much medical information.

Figure 7 Asking questions and doctors' attitude

Asking questions	Doctors' attitude			Total
	+	0	-	
Yes	7	19	20	46
No	14	19	1	34
Total	21	38	21	80

Chi² = 18.135 df = 2 p < .001

The data in Figure 7 show that the doctors' attitude towards the patients asking questions was less favorable. The doctors showed a clear positive attitude to 21 patients, while in this group there were twice as many who did not ask any questions than patient who did ask questions. The doctors attitude was certainly negative towards 21 patients; all but one of them asked questions. These results could be synthesized with those obtained by Roter to justify the conclusion that the doctors' attitude is the very factor responsible for patients' difficulties in asking questions during medical interview.

The results of this study are especially valuable in Poland. Our possibilities of changing doctor behavior are modest, because our doctors are rather resistant to psychological influence. But this study has shown that the course and outcomes of doctor-patient communication may be changed by influencing patient behavior. However, the price payed was a less positive attitude of the doctor towards the patient with increased participation during the doctor-patient encounter.

References

- BALES, R.F. Interaction process analysis: A method for the study of small groups. Cambridge: Addison-Wesley, 1950
- DAVIS, M.S. Variations in patients' compliance with doctors advice: An empirical analysis of patterns of communication. *Amer.J. publ. Hlth.*; 58, 1968, p. 274-288
- DAVIS, M.S. Variations in patients' compliance with doctors' orders: Medical practice and doctor-patient interaction. *Psychiat. Med.*; 1971, 2, p. 31-54
- FRANCIS, V., B.M. KORSCH, M.J. MORRIS. Gaps in doctor-patient communication: Patients' response to medical advice. *New Engl. J. Med.*, 280, 1969, p. 535-540
- FREEMON, B., V.F. NEGRETE, M.S. DAVIS, B.M. KORSCH. Gaps in doctor-patient communication: Doctor-patient interaction analysis. *Pediat. Res.*, 5, 1971, p. 298-311

- HESZEN-KLEMENS, I. Patients' noncompliance and how doctors manage this. *Soc. Sci. Med.*, 24, 1987, p. 409-416
- ROBINSON, E.J., M.J. WHITFIELD. Participation of patients during general practice consultation. *Psych. Hlth*, 1, 1987, p. 123-132
- ROTER, D.L. Altering patient behavior in interaction with provider. In: D.J. Osborne, M.M. Gruneberg, J.R. Eiser (eds). *Research in Psychology and Medicine*, vol 2. London: Academic Press, 1979
- STILES, W.B. Verbal response modes and dimensions of interpersonal roles: A method of discourse analysis. *J. Pers. soc. Psychol.*, 36, 1978, p. 693-703
- STILES, W.B., S.M. PUTMAN, M.H. WOLF, S.A. JAMES. Verbal response mode profiles of patients and physicians in medical screening interviews. *J. med. Educ.*, 54, 1979, p. 81-89

REFERENCES

- BALES, R.F. *Interaction process analysis: A method for the study of small groups*. Cambridge: Addison-Wesley, 1950
- BASS, M.J., G. BUCK, L. TURNER ET AL. The physician's actions and the outcome of illness in family practice. *J Fam Pract.*; 1986, 23, p. 43-47
- BENSING, J.M. *Doctor-Patient Communication and the Quality of Care. An observation study into affective and instrumental behavior in general practice (thesis)*. Utrecht: NIVEL, 1991
- BENSING, J.M. Doctor patient communication and the quality of care. *Social Science & Medicine*; 32, 1991, No. 11, p. 1301-1310
- BENSING, J.M., J. DRONKERS. Instrumental and affective aspects of physician behavior. *Medical Care*; 30, 1992, No. 4, p. 283-298
- CRIJNEN, A.A.M., H.F. KRAAN. Measuring patient satisfaction with communication. In: Kraan H.F., A.A.M. Crijnen. *The Maastricht History-taking and Advice Checklist*. Amsterdam: Lundbeck fond., 1987
- DAVIS, M.S. Variations in patients' compliance with doctors' orders: Medical practice and doctor-patient interaction. *Psychiat. Med.*; 1971, 2, p. 31-54
- DAVIS, M.S. Variations in patients' compliance with doctors advice: An empirical analysis of patterns of communication. *Amer.J. publ. Hlth.*; 58, 1968, p. 274-288
- DIMATTEO, M., D. DINICOLA. *Achieving patient compliance*. New York: Pergamon Press Inc., 1982
- DREW, J., J.D. STOECKLE, J.A. BILLINGS. Tips, status and sacrifice: Gift giving in the doctor-patient relationship. *Social Science & Medicine*; 17, 1983, p. 399-404.
- EISENTHAL, S., A. LAZARE. Expression of patient's request in the initial interview. *Psychological Reports*; 40, 1977, p. 131-138
- ELSTEIN, A., L. SHULMAN, S. SPRAFKA. *Medical problem solving*. Cambridge: Harvard University Press, 1978
- ENGEL, G. How much longer must medicine's science be bound by a seventeenth century world view? In: WHITE, K.L. *The task of medicine: Dialogue at Wickenburg*. California: Menlo Park, The Henry J. Kaiser Family Foundation, 1988
- FRANCIS, V., B.M. KORSCH, M.J. MORRIS. Gaps in doctor-patient communication: Patients' response to medical advice. *New Engl. J. Med.*, 280, 1969, p. 535-540
- FRASER, C. An analysis of face-to-face communication. In: *Communication between doctors and patients* (ed. A. Bennett) p. 5-29. Oxford: Oxford University Press, 1976

- FREEMON, B., V.F. NEGRETE, M.S. DAVIS, B.M. KORSCH. Gaps in doctor-patient communication: Doctor-patient interaction analysis. *Pediat. Res.*, 5, 1971, p. 298-311
- FREIRE, P. *Pedagogy of the oppressed*. New York: Seabury Press, 1970
- GROL, R.P.T.M. et al. *Huisarts en somatische fixatie: Theorie en praktijk van de preventie van somatische fixatie (General practitioner and somatic fixation: Theory and practice of the prevention of somatic fixation)*. Utrecht/Antwerpen: Bohn, Scheltema & Holkema, 1983
- HALL, J.A., D.L. ROTER, N.R. KATZ. Meta-analysis of correlates of provider behavior in medical encounters. *Medical Care*; 26, 1988, p. 657-675
- HAMBLETON, R.L., L.L. COOK. Latent trait models and their use in the analysis of educational test. *Journal Educat. Measur.*; 1977, 14, p. 95-96
- HESZEN-KLEMENS, I. Patients' noncompliance and how doctors manage this. *Soc. Sci. Med.*, 24, 1987, p. 409-416
- HILFIKER, D. *Healing the wounds: A physician looks at his work*. New York: Pantheon, 1985
- IDLER, E.L., S.V. KASL. Health perceptions and survival: Do global evaluations of health status really predict mortality? *Journal of Gerontology*; 46, 1991, S55-65
- INUI, T.S., W.B. CARTER. Problems and prospects for health services research on provider-patient communication. *Medical care*; 23, 1985, p. 521-538
- KLASS, P. *A not entirely benign procedure: Four years as a medical student*. New York: Signet, 1987
- KORSCH, B., V. NEGRETE. Doctor-patient communication. *Scientific American*; 227, 1972, p. 66-74
- KRAAN, H.F., A.A.M. CRIJNEN. *The Maastricht history-taking and Advice Checklist (MAAS)*. Amsterdam: Lundbeck fond., 1987
- KRAAN, H., A. CRIJNEN. *The Maastricht History Taking and Advice Checklist: Studies of instrumental Utility*. Amsterdam: Lundbeck, 1987
- KUYVENHOVEN, M.M. *Quality of performance of general practitioners confronted with patients' non-specific complaints*. Dissertation University of Utrecht 1988
- KUYVENHOVEN, M.M., H.M. PIETERS, H.M. JACOBS, F.W.M.M. TOUW-OTTEN, C. SPREEUWENBERG. The association between peer consultations and three aspects of clinical competence. *Family Practice*; 1990, 7, p. 163-167
- KUYVENHOVEN, M.M., H.M. JACOBS, F.W.M.M. TOUW-OTTEN, J.C. VAN ES. Written simulations of patient-doctor encounters 1. Comparison of the performance in the simulation with prescription and referral data in reality. *Family Practice*; 1984, 1, p. 25-29
- MCQUIRE, C.H., G. PAGE. *The assessment of clinical performance by written and oral simulation. Report to the faculty 1972-3*. Chicago/Illinois: Centre for Educational Development, University of Illinois College of Medicine, 1973

- MEEUWESSEN, L. Spreekuur of zwijguur: Somatische fixatie en sekse-asymmetrie tijdens het medisch consult (Speaking or being silent: Somatic fixation and gender asymmetry during the medical interview.) (Dissertation). Nijmegen: University of Nijmegen, 1988
- MELKER, R.A. DE, M.M. KUYVENHOVEN. Management of upper respiratory tract infections in Dutch family practice, *British Journal of General Practice*; 1991, 41, p. 504-507
- MITCHELL, S.K. Interobserver agreement, reliability and generalizability of data collected in observational studies. *Psychol. Bull.*; 1979, 86, p. 376-390
- NORMAN, G.R., P. TUGWELL, J.W. FEIGHTNER, L.J. MURRIN, L.L. JACOBY. Knowledge and clinical problem solving. *Med. Educ.*; 1985, 19, p. 344-356
- OSLER, W. The master-word in medicine. In: *Aequanimitas with other addresses to medical students, nurses, and practitioners of medicine*. Philadelphia: Blakiston, 1904
- PENDLETON, D. Doctor-patient communication: A review. In D. PENDLETON & J. HASLER (Eds.), *Doctor-patient communication*, p. 5-53. London: Academic Press, 1983
- PIETERS, R.M. Utrecht Consultation Assessment Method (UCAM). Reliability and validity of actor simulation in assessing consultation skills of trainees in the vocational training. Construction and validation of a measurement instrument. Utrecht: Invitational workshop on Doctor-patient communication and the quality of care, 1991
- PUTMAN, S., W. STILES, M. JACOB, S. JAMES. Patient exposition and physician explanation in initial medical interviews and outcomes of clinic visits. *Medical Care*; 23, 1985, p. 74-83
- ROBINSON, E.J., M.J. WHITFIELD. Participation of patients during general practice consultation. *Psych. Hlth*, 1, 1987, p. 123-132
- ROTER, D.L. Altering patient behavior in interaction with provider. In: D.J. Osborne, M.M. Gruneberg, J.R. Eiser (eds). *Research in Psychology and Medicine*, vol 2. London: Academic Press, 1979
- ROTER, D.L. *The Roter Method of Interaction Process Analysis*. Baltimore: The Johns Hopkins University, 1989
- SACKS, H.E., E.A. SCHLEGLOFF, A. JEFFERSON. A simplest systematics for the organisation of turntaking for conversation. *Language*; 50, 1974, 4, p. 696-735
- SCHOUTEN, J. *Anamnese en advies*. Alphen a/d Rijn, Brussel: Stafleu, 1982
- SCHRÖER, C. *Geleidelijk onderhandelen (Gradual negotiation)* (Unpublished doctoral thesis). Nijmegen: University, Department of Clinical Psychology, 1985
- SENIOR, J.R. *Towards the measurement of competence of medicine*. Philadelphia: National Board of Medical examiners, 1976
- SPRECHER, P., E. THOMAS, L. HUEBNER, B. NORFLEET, K. JACOBY. Effects of increased physician-patient communication on patient anxiety. *Professional Psychology: Research and Practice*; 14, 1983, p. 251-255

- SPRINGORUM, D. Spreken in gesprekken (Speaking in conversations). Groningen: Wolters-Noordhoff, B.V., 1981
- STILES, W.B., S.M. PUTMAN, M.H. WOLF, S.A. JAMES. Verbal response mode profiles of patients and physicians in medical screening interviews. *J. med. Educ.*, 54, 1979, p. 81-89
- STILES, W.B., S.M. PUTNAM, M.C. JACOB. Verbal exchange structure of initial medical interviews. *Health Psychology* 1; 4, 1982, p. 315-336
- STILES, W.B. Verbal response modes and dimensions of interpersonal roles: A method of discourse analysis. *Journal of Personality and Social Psychology*; 36, 1978, p. 693-703
- STILES, W.B., S.M. PUTNAM, M.H. WOLF, A.J. SHERMAN. Interaction exchange structure and patient satisfaction with medical interviews. *Medical Care*; 17, 1979, p. 667-681
- STILLMAN P., P., M. REGAN, D. SWANSON. A diagnostic fourth year performance assessment. *Archives of Internal Medicine*; 147, 1987, p. 1981-1985
- SWANSON, D. A measurement framework for performance-based-test. In: *Further Developments in Assessing Clinical Competence* (Ed. I. Hart & R. Harden) p. 13-36. Montreal: Heal Publications, 1987
- TAN, L.H.V., C.A.G.M. FOOLEN, G. VAN GELDROEP. Towards a national assessment of skills in the vocational training for general practice. In: Scherpbier, A.J.J.A., W. Bender, R. Zwierstra (eds.) *Proceedings 3e Ottawa Conference on assessing clinical competence - Groningen, Boekwerk 1989*
- THIEL, J., H.F. KRAAN, C.P.M. VAN DER VLEUTEN. Reliability and feasibility in measuring medical interviewing skills with the revised Maastricht history-taking and advice checklist (MAAS-R). *Medical Education*; 1991, 25, 224-229
- THIEL, VAN J., H. KRAAN, C. VAN DER VLEUTEN. Reproducibility of scores with the revised Maastricht history-taking and advice checklist (MAAS-R). Utrecht: Invitational workshop on Doctor-patient communication and the quality of care, 1991
- THORNDIKE, R.L. *Applied psychometrics*. Boston: Houghton Muffin cie, 1982
- TIELENS, V.C.L. Consultation length and the patient's view: aspects of quality. Utrecht: Invitational workshop on Doctor-patient communication and the quality of care, 1991
- TUCKETT, D., M. BOULTON, C. OLSON, A. WILLIAMS. *Meetings between experts*. New York: Tavistock Publications, 1985
- TUCKETT, D. *Meetings between experts: An approach to sharing Ideas in medical consultations*. London: Tavistock, 1985
- VAN THIEL, J., H.F. KRAAN, C.P.M. VAN DER VLEUTEN. Reliability and feasibility of measuring medical interviewing skills: the revised Maastricht History-taking and Advice checklist. *Med. Educ.*; 1991, 25, p. 224-229

- VLEUTEN, C. VAN DER, D. SWANSON. Assessment of clinical skills with standardized patients: state of the art. *Teaching and Learning in Medicine*; 2, 1990, p. 58-76
- WAITZKIN, H., J.D. STOECKLE. The communication of information about illness: Clinical, sociological and methodological considerations. *Advances in psychosomatic Medicine*; 8, 1972, p. 180-215
- WEST, C. *Routine complications: Troubles with talk between doctors and patients*. Bloomington: Indiana University Press, 1984
- WILLSON, P., J.R. MCNAMARA. How perceptions of a simulated physician-patient interaction influence intended satisfaction and compliance. *Social Science & Medicine*; 16, 1982, p. 1699-1703
- WRIGHT, B.D. Solving measurement problems with the Rasch model. *Journal Educat. Measur.*; 1977, 14, p. 97-116

