The Decisive Moment

Making diagnostic decisions and designing treatments

ico

THE DECISIVE MOMENT

MAKING DIAGNOSTIC DECISIONS AND DESIGNING TREATMENTS

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Twente,
op gezag van de rector magnificus,
prof. dr. H. Brinksma,
volgens besluit van het College voor Promoties
in het openbaar te verdedigen
op vrijdag 10 september 2010 om 16:45 uur

door

Marleen Groenier
geboren op 29 september 1979
te Groningen

Dit proefschrift is goedgekeurd door de promotoren:

Prof. dr. J.M. Pieters
Prof. dr. C.L.M. Witteman

CONTENTS

CHAPTER 1

General Introduction

It has always seemed to me a particular duty of the psychologist from time to time to leave his laboratory and with his little contribution to serve the outside interests of the community. Our practical life is filled with psychological problems which have to be solved somehow, and if everything is left to commonsense and to unscientific fancies about the mind, confusion must result, and the psychologist who stands aloof will be to blame. (Münsterberg, 1914, p. vii)

As early as the 1890's, Hugo Münsterberg (1899), then president of the American Psychological Association, recognized that the application of fundamental psychological knowledge is called for to solve everyday problems. As an applied scientist, Münsterberg valued the scientific method and pointed out that only following the scientific method results in useful knowledge. With the scientific method knowledge about phenomena in the world (such as gravity) is generated by deriving hypotheses from observations of this phenomenon, testing these hypotheses through experiments, evaluating the results of these experiments, and drawing conclusions about the phenomenon under study.

However, there seems to be a gap between scientific theory and its application – in particular in the field of clinical psychology. Although the majority of psychologists judge scientific knowledge useful for clinical practice (Beutler, Williams, Wakefield, & Entwistle, 1995), the resources they use mainly consist of professional newspapers, practice oriented journals, and popular books; not of scientific journals (Beutler, Williams, & Wakefield, 1993). Furthermore, scientific theories and studies reported in scientific journals do not answer the questions psychologists in clinical practice have, such as how to treat patients with multiple disorders or how to resolve clinical impasses (Persons & Silberschatz, 1998). Psychologists give more weight to their clinical experience and that of their colleagues than to empirical evidence in deciding upon treatment for individual cases (Beutler, et al. 1995; Stewart & Chambless, 2007). Thus, even though the intentions of scientists and psychologists about applying scientific findings to clinical practice are honorable, unscientific, and possibly unsound, methods seem to find their way into psychologists' practices (Dawes, 1996; Lilienfeld, Lynn, & Lohr, 2003).

The need for a scientifically based clinical practice was first stated at the Boulder Conference on Graduate Education in Clinical Psychology in 1949 (Committee on Training in Clinical Psychology, 1947). Clinical psychologists needed to be educated as scientists as well as practitioners, doing both research and clinical work. Thus, scientist-practitioners would use their own practices as experimental situations and their clients as subjects to scientifically investigate the phenomena they were interested in and they would report their findings in scientific journals. For example, they could examine the effectiveness of a treatment with a particular group of clients. However, the implementation of science and the scientific method into practice received little consideration and proved to be difficult (Shapiro, 2002). Consequently, the Boulder scientist-practitioner model has gradually been adapted to fit better with the demands of clinical practice. It became a more lenient model of the clinician as an applied scientist (e.g. Shapiro, 1967; Spengler, Strohmer, Dixon, & Shivy, 1995).

An applied scientist works scientifically in two ways: (1) by using validated methods of assessment or treatment when available, and (2) when lacking these methods, by applying the scientific method of observation, hypothesis generation and hypothesis testing (Newnham & Page, 2009; Shapiro, 1967). Because of a limited body of knowledge about disorders and possible treatments (cf. Stricker & Trierweiler, 1995) and insufficient well-validated methods of assessment and treatment (Cicchetti & Sroufe, 2000; Shapiro, 1985),

the first way is impossible in most cases and psychologists can do no better than follow the scientific method. The scientific method can be applied embedded in a problem solving approach. Psychologists work as applied scientists to find and implement a successful treatment for client problems. In the clinical process, psychologists follow a problem-solving or 'engineering' approach (Münsterberg, 1913; Sloves, Doherty, & Schneider, 1979; Van Strien, 1997). The engineering approach focuses on finding and implementing a solution to a problem and on the decision making process required to do so (Van Strien, 1997). In clinical psychology, it consists of defining and analysing client problems, designing a treatment, and implementing and evaluating a treatment. In each of these phases, psychologists apply the scientific method to gain knowledge which can be used in the next phase. The goal of the engineering approach is different from the scientist-practitioner model in that it does not aim for generalization of the knowledge generated in the process; the knowledge gained is specific for the problem at hand. Furthermore, it explicitly acknowledges the psychologist as an active participant in the research, instead of an objective observer, and encourages the use of the psychologist's experience in solving the problem.

In clinical practice, psychologists face the problem of deciding which treatment is most effective for a particular client with specific complaints and problems (Paul, 1967). In order to decide which treatment is best for a client, psychologists should thus perform an assessment of client characteristics, complaints and problems. About half a century ago, only 17% of the clinicians considered assessment relevant to treatment planning (Meehl, 1960). Most of them considered a psychologist's warmth, empathy and personality more important for treatment success. Currently, about 90% of psychologists performs assessment in their practices (Musewicz, Marczyk, Knauss, & York, 2009; Watkins, Campbell, Nieberding, & Hallmark, 1995). Over time, it has become a core and defining feature of clinical practice (Groth-Marnat, 2003).

Psychological assessment is the result of a diagnostic decision making process. In this process, psychologists should work as applied scientists to achieve a thorough analysis of the problem. They systematically gather information about the client, integrate this information with existing psychological, scientific knowledge into a coherent mental model of the client and test this model or parts of it (Nezu & Nezu, 1995; Tarrier & Calam, 2002; Van Aarle & Van den Bercken, 1992). The central idea is that diagnostic decision making is "a special case of the activity involved in the establishment of scientific explanations of human behaviour in general." (Van Aarle & Van den Bercken, 1992, p. 184). Therefore, the scientific method may be tailored to guide the diagnostic process (cf. De Bruyn et al., 2003; Van Aarle & Van den Bercken, 1992; Westmeyer, 1975). In the diagnostic process, an explanation should fit only one person, the client, instead of a group. Furthermore, the information gathered in the diagnostic process is not only used to explain past behaviour but is also needed to predict future behaviour, for example response to treatment. In addition, the diagnostic process depends on the psychologist's experience and training.

Following the scientific method within the engineering approach should help psychologists perform their tasks in a structured and careful manner while increasing the effectiveness of their practices and controlling possible sources of decision errors (De Bruyn, Ruijssenaars, Pameijer, & Van Aarle, 2003; Fernández-Ballesteros et al., 2001; Nezu & Nezu, 1995). Decision errors might occur because most tasks that psychologists perform in clinical practice require some form of subjective clinical judgement, whether these are decisions about which kind of data to gather, which tests to administer, or which therapeutic technique should be applied in a therapy session. In

these decision processes psychologists can be influenced by personal biases or experiences.

In this thesis, psychologists' diagnostic decision making processes and their relationship with treatment decisions are examined. Psychologists should follow a specific sequence of decisions during the diagnostic process to make sure that the right kind of information to form a mental model of the client is gathered, processed and tested (cf. Nezu & Nezu, 1995; Ruiter & Hildebrand, 2006; Witteman, Harries, Bekker, & Van Aarle, 2007). To what extent these decision steps are actually performed in clinical practice is unclear.

The diagnostic process as it should proceed is described first. After that, I will discuss the restrictions that clinical practice imposes on performing the diagnostic process optimally and the discrepancy between the diagnostic process in theory and in practice. Finally, I will present the research questions and an overview of the studies performed to answer these questions[1].

THE DIAGNOSTIC PROCESS

In the diagnostic process, information about the client's complaints, problems and background is gathered using several methods, such as interviews, tests or questionnaires. The aim of the diagnostic process is to form a mental model of the client's problems which includes an explanation of those problems, and to use this model as the basis for treatment decisions (Gough, 1971; Haynes & Wiliams, 2003). The mental model is the result of two processes: categorical diagnosis, or classification; and explanatory diagnosis, or case formulation (Witteman et al., 2007).

Classification

Classification includes a description of the client's problems and their severity as well as categorization of the client's problems into one or more mental disorders (De Bruyn et al., 2003; Krol, De Bruyn, & Van den Bercken, 1992). Classification of a mental disorder is based on assessment of client symptoms. Symptoms are indications of the presence of a disease or condition. They can be self-reported by the client or inferred by the psychologist from overt behaviour, affect, cognition, perception, or other characteristics (Kazdin, 1983). For a client to be classified as having a particular disorder, combinations of symptoms should be present; the diagnostic criteria for that disorder should be met.

Classifying client problems is helpful because it allows quicker and better prediction of future events or behaviour (Smith & Medin, 1981). For example, by knowing the kind of depression a client has, a psychologist is better able to estimate that client's risk of relapse (Kessing, 2003). Furthermore, classification restricts the search for possible explanations for the client's problems (Haynes, Spain, & Oliveira, 1993; Krol et al., 1992; Vermande, Van den Bercken, & De Bruyn, 1996). For example, by knowing that the client has a depression instead of an anxiety disorder, the number of possible causal mechanisms to be considered is reduced.

A limitation of classifying mental health problems is that the categories used are not always well-defined with clear boundaries (Cooper, 2004). The same symptoms can be

---

[1] In this thesis, the following terms are used interchangeably:
- *psychologist* and *clinician*;
- *(diagnostic) decision steps* and *diagnostic decisions*;
- *explanatory diagnosis, diagnostic formulation* and *case formulation*.

indicators of different disorders, e.g. sleeping problems are a symptom of both depression and anxiety disorder. To aid psychologists in distinguishing disorders from one another, several classification systems are available with symptom checklists, such as the International Statistical Classification of Diseases and Related Health Problems 10th edition (ICD-10; WHO, 1993) and the Diagnostic and Statistical Manual of Mental Disorders 4th revised edition (DSM-IV-TR; APA, 2000). Classification of client problems guides the generation of hypotheses about possible explanations for these problems, that is: case formulation (Krol et al., 1992).

Case Formulation

Case formulation consists of a causal explanation, relating the client's problems to factors that cause and sustain them, while taking the unique situation and characteristics of the client into account (Haynes & Williams, 2003; Kuyken, Fothergill, Musa, & Chadwick, 2005). A case formulation "… aims to describe a person's presenting problems and to use theory to make explanatory inferences about causes and maintaining factors that can inform interventions" (Kuyken et al., 2005). It is a useful tool to organize complex and at times contradictory information from a client. Several models from different theoretical perspectives have been proposed, each prescribing what should be included in a case formulation (e.g. Curtis, Silberschatz, Sampson, & Weiss, 1994; Haynes & O'Brien, 1990; Persons & Tompkins, 2007). Though differences between these models have been reported (Eells, 2007), they also have several aspects in common. A case formulation should consist of a description of the client's overt problem(s), disorder(s) or symptoms, a relevant developmental history of the client, an explanatory mechanism linking causal and maintaining factors that explains the problem(s), coping strengths and weaknesses and guides for intervention (cf. Bieling & Kuyken, 2003; Eells, 2007; Perry, Cooper & Michels, 1987).

Case formulation is helpful because it supports the linking of the client's problems to possible explanations and the assessment of which explanation fits a particular client best. Furthermore, it helps to establish the therapeutic relationship by creating a shared understanding with the client (Eells, 2007; Tarrier & Calam, 2002). Together, classification and case formulation determine treatment decisions by identifying client problems and underlying causal factors and mechanisms of change which can be matched to therapeutic methods and techniques (Haynes, 1993).

A structured and thorough diagnostic process which includes classification and case formulation should help psychologists make better treatment decisions (Nelson-Gray, 2003). However, so far, the expected benefit of such a systematic and thorough diagnostic process has not been established (cf. Witteman et al., 2007). Research showed that following a structured method for classification, such as structured interviews based on DSM-IV or ICD-10, does lead to improved classification decisions (Sartorius et al., 1993; see Garb, 2004, for a review). Therefore, a structured and thorough diagnostic process which includes both classification and case formulation could result in improved treatment decisions, especially for complex cases (Haynes & Williams, 2003; Kuyken et al., 2005) or when psychologists need to decide between multiple evidence based treatments (Nelson-Gray, 2003).

Performing the diagnostic process effectively is not as straightforward as it seems. The validity and reliability of psychologists' diagnostic judgements and treatment decisions are low (see Garb, 1998, for an extensive overview). This low validity and reliability seems to result from the restricting circumstances encountered in clinical practice (cf. Gambrill,

2005) and psychologists' use of mental short-cuts (heuristics) to cope with these circumstances (cf. Garb, 1998). In the next two sections, I will discuss the constraints of clinical practice and how these affect decision making processes and the use of heuristics by psychologists.

THE DIAGNOSTIC TASK IN CLINICAL PRACTICE

The diagnostic situation is complex and dynamic (cf. Klein, Orasanu, Calderwood, & Zsambok, 1993). The information gathered is often incomplete and ambiguous, problems can be explained by multiple causes, and the relation between diagnosis and treatment is far from obvious (Lichtenberg, 1997). In clinical practice, the diagnostic task is complicated by limited time to gather and interpret information, lack of an objective benchmark to assess decision accuracy about diagnosis, and insufficient instruments to assess problems and causal factors. These aspects influence psychologists' diagnostic processes differently and could thus lead to unwelcome differences in their treatment plans resulting in low validity and reliability of diagnostic decisions.

Time
Time pressure is intrinsic to the diagnostic process (Meehl, 1954). An interview with a client cannot be interrupted each time the psychologist would like to reflect on what was said by the client. Time pressure results in fewer alternative hypotheses being considered (Dougherty & Hunter, 2003; Thomas, Dougherty, Sprenger, & Harbison, 2008). Psychologists focus on only a few possible hypotheses and do not search elaborately for information to support or refute other hypotheses. Time pressure could thus result in missing important information or an inaccurate interpretation of the information.

Feedback
The diagnostic task is complicated further because a definitive criterion to evaluate the accuracy of a diagnostic decision is absent. Unlike the medical domain, where in most cases a pathologist can confirm or rule out physicians' diagnoses with high certainty, there is no ultimate test to verify the presence of a mental disorder. Therefore, there is no 'gold standard' against which to test the accuracy of a diagnosis. Psychologists thus receive minimal feedback on the accuracy of their diagnoses, and if they receive feedback it is often too late to be effective (Dawes, 1996; Garb, 1989). Lack of feedback seems to lead to decision errors (Dawes, 1996; Garb & Boyle, 2003). For example, psychologists' judgements of treatment success are likely to be biased because they usually only receive feedback about the clients who complete treatment. They do not receive feedback on those clients that drop-out for various reasons, who may recover just as well without treatment. This might lead psychologists to believe that treatment is always necessary to overcome problems. Different judgements about a client are likely to result in different decisions about the type of treatment for a client.

Instruments
To support psychologists in gathering and interpreting information and to counter undesired influences from time pressure and lack of feedback, diagnostic instruments have been developed. However, these instruments are either insufficient, for example the classification of problems with the aid of manuals such as the DSM (Caspar, 1997), or unavailable, for example the identification of relevant causal factors (Haynes, Spain, &

Oliveira, 1993). Although the DSM classification system has been criticized for low construct validity and reliability (e.g. Follette & Houts, 1996), the main criticism about applying the system in clinical practice concerns the categorical distinctions between disorders (Cooper, 2004). There is no evidence for natural boundaries between the categories (Borsboom, 2008; Widiger & Samuel, 2005), meaning that the symptoms of mental disorders overlap. The amount of overlap between disorders determines the ease or difficulty of inferring the presence of one disorder rather than another from a set of symptoms presented by a client. For most disorders, this overlap is substantial (Widiger & Samuel, 2005), thus complicating the process of making a diagnostic decision.

Also, knowing what disorder a client has is usually not enough to identify the relevant causal variables or to select a treatment. For most disorders, many possible causes can explain the symptoms even when the causal mechanisms are unknown (Haynes, 1993). To select a treatment, relevant causal factors for a particular client have to match the mechanisms of change of a treatment. As objective and validated instruments to assess these causal factors are lacking, psychologists have to rely on their own subjective judgements. This is a complex task even when causal theories about a disorder are available, because psychologists then have to differentiate between these, often competing, theories and find out which one fits a particular client best.

HEURISTICS

Even though it has not been empirically established yet, lack of time, of targeted feedback and of appropriate instruments seem to contribute to exceeding of the psychologists' cognitive capacities for information processing. In such situations, the likelihood of biased judgements increases and the quality of decisions decreases (cf. Faust, 1986). Especially in situations where the outcome is unknown and the stakes are high, such as the diagnostic process, these cognitive limitations are most apparent (Newell & Simon, 1972; Van Merriënboer & Sweller, 2009). Taking the task and its circumstances into account, psychologists face an unfeasible mission. To perform this mission to the best of their abilities, they develop mental short-cuts, also called heuristics (Garb, 1996; Tversky & Kahneman, 1974).

In unaided decision situations, such as the diagnostic situation, heuristics help to make quick decisions based on limited information (Gigerenzer, Todd, & The ABC Research group, 1999; Kahneman & Klein, 2009). Prescriptive decision theories warn against heuristic decision making. They assume an ideal decision maker who is fully informed, able to compute with perfect accuracy, fully rational and with plenty of time available (Klein et al., 1993). As psychologists are usually not fully informed, fully rational or able to make perfect calculations, decision theories fail to accurately predict their decision behaviour. Though heuristics also do not always accurately *predict* decision behaviour, they *describe* decision behaviour rather well compared to decision theories (Plous, 1993).

Heuristics have certain advantages: decisions can be made fast, because little cognitive effort is required, and decisions can be made using only part of the information available (Gigerenzer & Brighton, 2009). They rely on prior knowledge about certain events and their outcomes, acquired in a particular task or environment. Using heuristics can lead to successful outcomes. A study by Green and Mehr (1997) showed that by applying a heuristic strategy, unnecessary, excessive referral by physicians to a critical care unit decreased significantly. Physicians used only a few cues to determine whether a patient should be admitted to the critical care unit, while the expert system used in the study

weighted and integrated about 50 cues. Decision accuracy was not affected: physicians using a heuristic decision strategy performed similar to the expert system using all available information. In the Green and Mehr (1997) study the heuristic decision strategy was made explicit to the physicians: they knowingly reduced the amount of information and time needed to make a decision. When the physicians were offered to use either the expert system or their own decision strategies again, they continued to use the heuristic decision strategy. In the diagnostic task, psychologists might also, based on their experience, explicitly and knowingly reduce the number of decisions.

However, the use of heuristics can also lead to judgement bias and decision errors (Dumont, 1993; Garb, 1998). A study by Garb (1996) showed that psychologists used the representativeness heuristic in diagnostic decisions. The representativeness heuristic is said to be descriptive of psychologists' decision strategies when they make a decision about a client by comparing that client to another one, a stereotypical or prototypical client (Kahneman & Tversky, 1974). Psychologists in Garb's study reached a diagnosis by comparing the client's complaints and symptoms with those of a prototypical client. Likelihood ratings of disorders for a particular client were highly correlated with similarity ratings. For example, psychologists who judged that the current client was very similar to a psychotic client, also indicated that it was very likely that the current client was psychotic. Only 27% of the participants correctly classified the client problems; they adhered to DSM-IV criteria. Prototypes vary between psychologists (Krol et al., 1992) and also differ from DSM criteria (Garb, 1996), which seems to lead to differences in diagnoses.

The development and use of prototypes is an example of the implicit use of heuristics. Psychologists cannot consciously choose which prototypes are developed, unlike the deliberative use of heuristic decision strategies such as presented in the Green and Mehr (1997) study. The former type of heuristics are partially automatic processes and are unconsciously activated (cf. Glöckner & Witteman, 2010). Implicit use of heuristics might more often lead to decision errors than deliberate use of heuristics. Psychologists are unaware of the influence of these heuristics on their decision processes and thus unable to correct them if necessary.

Heuristics are valuable because they make the diagnostic task manageable for psychologists. Although the use of a heuristic might lead to a non-optimal decision for an individual case, people who use heuristics might perform quite well across many cases. For example, adopting a confirmation strategy, i.e. seeking information that confirms rather than falsifies a hypothesis, is often judged to be an erroneous decision strategy (e.g. see Dumont, 1993; Nickerson, 1998). However, it can be very useful to seek confirmation in situations where the occurrence of events is uncertain, feedback about events is probabilistic, or time pressure is high, such as the diagnostic situation (Klayman & Ha, 1987). In those situations, the diagnosticity of the information gathered is relevant. If the initial diagnosis considered by the psychologist is depression, it is more informative to find out whether a client has suicidal thoughts than to find out whether this client has anxiety complaints. Under time pressure, the verification of a hypothesis can then be more informative and successful than its falsification. Though the application of such a strategy easily leads to 'false positives', i.e. persons diagnosed as depressed who are actually not depressed, the cost of missing one diagnosis of depression could be considered more serious than the cost of further testing and treatment of a person who is not depressed. The pragmatic confirmation of a diagnosis which is judged to be most likely, based on experience, will eventually lead to the best possible outcome under those specific circumstances.

THE DIAGNOSTIC PROCESS AND DESIGNING TREATMENTS

In clinical practice, optimal performance of the diagnostic task is hampered by the complexity and the dynamic nature of the situation, and it is constrained by limited time and because the cognitive capabilities of the psychologists are exceeded. Understandably, in the diagnostic decision making process psychologists therefore also rely on resources other than the scientific method, such as their own beliefs about disorders and their causes (Kim & Ahn, 2002), the theoretical orientation within which they were trained (Witteman & Koele, 1999) and previous experiences with similar clients (Garb, 1996). Thus far, it remains unclear to what extent psychologists follow diagnostic models' prescriptions based on the scientific method in their practices.

Prescriptive diagnostic models are based on the assumption that following the scientific method within an engineering approach improves the decision outcome. Furthermore, a thorough and complete assessment of the client's complaints and problems is supposed to be essential for making an appropriate treatment decision (Eells, 2007; Fernández-Ballesteros et al., 2001; Haynes & Williams, 2003). These two assumptions taken together imply that the treatment plan depends on the outcome of the diagnostic decision process and that this outcome in turn depends on the kind of decisions considered and made during this process. This thesis focuses on the role of the diagnostic decision making processes in designing treatments and aims to answer two research questions derived from these assumptions:

1. *What characterizes the diagnostic decision making process in clinical practice?*
2. *What is the role of the diagnostic decision making processes in designing treatments?*

The answers to these research questions will provide insight into the influence of the constraints of clinical practice on psychologists' diagnostic decision making processes, into the treatment utility of the diagnostic process, and into the applicability of diagnostic decision models in clinical practice. Knowledge about the characteristics of psychologists' decision processes can be used for training, to improve the quality of treatment decisions, and for the development of tools supporting or improving their natural decision processes.

OUTLINE OF THIS THESIS

In this thesis I will describe four studies that aim to answer the two research questions from different methodological perspectives. These different methodological approaches are used to be able to verify results and overcome the limitations of one particular method.

In chapters 2 and 3, the first research question is addressed and the diagnostic process is examined by comparing psychologists' diagnostic processes to the decisions described in prescriptive theoretical models. The little research there is on psychologists' diagnostic processes has mainly focused on the personal descriptions of psychologists of their diagnostic process, for example through verbal protocols (Witteman & Kunst, 1997). A drawback of such studies is that the terms used by the psychologists to describe their diagnostic activities cannot be compared. Providing psychologists with a common language as a frame of reference has been advocated by Beutler (1991) to overcome these limitations. To be able to identify and compare the diagnostic activities I constructed a

questionnaire with lists of diagnostic decision activities prescribed by theoretical models, as frames of reference for the psychologists to make their diagnostic processes explicit. In chapter 2 the kind of decisions made by psychologists in the diagnostic process are described and compared to the prescribed decisions. In chapter 3 the sequence of decisions made, adherence to the prescribed sequence of decisions, and agreement among psychologists about the sequence of decisions are examined.

In chapter 4, both research questions are addressed: the diagnostic process and its relationship with the treatment decision. In this study, psychologists performed the diagnostic process in an authentic diagnostic situation. Most studies have used written case descriptions (such as the study described in chapters 2 and 3 of this thesis; but see also Eells, Lombart, Kendjelic, Turner, & Lucas, 2005; Hillerbrand & Claiborn, 1990) instead of more authentic assessment tasks. The use of written case descriptions creates an artificial situation because the task is often self-paced and complete case descriptions are available. Psychologists have practically unlimited time and resources to examine the case information and make a diagnostic decision. The use of a diagnostic interview and a stimulated recall procedure allows me to investigate how psychologists cope with the restrictions of time and resources in actual practice.

In chapter 5, I further investigate the second research question and examine which part of the diagnostic process predicts the treatment decision better: classification or case formulation. In addition, a specific part of the diagnostic process, case formulation, is investigated in further detail.

In the final chapter, the main findings of all four studies are summarized and discussed, the concept of a decision support tool is described, and suggestions for further research are made.

CHAPTER 2

Psychologists' Judgements of Diagnostic Activities: Deviations From a Theoretical Model

ABSTRACT

In this article we describe an investigation into the diagnostic activities of practicing clinical psychologists. Two questionnaires were filled in by 313 psychologists. One group of psychologists (N=175) judged the necessity of diagnostic activities; the other group (N=138) selected the activities they would actually perform. Results show that more participants thought that diagnostic activities were necessary than there were participants who intended to actually perform those activities. Causal analysis, by generating and testing diagnostic hypotheses to form an integrated client model with an explanation for the problem, was judged least necessary and would not be performed. We conclude that a discrepancy exists between the number and kind of activities psychologists judged to be necessary and they intend to actually perform. The lack of attention for causal analysis is remarkable as causal explanations are crucial to effective treatment planning.

The goal of psychodiagnosis is to understand the complaints of a client and to provide an indication for their treatment. In the psychodiagnostic process, information about the client's complaints, problems and environment is gathered in interviews and through tests, until a classifying and explanatory diagnosis is reached and treatment decisions can be made (De Bruyn, Ruijssenaars, Pamijer, & Van Aarle, 2003; Ruiter & Hildebrand, 2006). The goal of the psychodiagnostic process is to form an integrated picture of the client, with a problem description and an explanation for the problem, and to propose a possible treatment for the problem based on this integrated picture. Psychologists may use several methods to collect relevant information, such as diagnostic interviews, tests or questionnaires. The final diagnosis is the result of an integration of the information gathered and the decisions made along the way. Theoretical models have been developed to aid psychologists in organizing and judging the importance of client information. These models usually contain several sequential phases – from describing the problem to selecting a treatment method (De Bruyn et al., 2003; Vertommen, Ter Laak, & Bijttebier, 2005). This paper focuses on the question which diagnostic activities are considered theoretically necessary in diagnosing a client and which would be actually used. As further treatment planning depends on an accurate diagnosis and an effective diagnostic process, research into diagnostic activities can be used to improve both the diagnostic process and the diagnosis.

Since Meehl (1954) challenged the value of intuitive clinical judgement, prescriptive methods for collecting and interpreting information in psychodiagnosis have been proposed to counteract the low reliability and validity of diagnostic judgement (Garb, 1998). The central idea of prescriptive psychodiagnostic models such as the Diagnostic Cycle is that psychodiagnosis should adhere to the scientific method to obtain knowledge in psychology by generating and testing hypotheses (De Bruyn et al., 2003). The Diagnostic Cycle prescribes three phases: observations of the client, formulating and testing hypotheses about the problem and possible causes of the problem based on these observations, and an evaluation of the outcomes of testing these hypotheses (Van Aarle & Van den Bercken, 1999). For example, a psychologist may see a child who is easily distracted and at times aggressive. A hypothesis is generated about the origin of the aggressive behaviour and a test is performed showing that the child has limited social abilities. Based on studies that show that limited social abilities may result from deprived sensory stimulation in early development, the psychologist then hypothesizes that the child may have lacked physical contact in her early years. This hypothesis is confirmed by the child's parents who explain that due to an illness the child had to be physically restrained and was not to be cuddled for a short period after birth. The goal of formulating and testing hypothesized explanations of a client's problem is to make sure that a plausible explanation is found by explicitly considering and ruling out other possible causes, and consequently a focus in treatment can be selected on a firm foundation (De Bruyn et al., 2003). Identifying causal factors that affect the problem is necessary to plan effective treatment (Haynes & Williams, 2003). Although formulating an explanation for a problem is not always necessary to start treatment, it provides much needed insight to direct treatment if the problem is complex or the first choice treatment method is not working as expected and the intervention needs to be adjusted.

The problem with most prescriptive models, including the psychodiagnostic models, is that they are rather time-consuming. They propose strict and lengthy procedures which require a lot of mental effort (Van Aarle & Van den Bercken, 1999). Also, immediate feedback on the hypothesis testing process necessary to improve diagnostic performance is

lacking (Dawes, 1996; Garb, 1989). Psychologists receive minimal feedback on the accuracy of their diagnoses or on the quality of the hypotheses they generate, and if they receive feedback it is often too late to be effective. In clinical practice, cognitive and time limitations force psychologists to use their mental resources efficiently. Psychologists often generate mental short-cuts (heuristics) to quickly diagnose a client (see Garb (1998) for an extensive review of the use of heuristics in clinical psychology). Using short-cuts in reasoning is not uncommon in other fields. Research on solving chess and medical problems showed that chess players and physicians do not always adhere strictly to theoretical problem solving models to solve the problems they face (Boshuizen & Schmidt, 1992; Patel, Arocha, & Zhang, 2005). Several studies have compared the theoretical problem solving approach with the actual practice of chess players (see Ericsson & Lehmann (1996) for a review). Results showed that successful chess players did not extensively search for possible moves, as prescribed by the theoretical model, but rather selected moves based on cued recall from memory. In the medical field, it was assumed that physicians used some form of hypothesis testing in diagnostic problem solving (Elstein & Schwarz, 2002). However, empirical studies showed that physicians' diagnostic reasoning was also influenced by rapid pattern recognition processes (Lesgold, Glaser, Rubinson, Klopfer, Feltovich, & Wang, 1988). Deviations from a theoretical model are related to clinical experience. The reasoning strategies used by experienced professionals differed from those used by novices (Shanteau, 1988). Reasoning strategies thus seem to change as clinical experience increases and new ways to cope with time and cognitive limitations are created.

Empirical studies suggest that the same is true for clinical psychologists. As experience increases, they approach the psychodiagnostic process in a more flexible way, based on the clinical knowledge they have acquired in practice (Brammer, 1997; Bus & Kruizenga, 1989; Hillerbrand & Claiborn, 1990). Bus and Kruizenga (1989) showed that diagnosing a client becomes a routine process. They expected that the diagnostic process would follow the same procedure as scientific problem solving. However, the psychologists in their study seemed to gather information without any hypotheses or explicit goal. Also, recommendations could not be traced back to the diagnoses the psychologists formulated. This finding was confirmed by research by Witteman and Koele (1999), who found that there was no relation between the psychologists' arguments and the treatment proposals. Hillerbrand and Claiborn (1990) claimed that this routine process of psychologists is based on their knowledge organization. They argued that the psychologists' organization of their knowledge base changes through clinical knowledge they acquire in practice, which would result in clearer and more accurate problem representations. A more accurate problem representation could increase diagnostic accuracy. A study by Brammer (1997) confirms these findings. He found that more experienced psychologists asked fewer questions but that these questions were more often related to diagnostic categories. He argued that these questions were based on implicit theories psychologists had formed about the clients and that they used these questions to fill up the gaps in their theories. However, in these studies it remains unclear which steps are actually performed in the diagnostic process.

We aim to fill in the gap in the existing knowledge about clinical psychologists' diagnostic reasoning by comparing their actual diagnostic process, from registration to treatment selection, to the activities described in the theoretical models they are taught during training. The little research there is has mainly focused on the personal descriptions of psychologists about their diagnostic process, for example through verbal protocols (De Kwaadsteniet, Krol, & Witteman, subm.; Witteman & Kunst, 1997). A drawback of these

studies is that the terms used by the psychologists to describe their diagnostic activities cannot be directly compared. Providing psychologists with a common language as a frame of reference has been advocated by Beutler (1991) to overcome these limitations. This is what we undertake in this study. To be able to identify and compare the diagnostic activities we used lists of diagnostic activities prescribed by theoretical models as frames of reference for the psychologists to make their diagnostic process explicit.

The current study aims to establish which diagnostic activities clinical psychologists judge to be theoretically necessary and which activities they intend to actually perform themselves. A distinction is made between judgements of the necessity of diagnostic activities and the intention to actually perform these activities to control for possible social desirability effects. Several review and meta-analytical studies (Ajzen, 2001; Ajzen & Fishbein, 1977; Glasman & Albarracín, 2006) have shown that there is a difference between what people consider necessary and what they actually do. Although measuring the intention to perform activities is not equal to measuring the actual behaviour, it approximates the actual behaviour best.

METHOD

Participants

Participants for both questionnaires were 313 members of the Dutch Institute of Psychologists (NIP) mental health care division. The mean age of the participants was 44.29 years (*SD* = 11.21; *range* = 23-79 years). The majority of the participants had completed post-graduate education (87%), was a registered mental health care psychologist (32%), had a BIG-registration[2] (78%), worked part-time (53%) and was employed in mental health care (48%). The theoretical orientation of the majority of the participants was cognitive-behavioural (55%). They worked with adult clients (50%) and with clients with personality disorders. On average the participants spent most of their time treating clients, next on diagnosing clients, then on executive tasks and they spent least time on scientific research.

175 psychologists filled in the Questionnaire Necessary Activities (the NA-group) and 138 psychologists filled in the Questionnaire Performed Activities (the PA-group; see below: Materials). Except for clinical setting, with more psychologists working in a hospital in the NA-group than in the PA-group (χ^2 = 16.70, df = 7, *p* = .019), the groups did not differ on any other background variable.

Procedure

By email we invited all members of the NIP mental health care division to take part in the study. Participants who accepted the invitation were sent a second email with a hyperlink to one of the two web-based questionnaires (see below: Materials; Quaestio Survey Manager, 1993). The participants were randomly assigned to one of the two questionnaires.

Psychodiagnostic Model

Lists of diagnostic activities used in this study as frames of reference for responding

---

[2] The Individual Health Care Professions Act, known through the Dutch acronym as the BIG Act, regulates the provision of care by health care professionals. Only registered individuals may use the legally protected title. The register enables the expertise of the registered practitioners to be recognized by all.

were derived from the Diagnostic Cycle (De Bruyn et al., 2003). The DC was chosen because it provides a clear specification of the diagnostic activities a psychologist ought to perform. The wording used in the DC is based on generic terms recognizable both for participants educated with the DC and for participants educated before the DC was introduced. Also, the wording is similar to that in other Dutch theoretical models used in educational programs, such as the diagnostic model proposed by Vertommen, Ter Laak and Bijttebier (2005).

Based on De Bruyn et al.'s DC (2003), we distinguished six main categories and 63 diagnostic activities within the main categories (see Appendix A). The first main category, Registration (11 activities), has the objective to decide whether or not the assessment process is continued. The goal of the second main category, Complaint analysis (11 activities), is to identify and summarize the client's complaints and describe them in behavioural terms. In the third main category, Problem analysis (10 activities), the problematic behaviour of the client is explored and the problem is stated in objective, testable terms. In the fourth main category, Explanation analysis (11 activities), alternative diagnostic hypotheses are generated and tested so that an integrated picture of the client with an explanation for the problem can be formed. After that, a method of treatment is selected in the fifth main category, Indication analysis (15 activities). The final and sixth main category, Diagnostic Scenario (5 activities), has the objective to formulate a plan to answer the client's diagnostic questions.

Materials

We developed two web-based questionnaires. One questionnaire asked participants to judge the necessity of the diagnostic activities derived from the DC (the Questionnaire Necessary Activities), the other questionnaire asked participants to select the diagnostic activities they actually intend to perform in diagnosing a client (the Questionnaire Performed Activities), to be referred to as the NA-group and the PA-group.

Each questionnaire started with a description of the purpose of the study and the structure of the questionnaire. Then a case description was presented (see Appendix B). This case was selected to be recognizable for every participant; this was checked with three experienced psychologists. The participants had to keep this particular client in mind while filling in the questionnaire. The participants could also consult a list with explanations of the concepts used in the questionnaire.

The next part was different for the two questionnaires. The main categories and diagnostic activities within the main categories were both presented in a fixed randomized order to the participants. The NA-group was asked to "indicate, for each activity, to what extent you deem that activity necessary in diagnosing the client described in the case vignette" on a 4 point Likert-scale ranging from 'absolutely unnecessary' to 'absolutely necessary'. The PA-group was asked to: "select the diagnostic activities from each main category that you actually intend to perform with the client described in the case vignette". Activities the participants did not intend to perform could be skipped.

Both questionnaires contained 14 open-end and multiple choice questions about the background and job characteristics of the participant. These questions asked about gender, age, work experience, BIG-registration, part-time/fulltime appointment, clinical setting, theoretical orientation, client population, specialization in disorders, post-graduate education, and time spent on diagnosis, treatment, executive tasks, and scientific research. Each questionnaire ended with a request to participate in future research and thanking the participants for their cooperation.

Analysis

To facilitate the comparison of the results of the two questionnaires, the measurement scale of the Questionnaire Necessary Activities was adjusted. For this purpose, the response options "absolutely unnecessary" and "unnecessary" were recoded into "(absolutely) unnecessary". Likewise, "absolutely necessary" and "necessary" were recoded into "(absolutely) necessary".

To establish which diagnostic activities psychologists considered necessary and which activities they intend to actually perform, percentages were calculated. An independent samples t-test was performed to test for differences between the answers on the two questionnaires. To test for differences between main categories within each questionnaire, ANOVA's were performed. A Bonferroni procedure was used to maintain an overall significance level of .05.

Also, background characteristics considered theoretically relevant were selected and their influence on the selection of activities was investigated. Work experience, training, theoretical orientation and setting were entered into a multiple regression analysis.

RESULTS

Figure 2.1 shows the percentage of participants in the NA-group who considered an activity (absolutely) necessary (dotted line) and the percentage of participants in the PA-group who actually intended to perform that activity (straight line).
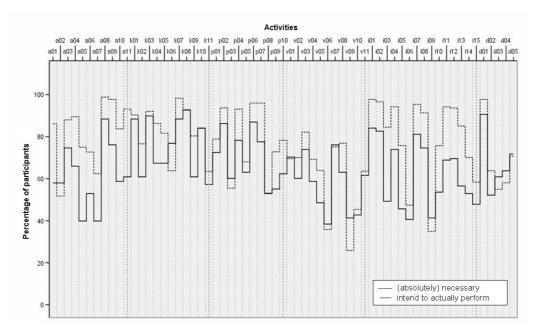

Figure 2.1. Percentages of participants who either judged an activity (absolutely) necessary (dotted line) and who intended to actually perform that activity (straight line), with the diagnostic activities (see Appendix A) on the horizontal axis.

The percentages of participants differed for the two questionnaires, as can be seen in Figure 2.1. Percentages of the NA-group are on average higher than percentages of the PA-group (76% and 65% respectively). This means that, for any activity, about three-fourth of the NA-group judge that activity (absolutely) necessary, while about two-third of the PA-group intends to perform that activity.

To compare main categories of activities, results from Figure 2.1 were comprised into an overview of these categories. Table 2.1 shows the mean percentages of participants for each main category, per questionnaire.

Table 2.1. Percentages of Participants and Standard Deviation for Each Main Category by Questionnaire Type.

	Questionnaire Type			
	Questionnaire Necessary Activities (N=175)		Questionnaire Performed Activities (N=138)	
Main Category	Percentage	Standard Deviation	Percentage	Standard Deviation
Registration	81.8	38.64	61.2	48.75
Complaint analysis	82.7	37.87	75.8	42.83
Problem analysis	78.6	41.06	69.5	46.06
Explanation analysis	61.6	48.65	57.7	49.42
Indication analysis	79.7	40.24	61.5	48.67
Diagnostic Scenario	69.0	46.26	67.8	46.75
Total	76.4	42.47	65.1	47.68

First, an independent samples t-test with percentages of the main categories as dependent variables and questionnaire type as a grouping factor was performed to test for differences between the two questionnaires. Significant differences were found for Registration ($t(299) = 6.64$, $p < .001$), Complaint analysis ($t(307) = 2.61$, $p = .01$), Problem analysis ($t(309) = 3.31$, $p = .003$), and Indication analysis ($t(306) = 6.48$, $p < .001$). It can be seen in Table 2.1 that the percentages of the NA-group are higher than those of the PA-group. This means that for the activities of these main categories a significantly larger part of the participants judged these activities necessary than participants from the other group intended to actually perform them.

Second, two ANOVAs were performed, one for each group, to test for differences between the main categories. The percentage of participants was the dependent variable and the main category was the fixed factor (six levels). The results will be discussed for the two groups separately.

For the NA-group, a significant effect of main category was found ($F(5, 10944) = 72.22$, $p < .001$). *Post hoc* analyses showed that Complaint analysis (83%), Registration (82%) and Indication analysis (80%) did not differ significantly from each other. Problem analysis (79%) differs significantly from Complaint analysis but not from Registration and Indication analysis. Diagnostic scenario (69%) and Explanation analysis (62%) differ significantly from every other main category. As can be seen in Table 2.1, the activities from the Complaint analysis, Registration, and Indication analysis were judged necessary by more participants than activities from the other main categories. The activities from the main categories Diagnostic Scenario and Explanation analysis are judged necessary by the least percentage participants.

For the PA-group also, a significant effect of main category was found (F(5, 8688) = 30.34, $p < .001$). *Post hoc* analyses showed that Complaint analysis (76%) differed significantly from every other main category. Next, Problem analysis (69%) and Diagnostic Scenario (68%) differed significantly from every other main category but not from each other. Indication analysis (62%), Registration (61%) and Explanation analysis (58%) also differed significantly from the other three main categories but not from each other. In Table 2.1 it can be seen that activities from the Complaint analysis would be performed by the largest part of the participants. Activities from the Indication analysis, Registration and Explanation analysis would be performed by the least number of participants.

It should be noted that the participants gave the activities from the Explanation analysis the lowest score on both questionnaires. This means that these activities are judged least necessary and that participants intended to actually perform them least often.

A multiple linear regression analysis was performed to investigate the influence of work experience, training, theoretical orientation, and setting on the percentages of participants selecting an activity. These predictors accounted for 10 % of the variance in percentages for the Questionnaire Necessary Activities ($R^2 = .099$), which was statistically significant (F(17, 9406) = 61.85, $p < .001$). For the Questionnaire Performed Activities, these predictors accounted for 7 % of the variance in percentages ($R^2 = .073$), which was statistically significant (F(17, 7164) = 34.35, $p = .001$).

CONCLUSIONS AND DISCUSSION

With the current study we aimed to investigate the diagnostic activities that psychologists in practice judge necessary and would actually perform. Results show that activities considered necessary and to be actually performed differ in number and kind.

In general, activities were more often judged necessary than that people would actually perform them. More specifically, more participants judged the activities from Registration, Complaint analysis, Problem analysis and Indication analysis necessary than there were participants who intended to actually perform these activities. It appears that what is considered necessary in theory is not always what would be done in practice.

Furthermore, the results show that activities from Registration, Complaint analysis, and Indication analysis were judged equally necessary, while activities from the Complaint analysis were most often intended to be actually performed. Activities from the Explanation analysis were judged least necessary and were also least likely to be actually performed. It seems that psychologists mainly focus on deciding whether or not to continue the diagnostic assessment process (Registration), identifying and summarizing the client's complaints (Complaint analysis) and on selecting a treatment method (Indication analysis). Generating and testing alternative diagnostic hypotheses to form an integrated model of the client with an explanation for the problem (Explanation analysis) gets much less attention.

The theoretical diagnostic model used as a frame of reference for the activities to be judged, the Diagnostic Cycle (DC), assumes that each part of the diagnostic process is equally important. Results show that the relevance and intention to actually perform the diagnostic activities were judged differently.

More specifically, the lack of focus on the Explanation analysis is noteworthy. An integrated model of the client including possible causal explanations for the problem behaviour, i.e. the end result of the Explanation analysis, is an essential condition for further treatment planning (Kendjelic & Eells, 2007; Krol, Morton, & De Bruyn, 2004; Kuyken, Fothergill, Musa, & Chadwick, 2005). While this is true theoretically, explanation

does not receive much attention from the participants in our study. A possible explanation could be that psychologists do not use causal reasoning to generate possible explanations of the problem behaviour. Psychologists could be building up a schema with explanations directly upon seeing the symptoms (Mayfield, Kardash, & Kivlighan, 1999). Recognizing the pattern of these symptoms might activate the schema's of the disorders, which include diagnostic explanations. Explicit causal analysis about explanations then becomes unnecessary. An alternative explanation could be that the participants use causal analysis implicitly. This explanation is supported by research by Kim and Ahn (2002) who found that psychologists' diagnostic reasoning is based upon personal, implicit causal theories about disorders. These causal theories may correspond to Brammer's (1997) implicit theories. Based on a few observations, psychologists appear to form a theory about the client's problem. They then use this theory to guide further information gathering (Brammer, 1997). These implicit theories preclude the necessity to explicitly reason causally. Thus, psychologists might use pattern recognition to see whether the pattern of complaints and problem behaviour of a specific client fits their personal, implicit, causal theory. If so, then explicitly generating and testing possible explanations would be redundant.

The regression analysis showed a significant influence of the background characteristics on the selection of activities and offers insight into the role of the psychologists' background on the decision making process. Nevertheless, this result needs to be regarded with some caution. The psychologists' background characteristics do determine the diagnostic decision making process to some extent. However, individual contributions of work experience, training, theoretical orientation, and setting to the diagnostic decision making process were not determined due to heterogeneity of the predictors used and limitations of the data collected. The influence of the individual predictors should certainly be explored further in future research.

It should be noted that there was a difference in clinical setting between the NA-group and the PA-group. As there were more psychologists working in a (general) hospital in the NA-group than the PA-group this might have resulted in differences in the decision making process, for example psychologists working in a hospital might be used to diagnosing more complex and severe problems.

Implications

Clinical psychologists do not seem to practice what they preach. By comparing their diagnostic activities to a theoretical model, the DC, we saw that one activity in particular seemed to be neglected: the explanation analysis. Since proper treatment planning depends on proper explanation, this activity should be the focus of further studies: when do psychologists engage in explanatory diagnosis, and what are the consequences for treatment planning both when they do and don't explicitly look for explanations of their clients' problems? Also, more attention could be paid to designing educational aids to training psychologists to follow the prescriptions of a diagnostic process model, and specifically to reason causally about their clients' complaints.

CHAPTER 3

Structuring Decision Steps in Psychological Assessment:
A Questionnaire Study

ABSTRACT

We investigated the structure of the diagnostic decision making processes followed by practicing clinical psychologists. Psychologists rank ordered decision steps they intended to perform in making a diagnosis. The first steps in psychologists' decision processes are identifying, summarizing and classifying the client's complaints and symptoms. However, the position of the causal analysis in the diagnostic process is unclear. Also, agreement among psychologists about the order of decision steps to be taken next and agreement with a prescriptive model is low. A trend is observed that as experience increases agreement decreases. We conclude that a prescriptive model is only partly used in practice, and that continuing education should remind psychologists of the prescription, especially to look for explanations and formulate an adequate treatment plan.

Psychodiagnosis is a complex decision making situation (Lichtenberg, 1997; Witteman & Kunst, 1997). Its aim is to form a mental model of the client's problems which includes an explanation of those problems, and to use this model to inform treatment decisions (Gough, 1971). The mental model is the result of two processes: classification, or categorical diagnosis, and explanatory diagnosis (Witteman, Harries, Bekker, & Van Aarle, 2007). Classification includes a description of the client's problems and their severity as well as categorization of those problems into a disorder (De Bruyn, Ruijssenaars, Pamijer, & Van Aarle, 2003; Krol, De Bruyn, & Van den Bercken, 1992). Classification guides the generation of hypotheses about possible explanations for the client's problems (Krol et al., 1992). Explanatory diagnosis consists of a causal explanation, relating the client's problems to factors that cause and sustain them (Haynes & Williams, 2003; Kuyken, Fothergill, Musa, & Chadwick, 2005). Together, classification and explanatory diagnosis guide treatment decisions (Haynes, 1993). It is crucial that correct psychodiagnostic decisions are made, since effective treatment is very important to the client's welfare. Treatment decisions depend on the outcome of the diagnostic process and the outcome of the diagnostic process in turn depends on the type and sequence of diagnostic decisions made during this process. How psychologists structure this diagnostic decision process is addressed in this paper.

Psychodiagnosis takes place in a suboptimal situation: it is an open-ended task in an environment with multiple, interdependent causal factors, in which information is often incomplete and ambiguous, and that usually proceeds under considerable time stress (Klein, Orasanu, Calderwood, & Zsambok, 1993). Methods to assist in the collection and interpretation of client information are either unavailable, for example the identification of relevant causal factors (Haynes, Spain, & Oliveira, 1993), or insufficient, for example the classification of problems with the aid of manuals such as the Diagnostic and Statistical Manual of Mental Disorders (Caspar, 1997). Thus, making a well-founded decision is fairly complicated. Only when clinicians use the same standardized diagnostic interviews to classify psychological disorders are they capable of achieving an acceptable level of inter-clinician reliability (Sartorius et al., 1993).

Prescriptive decision making models have been put forward to help psychologists to effectively organize and judge the information gathered in the diagnostic process, irrespective of theoretical backgrounds (e.g. Fernández-Ballesteros et al., 2001; Nezu & Nezu, 1995). Witteman et al., (2007) state that prescriptive models are called for, given the suboptimal situation. All of these models have several decision steps in common which are derived from more general problem solving and decision making steps of representing and understanding the problem, generating a solution, testing a solution and evaluating a solution (Newell & Simon, 1972; Pliske & Klein, 2003). An essential decision step in the decision process in general and in psychodiagnosis in particular is explaining the problem, because it helps to narrow down the number of solutions when more than one can be applied (Haynes & Williams, 2003). Eells, Lombart, Kendjelic, Turner & Lucas (2005) showed that psychologists who used a systematic model to organize and structure the information in their case formulations, produced higher quality case formulations. In a previous study (Groenier, Pieters, Hulshof, Wilhelm, & Witteman, 2008), we found that psychologists who do not use such a model focus on identifying complaints and problems, that is: categorical diagnosis, rather than on generating and testing alternative explanations for a client's problem: explanatory diagnosis (cf. Eells, Kendjelic, & Lucas, 1998).

Although decision making models can be useful for structuring the psychodiagnostic process, Van Aarle and Van den Bercken (1999) state that these models place a high

demand on the clinicians' cognitive capacities and applying these models in clinical practice might be difficult. Unwelcome differences in diagnoses and treatment plans might follow upon differences in the type and sequence of decision steps psychologists take in the diagnostic process, i.e. their decision strategy. A study by Bus and Kruizenga (1989) showed that psychologists' problem solving approaches for the same case vary and as experience increases psychodiagnosis becomes a routine process. This raises the question how psychologists structure the diagnostic decision making process, i.e. the constituent decision steps and their sequence. We address this question in the current paper.

Different decision steps in psychological assessment have been studied separately, such as classification (Krol et al., 1992), causal analysis (Eells & Lombart, 2003) and treatment decisions (Witteman & Kunst, 1997). However, the sequence in which psychologists perform these decision steps and the place of a causal analysis in the entire assessment process is still unclear. Conclusions drawn based on the outcome of one step are used as input for the next step. This output is necessary input for the next step. Therefore, the decision steps are logically connected and determine the kind of information gathered and decisions made. A more effective information gathering and interpretation approach increases diagnostic accuracy (Coderre, Mandin, Harasym, & Fick, 2003). Knowledge about the characteristics of psychologists' practice and structure of their decision making process can be used directly for the training of novices (Shanteau, 1988) and to improve the quality of treatment decisions.

In the current study, the diagnostic decision making process is investigated by comparing the psychologists' diagnostic decision steps with the steps of prescriptive models they were taught during training (cf. Groenier et al., 2008). To be able to identify and compare the participant's responses we used a list of decision steps prescribed by a decision making model, the Diagnostic Cycle (DC; De Bruyn et al., 2003), as a frame of reference. Providing psychologists with a frame of reference using a common language has been advocated by Beutler (1991) to overcome limitations of comparability of responses in studies that allow psychologists to generate their own steps.

We chose the Diagnostic Cycle as a survey instrument because it provides a complete and comprehensive inventory of decision steps that could be taken in the assessment process. Furthermore, the DC is explicitly not related to a particular theoretical or practical orientation. Psychologists from different theoretical backgrounds are thus able to recognize and identify the decision steps. Our previous study (Groenier et al., 2008) showed that theoretical orientation, together with other background variables, explained only 7 – 10 % of the variance in the number and kind of decision steps psychologists would take. Also, the DC encourages testing hypotheses and using scientific knowledge in every decision step.

We investigated the number, kind and sequence of decision steps psychologists would take in psychological assessment. Furthermore, psychologists' agreement about the sequence of decision steps of a theoretical model and their agreement with each other about the sequence of decision steps to be taken in psychological assessment were examined. In addition, the influence of experience on psychologists' agreement with a theoretical model and agreement among psychologists is examined. Although measuring the intention to perform decision steps is not equal to measuring the actual behaviour, it is a reliable approximation of the actual behaviour (Glasman & Albarracín, 2006).

Based on our previous findings we expect that a) an analysis of the complaints and problems of a client will be performed more often than a causal analysis (Groenier et al., 2008), b) psychologists will perform a causal analysis early in the diagnostic process as this

step is theoretically considered to be a crucial and informative step (Haynes, 1993), c) conformity with a prescriptive model's sequence of decision steps is low because psychologists focus only on the steps they judge important, d) agreement among psychologists about the sequence of decision steps is low because of the lack of a decision making model they can usefully apply in clinical practice (Van Aarle & Van den Bercken, 1999) and the use of individually developed decision strategies, and e) conformity with the sequence of decisions from a theoretical model and agreement among psychologists about the sequence decreases as experience increases because over time psychologists develop their own routines (Bus & Kruizenga, 1989).

METHOD

Participants

Using email, we invited psychologists from the mental health care division of the Dutch Institute of Psychologists (NIP) to take part in the study. Psychologists who accepted the invitation were sent a follow-up email with a hyperlink to a web-based questionnaire (see Materials below). Our sample consisted of 138 psychologists from all parts of the country.

The mean age of the participants was 43.8 years ($SD = 12.1$; *range* = 26 – 79 years). 26 participants had less than five years of work experience, 48 participants between 5 and 15 years and 64 participants over 15 years work experience. The majority of the participants had completed post-graduate education (84%), was registered mental health care psychologists (74%), was certified (78%), worked part-time (49%) and was employed in mental health care (49%). The theoretical orientation of the majority of the participants was cognitive-behavioural (54%). They mainly worked with adult clients (89%) with personality disorders. On average, the participants spent most of their time treating clients, somewhat less on diagnosing clients, then on management tasks. They devoted least time to scientific research.

Psychodiagnostic Model

Diagnostic decision steps used in this study were derived from the Diagnostic Cycle (DC; De Bruyn et al., 2003). We distinguished six decision steps. The first step, Registration, has the objective to decide whether or not the assessment process is continued. The goal of the second step, Complaint analysis, is to identify and summarize the client's complaints and describe them in behavioural terms. The third step, Diagnostic Scenario, has the objective to formulate a plan to answer the client's diagnostic questions. In the fourth step, Problem analysis, the problematic behaviour of the client is explored and the problem is stated in objective, testable terms. In the fifth step, Explanation analysis, alternative diagnostic hypotheses are generated and tested so that an integrated picture of the client with an explanation for the problem can be formed. After that, a method of treatment is selected in the final step, Indication analysis.

Materials

We developed a web-based questionnaire. This started with a description of the study's purpose and the questionnaire's structure. Then a case description was presented portraying a female client with either a depressive or an adjustment disorder (see Groenier et al., 2008, for a full description). The case was selected to be recognizable for every participant; this was checked beforehand with three experienced psychologists. The participants had to keep

the particular client in mind while filling in the questionnaire. The participants could also consult a list of explanations of the concepts used in the questionnaire.

The decision steps were presented to the participants in a fixed randomized order. Descriptions of the six decision steps were presented simultaneously on the computer screen. Participants were asked to rank order the decision steps according to which step they would perform first, second, third and so on. Those decision steps that participants did not intend to perform could be skipped.

The questionnaire further contained 14 open-ended and multiple-choice questions about the background and job characteristics of the participant, asking about gender, age, work experience, certification, part-time/fulltime appointment, clinical setting, theoretical orientation, client population, specialization in disorders, post-graduate education and time spent on diagnosis, treatment, management tasks and scientific research.

Analysis

First, percentages of participants selecting each decision step were calculated to examine differences in numbers and types of decision steps. A Cochran's Q test for binary data (selected vs. non-selected) was performed to test for differences between the decision steps.

Second, frequency distributions of the percentage of participants selecting a rank number for each decision step were calculated. The frequency distributions were inspected to determine the sequence of decision steps.

A consequence of participants selecting only those decision steps they intended to perform is that, for some participants, some decision steps did not get a rank score. Considering the type (non-random missing data) and amount of missing data (12%) in our study, we replaced the missing data and were able to compare rank scores with statistical techniques suited for complete datasets (Schafer & Graham, 2002). We replaced the unranked decision steps with the mean of the missing rank scores. For example, if a participant rank ordered four decision steps and skipped two, the two unranked decision steps were both assigned a rank score of 5.5. If a participant rank ordered 3 decision steps and skipped three, the three unranked decision steps were assigned a rank score of 5. The statistical analyses described next were all performed on the dataset with the missing data replaced.

Third, participants' rankings were compared with the prescriptive model's ranking to establish to what extent participants agreed with the sequence in which decision steps should be performed according to the prescriptive model. First, a Spearman's rank correlation was calculated between the model's ranking and each of the participants' rankings. Then, rank correlations were averaged across participants and across the three work experience groups.

Finally, the participants' rankings were compared to each other to establish to what extent participants agreed with each other about the sequence in which the decision steps should be performed. Kendall's W was calculated to assess agreement about the rankings of decision steps between all participants. Also, Kendall's W was calculated for each work experience group separately. Kendall's W is used because agreement between more than two raters is assessed. It can range from 0 (no agreement) to 1 (complete agreement) and can be interpreted as a correlation coefficient.

RESULTS

The total percentage of participants selecting a decision step and the percentages of participants selecting a particular rank number for a decision step are displayed in Table 3.1.

To examine differences in numbers and types of decision steps, percentages of participants selecting a decision step were compared. The percentages of participants selecting a decision step differed significantly between the decision steps (Cochran's $Q =$ 86.1, $df = 5$, $p < .001$). Complaint analysis was selected most often (98%) and Registration least often (71%).

Table 3.1. Percentage of participants assigning a rank number and skewness coefficients for each category.

Decision step	Rank number							Skewness	
---	1	2	3	4	5	6	Total	Skewness	SE
Registration	33	5	9	10	7	7	71	0.62	0.24
Complaint analysis	51	32	12	2	1	0	98	1.22	0.21
Diagnostic Scenario	4	13	19	20	14	8	78	- 0.01	0.23
Problem analysis	13	28	26	16	4	2	89	0.45	0.22
Explanation analysis	2	18	21	23	26	3	93	- 0.16	0.21
Indication analysis	1	4	10	18	23	40	96	- 0.88	0.21

To examine the sequence of decision steps the participants would perform, the frequency distributions of the rank numbers were inspected. The frequencies and frequency distributions of the rank numbers are different within and between the decision steps. The frequency distributions of Registration, Complaint analysis, Problem analysis and Indication analysis appear to be skewed, while the frequency distribution of Diagnostic Scenario appears to be normal and that of the Explanation analysis uniform. More participants assign a low rank number (1 or 2) to the decision steps Registration and Complaint analysis, these steps would be performed earlier on in the diagnostic process, and a high rank number (5 or 6) to Indication analysis, which would be performed later on in the diagnostic process. Skewness coefficients were calculated and tested to determine the degree of skewness of the frequency distributions (Field, 2000). The frequency distributions of the steps Registration ($z = 2.52$, $p < .05$), Complaint analysis ($z = 5.84$, $p < .001$), Problem analysis ($z = 2.06$, $p < .05$) are significantly skewed to the left and Indication analysis ($z = -4.16$, $p < .001$) is significantly skewed to the right.

There was a significant rank correlation between the model's ranking of the decision steps and the participants' rankings ($r_s(135) = .35$, $p < .001$). Rank correlations decreased as work experience increased. Rank correlations were .45, .36 and .30 for participants with less than 5 years work experience, 5 to 15 years work experience and over 15 years work experience respectively.

Agreement among participants about the sequence in which decision steps would be performed was low (Kendall's W = .34, $\chi^2 = 229.6$, $df = 5$, $p < .001$). Agreement decreased as work experience increased. Kendall's W was .39, .37 and .31 for participants with less than 5 years work experience, 5 to 15 years work experience and over 15 years work experience respectively.

DISCUSSION

In the current study, we investigated psychologists' opinions about following decision steps in the psychological assessment process. Almost all psychologists (98%) would perform an analysis of the client's complaints as a first decision step in the psychological assessment process, whereas at least a quarter of the participants would omit the decision step of whether or not to continue the assessment procedure. These results confirm our expectation that an analysis of the complaints and problems of a client will be performed more often than a causal analysis. The results support our previous finding that psychologists do not judge every decision step in the assessment process equally important (Groenier et al., 2008).

Looking at the sequence of diagnostic decision steps that psychologists would take, results show that they would first identify and summarize the client's complaints (Complaint analysis). Next, they would classify the problem (Problem analysis) and finally they would select an appropriate treatment method (Indication analysis). The position of the Diagnostic Scenario and Explanation analysis in the diagnostic process is less clear. It is noteworthy that some participants would generate and test alternative explanations for the client's problem (Explanation analysis) early on in the diagnostic process (18% of the participants select rank number 2), while others would perform this decision step late in the diagnostic process (26% of the participants select rank number 5). Deviating from the prescribed sequence of decision steps could result in a loss of information because the necessary input for the next step is missing and decisions are based on incomplete information. Further research is needed to gain more insight into the reason of the disagreement about the place of this decision step in the diagnostic process.

The low agreement with the prescriptive model's sequence of decision steps in our study indicates that this model is not a sufficient aid for structuring the diagnostic decision process in clinical practice. This finding supports the statement of Van Aarle and Van den Bercken (1999) that although prescriptive decision models are developed to help psychologists effectively structure information gathering and interpretation, in practice they are not applied as intended. Furthermore, the low agreement among psychologists about the sequence of the decision steps suggests that to some extent they develop their own personal decision strategies. Further research is needed to examine which steps psychologists actually take in diagnostic decision making and to what extent they develop personal decision strategies.

More specifically, we found a trend that agreement with a prescriptive model's sequence of decision steps and inter-clinician agreement both decrease as clinical experience increases. Adapting the decision process to the task and situation at hand is considered an important characteristic of expertise (Ericsson & Lehmann, 1996). As a result of adapting the decision process, experts deviate more from the sequence of decision steps they were taught than novices (Bus & Kruizenga, 1989) and are likely to disagree more with each other on the steps involved in the process than novices. Our results suggest that the diagnostic decision process does indeed become more divergent. With increasing experience, psychologists seem to adapt their diagnostic decision process to the task demands and to the specific situations they encounter in their clinical practices.

The difficulty with research that asks participants to rank order objects is that some participants may just leave objects unranked, for example because they rank only relevant, applicable or familiar objects, or objects worth ranking. The resulting missing data could be avoided by forcing participants to rank every single object. However, this is not always

desirable because the rankings of those objects may then not be meaningful. To be able to perform statistical analyses suited for complete datasets, we replaced the missing data in our study. Replacing missing data is more efficient than other methods, such as case deletion, because it helps to prevent a loss of statistical power resulting from a smaller sample size (Schafer & Graham, 2002).

A limitation of the current study is the use of only one case description. Psychologists might structure their diagnostic processes differently depending on case characteristics, such as complexity of client problems. However, the case description in the current study presented a client with problems that are commonly encountered in clinical practice.

A second limitation of the current study is the self-selection of the participants. This may have resulted in a sample that is not representative of the target population. Despite an estimated low response rate, the sample size of the current study was large enough for statistical analysis. Generalization of results to the population of interest should be done with caution.

Implications

By comparing the psychologists' approaches with a prescriptive model, namely the Diagnostic Cycle, we saw that they would identify, summarize and classify the client's complaints first and select a treatment method last in the diagnostic process, while the position of generating and testing alternative explanations for the client's problem in the diagnostic process is indefinite. This deviation from the prescriptive model was reflected in the low agreement with the prescriptive model's sequence of decision steps and low agreement among the psychologists. Psychologists' diagnostic processes seem to diverge more as experience increases.

To make sure that a helpful and structured method, such as the DC, is actually used in practice, it is crucial that psychologists are aware of their deviations from such a method. Regular reflection, supervision and peer-to-peer discussions at every experience level are recommended to help psychologists identify and, if necessary, eliminate these deviations.

CHAPTER 4

Clinicians' Judgements: Decisions During a Diagnostic Interview

ABSTRACT

In mental health care, clinicians' treatment decisions depend on assessments of the clients' complaints, problems and causes for these problems. The quality of the assessment depends on clinicians' decision processes. In the current study, we examined (a) the decisions made by clinicians during a diagnostic interview with a simulated client and (b) to what extent these decisions are related to assessment reports based on these interviews. Subsequently, stimulated recall was used to assess the clinicians' reflections on these decisions. We noticed that clinicians considered the client's complaints, possible classifications, explanations, and treatment options right from the start. They agreed more with each other about classifications than about causal factors and treatment options. Assessment decisions described in the reports were only partly related to decisions made during the interview. We conclude that clinicians' decision processes deviate from the decision process described in prescriptive diagnostic decision models.

Psychological assessment is the result of a decision making process aimed at describing, classifying, explaining, predicting and often also changing the behaviour of a client (Fernández-Ballesteros, 1999). Traditionally, 'psychological assessment' is used as a synonym for psychological testing. However, in line with Fernández-Ballesteros et al. (2001), we understand the assessment process to include the use of different instruments, such as tests, interviews and observations, as well as different activities, such as diagnostic classification and diagnostic formulation. The terms 'assessment', 'diagnosis', and 'formulation' often refer to different activities (cf. Westhoff, Hagemeister, & Strobel, 2007). We refer to 'assessment' as the decision making process of finding out what the client's problem is and of selecting an appropriate treatment. We refer to 'diagnostic classification' as the decision of classifying client problems into one or more disorders (a DSM-IV or ICD-10 classification) and to 'diagnostic formulation' as the decision of formulating an explanation for client problems.

The result of the assessment process is an integrated picture of the client's complaints and problems with an explanation for the problems and a treatment proposal. Treatment selection and planning depends on the outcome of the assessment process (Nelson-Gray, 2003). Therefore, it is crucial that a thorough assessment is carried out and correct diagnostic decisions are made. Previous studies showed that psychologists pay less attention to the decision process of diagnostic formulation than to diagnostic classification or treatment selection (Groenier, Pieters, Hulshof, Wilhelm, & Witteman, 2008). Also, a study by Eells, Kendjelic, and Lucas (1998) showed that psychologists' diagnostic formulations mainly contained descriptive information about complaints and symptoms. Furthermore, in an extensive review, Garb (1998) concluded that reliability and validity of diagnostic classifications are low. Parts of the psychological assessment process appear not to be carried out properly and consequently the quality of the outcome might be inadequate. So far, it is still unclear which decisions psychologists consider in actual practice, in which sequence they perform these decisions, and to what extent psychologists base their assessment reports on these decisions.

A thorough and complete assessment of the client's complaints and problems is supposed to be essential to making an appropriate treatment decision (Eells, 2007; Haynes & Williams, 2003). The central idea of naturalistic decision theories (e.g. Klein, Orasanu, Calderwood, & Zsambok, 1993) and diagnostic decision models (e.g. Nezu & Nezu, 1995) is that in psychological assessment, psychologists should adhere to the scientific method of obtaining knowledge by generating and testing hypotheses and should perform diagnostic decisions in a specific sequence. Diagnostic decision models prescribe a set of distinct subsequent diagnostic decisions for a thorough and complete assessment. Though there are differences between these models (e.g. see Groth-Marnat, 2003; Nezu & Nezu, 1995), they also have several aspects in common. First, the client's complaints and referral question are identified and analyzed. Second, the severity of the client's problems is assessed and the problems are grouped into a disorder (diagnostic classification). Third, possible explanations for the problems are generated, tested, and the psychologist forms an integrated client model (diagnostic formulation). Diagnostic formulation models from various theoretical orientations emphasize different kinds of causal factors and mechanisms (see Eells, 2007, for an overview of formulation models). For example, behavioural therapy models emphasize relationships between antecedents and consequences, cognitive therapy models emphasize dysfunctional attitudes and thoughts, and psychodynamic models emphasize core conflictual relationships (cf. Persons, 1991). However, the main goal of the case formulation decision is the same in every model; to identify and analyze causal factors

and mechanisms underlying client problems. Finally, a treatment is selected based on the integrated model.

A thorough assessment including these decisions requires a considerable amount of time and time is limited in clinical practice. In addition, the information gathered is often incomplete and ambiguous, problems can be explained by multiple causes, and the relation between diagnosis and treatment is far from obvious (Lichtenberg, 1997). These characteristics make the psychological assessment process complex and dynamic (cf. Klein et al., 1993). In these kinds of task situations, cognitive capacities are easily exceeded, as was shown in various professional task situations (Ericson & Lehmann, 1996). Decision makers will then resort to strategies that limit the number of decisions made and thereby reduce the amount of processing capacity needed (Garb, 1996; Gigerenzer, 2000).

Bus and Kruizenga (1989) showed that psychologists did not follow the scientific method of hypothesis generation and testing. Psychologists seemed to rely more on routine and did not gather information based on hypotheses or specific goals. More recently, a study by Groenier et al. (2008) showed that psychologists' diagnostic processes deviated from a theoretical model's prescriptions in the number and kind of decisions. Psychologists would focus more on diagnostic activities related to analyzing the client's complaints and problems and selecting a treatment method and less on activities related to diagnostic formulation. Thus, psychologists' decision processes seem to be adapted to the constraints of clinical practice.

The prescriptive diagnostic models are based on the assumption that following the scientific method and a specific set of decisions improves the decision outcome. However, the complex task situation, constraints of practice, and psychologists' naturally limited cognitive capacities would explain why they would not follow the prescriptions, presumably necessary for a well-founded treatment decision. Thus, research is needed to investigate whether psychologists' diagnostic decision processes do or do not follow the prescriptions from theoretical models. Empirical studies about psychologists' diagnostic decision making processes can contribute to the identification of successful processes. A first step, in the current study, is to investigate psychologists' naturally occurring decision processes. A second step, for future research, would be to examine the effectiveness of these decision processes.

Knowledge about psychologists' diagnostic decisions in actual practice is scarce: most studies have used written case descriptions (such as Eells, Lombart, Kendjelic, Turner, & Lucas, 2005; Hillerbrand & Claiborn, 1990) instead of more authentic assessment tasks. The use of written case descriptions creates an artificial situation because participants can work at their own pace and complete case descriptions are available. In these studies, psychologists have practically unlimited time and resources to examine the case information and make a diagnostic decision. We used an authentic assessment situation resembling clinical practice as much as possible, which allows us to investigate how psychologists cope with the restrictions of time and resources in actual practice and to examine how they gather information to base their decisions on.

In this exploratory study, we investigate the decisions that psychologists make during an authentic diagnostic interview. The type and sequence of decisions made during the interview is examined and stimulated recall is used afterwards to achieve a reliable evaluation of the considerations that psychologists have (cf. Kagan, Krathwohl, & Miller, 1963). The diagnostic interview is constrained in time, lasting a maximum of at most 30 minutes. There are no time constraints during stimulated recall, therefore, the stimulated recall task situation resembles the artificial assessment tasks with written case descriptions

and results can be compared. Using an authentic diagnostic interview and stimulated recall allows us to examine the process of psychological assessment instead of just the outcome, as is often the case with written case descriptions. In addition, we investigate to what extent the decisions made during the interview are acknowledged in a written report based on the interview. Writing a report after a diagnostic interview is a common method to summarize considerations and to communicate these to others. The innovative design of the study allows us to study the assessment process in three different ways, based on the diagnostic interview, stimulated recall interview, and the written report, tracing the decision making processes of each subject.

We distinguish four types of decisions: complaint analysis, diagnostic classification, diagnostic formulation, and treatment selection (cf. Fernández-Ballesteros, 1999; Pliske & Klein, 2003; Van Strien, 1997). Type, sequence and content of these decision are investigated. The type and sequence of decisions made during the diagnostic interview under time constraints is compared with those made during the stimulated recall interview without any time constraints. We expect that, because of psychologists' cognitive limitations and time constraints, they do not follow the prescribed sequence of decisions during the diagnostic interview. We also expect that psychologists consider decisions in the stimulated recall interview they did not perform during the diagnostic interview. Furthermore, type and sequence of decisions made are examined for each interview separately. A study by Groenier et al. (2008) showed that in an artificial diagnostic situation (i.e. using a case description) psychologists judged complaint analysis and diagnostic classification more important than diagnostic formulation. We expect that psychologists focus more on complaint analysis and diagnostic classification than on diagnostic formulation during the artificial diagnostic situation in the current study, i.e. the stimulated recall interview. Furthermore, reflections on diagnostic classification, diagnostic formulation and treatment selection decisions during the stimulated recall interview are examined as well as descriptions of these decisions in the written reports. Studies by Persons and Bertagnolli (1995) and Kuyken, Fothergill, Musa, and Chadwick (2005) showed that psychologists agreed more with each other about descriptive information, such as diagnostic classification, than about more inferential information, such as diagnostic formulation. We expect psychologists to agree with each other about diagnostic classification and to disagree about diagnostic formulation and treatment selection. In addition, the relationship between the content of the decisions considered during the interview and the content of the report written afterwards is explored. We hypothesize that the consistency of the content of decisions between the stimulated recall and the written report is low. The written report is a summary of decisions made during the diagnostic interview, while during the stimulated recall interview psychologists are more inclined to focus on the details of specific decision moments.

METHOD

Participants

Participants were recruited from a group of psychologists participating in a previous study who had indicated that they could be contacted for future research. Additionally, a new group of psychologists that was part of the active networks of the first and second author was invited to participate. Participants came from all over The Netherlands. We aimed for a group of participants from various backgrounds, differing in theoretical orientation, work setting, and years of experience. Totally, 27 psychologists participated,

aged between 22 and 65 years old ($M = 44.9$, $SD = 14.3$); 12 male (mean age = 54.9, $SD = 6.5$) and 15 female (mean age = 36.9, $SD = 13.8$). On average, the participants had 14.8 years of work experience ($SD = 12.8$, *Median* = 12.0, range = 0 – 35 years). We found no significant effect of experience on the type and sequence of decisions, therefore, this factor was excluded from further analyses. Participants either worked in a mental health care centre (74%) or in a private practice (26%), and they mainly worked with adults (89%). Participants were asked to indicate whether they worked from a specific theoretical orientation. Of the 27 participants, 15 indicated they did not and 12 participants indicated that they worked from one or more theoretical orientations. These twelve participants worked from the cognitive-behavioural orientation (8), the family systems orientation (3), client centred orientation (1), and the psychoanalytical orientation (1). Two participants worked from two theoretical orientations. They both worked from the cognitive behavioural and family systems orientation.

The Diagnostic Interview

All participants interviewed the same simulated client for at most thirty minutes. The interview was videotaped. After the diagnostic interview, participants were instructed to write a report (see Procedure). Allowing the participants to take notes during the interview enhanced authenticity.

The participants interviewed one actor portraying the same role. The role was written especially for this study. To maximize authenticity, the role was based on a real client and adapted to the actor's personal situation. It contained information indicating the classification depressive disorder without psychotic features. To ensure comparable performance across interviews, the actor received intensive training before the experiment, reread the script prior to being seen by each new participant and was given feedback about the consistency of his performance after each interview.

Procedure

Before starting the interview, participants received information about the client's name, gender, date of birth, residence, reason for referral by the family doctor (depressive and suicidal thoughts), the client's occupation, psychiatric history (none), marital status (married), number of children (two, a boy aged 16 and a girl aged 14), physical history (recently diagnosed with arteriosclerosis), current medication (no antidepressants) and the request from the family doctor for assessment and further treatment. Immediately after the interview, participants wrote a report summarizing their findings in a diagnostic classification (DSM-IV classification; APA, 2000), a diagnostic formulation, and a treatment proposal. After that, the stimulated recall session started.

In the stimulated recall session, one of the experimenters instructed the participants to watch the videotape of their diagnostic interview and to reflect on their own actions by reporting any thoughts they had had during the actual interview. At the start of the stimulated recall session, the experimenter asked the participant two questions: 1) What did you think when you read the referral letter from the family doctor, and 2) What did you think when you first saw the client? Each time the participants indicated that they remembered a thought, the experimenter stopped the video and the participant verbally reported their thoughts. When participants did not report any thoughts for at least two minutes, the experimenter reminded them to verbalize their thoughts. The stimulated recall session was recorded on videotape. The duration of the simulated recall session varied from 30 to 90 minutes.

Finally, participants filled in a questionnaire with questions about their background (gender, age, work experience, theoretical orientation, work setting, and time spent on treating clients). Afterwards, participants received a gift, i.e. a game or a gift certificate, for their participation.

Coding Diagnostic Interview

To evaluate the participants' questions and remarks during the diagnostic interview, we developed a coding scheme based on the Dutch Guideline for Psychiatric Assessment with Adults by Sno, Hengeveld, and Beekman (2004). We chose this guideline because it provides a complete and comprehensive inventory of all topics a clinician could discuss with a client during a diagnostic interview. Descriptions of the topics can be found in Appendix C. Six independent coding assistants (bachelor degree medical and psychology students) divided the participants' questions or remarks into meaningful units and consecutively assigned them to one of the coding scheme's categories with the computer program Sequence Viewer (Dijkstra, 2002). We defined a meaningful unit as a sentence or part of a sentence that expresses a single idea and receives only one code. A small unit of meaning was chosen to gather as much information as possible and to make sure that no important information was missed. Two "gold standards" of coded diagnostic interviews were created by the first and second author. One of the gold standards was used to train the coding assistants and the other to assess each coder's reliability afterwards. Because of differences in the number of meaningful units identified by the coding assistants and that of the gold standard, Cohen's Kappa could not be used to determine each coder's reliability. Therefore, we used Sequence Viewer's Delta (Dijkstra & Taris, 1995) which takes differences in the number of meaningful units identified into account. Delta varied from 0.75 to 0.85 and was considered satisfactory.

Coding Stimulated Recall Interview

To evaluate the participants' reflections during the stimulated recall interview, we developed a coding scheme based on the Diagnostic Cycle (DC) by De Bruyn, Ruijssenaars, Pameijer, and Van Aarle (2003). We chose the DC because it provides a complete and comprehensive inventory of the decision making activities psychologists could consider. Descriptions of the categories can be found in Appendix D. Two independent coding assistants (bachelor degree psychology students) divided the participants' reflections into meaningful units with the computer program Sequence Viewer (Dijkstra, 2002). A small unit of meaning was chosen to gather as much information as possible and to make sure that no important information was missed. Again, two "gold standards" of coded stimulated recall interviews were created by the first and second author. One of the gold standards was used to train the coding assistants and the other to assess each coder's reliability. The reliability of the combined segmentation and coding proved to be low. Therefore, we decided to separate the segmentation and coding processes. Percentage coder agreement about segmentation was calculated according to the method proposed by Strijbos, Martens, Prins, and Jochems (2006). Percentages agreement varied from 77 to 90 % which was quite satisfactory. After that, two other independent coding assistants (bachelor degree psychology students) and the first author assigned the meaningful units to one of the coding scheme's categories. We calculated each coder's reliability by comparing each coding assistant's categorization of the same interview with the "gold standard" created by the first and second author. To determine each coder's reliability we used Cohen's Kappa, which was .75 for each coding assistant and was considered satisfactory (cf. Landis & Koch, 1977).

Content Coding Stimulated Recall Interview and Written Report

To be able to compare the content of the diagnostic classification, diagnostic formulation and treatment selection decisions in the stimulated recall interviews with the content of the same decisions in the written reports, the content of these decisions were coded into categories by the first author. In the stimulated recall interviews, the following categories were coded: (a) diagnosis and differential diagnosis (Diagnostic classification), (b) potential stressors and explanatory mechanisms (Diagnostic formulation), (c) treatment (Treatment selection). In the written reports, the following categories were coded: (a) the Axis I and II classifications in the written reports (Diagnostic classification), (b) stressors and predisposing experiences and explanatory mechanisms (Diagnostic formulation), and (c) the treatment methods and goals (Treatment selection). See Table 4.1 for a description of the categories.

Table 4.1. Descriptions of the Coding Categories for Each Type of Decision.

Type of Decision	Categories Diagnostic Interview	Categories Stimulated Recall Interview	Categories Written Report
Complaints & symptoms	Complaints & symptoms	Complaints & symptoms	N/A
Diagnostic classification	Diagnosis Differential diagnosis	Diagnosis Differential diagnosis	Axis I DSM-IV classification Axis II DSM-IV classification
Diagnostic formulation	Psychiatric history Family history Physical history Social history Biography Personality	Potential stressors Protective factors Explanatory mechanism	Stressors and predisposing experiences Explanatory mechanism
Treatment selection	Treatment	Treatment	Treatment method Treatment goal

Analysis

To facilitate comparison of the results of the two interviews (diagnostic and recall), categories from each interview were combined into the four types of decisions described in the introduction: Complaints & symptoms, Diagnostic classification, Diagnostic formulation and Treatment selection (see Table 4.1).

Types of decisions

To evaluate the number of questions, remarks and reflections about the four types of decisions Complaints and symptoms, Diagnostic classification, Diagnostic formulation, Treatment selection and their subcategories, we calculated the mean percentage meaningful units in each decision type.

Sequence of decisions

To evaluate the sequence of decisions that psychologists make and reflect on, we divided each diagnostic and stimulated recall interview into three equal parts: the beginning of the interview, the middle and the end. To do so, we calculated the total number of

questions, remarks or reflections a participant expressed during an interview. The first one-third of this total number constitutes the beginning of an interview, the second one-third constitutes the middle, and the final one-third constitutes the end. For each part and each type of interview, we examined the mean percentage meaningful units in each decision type.

Consistency of decisions: From interview to report

The content of the decisions made during the stimulated recall interview was compared with the content described in the written reports. We compared (a) the classifications mentioned in the stimulated recall interviews with the Axis I and II classifications from the written reports (Diagnostic classification), (b) the potential stressors and explanatory mechanisms with the stressors and predisposing experiences and explanatory mechanisms (Diagnostic formulation), and (c) the treatment methods and goals with the treatment proposals (Treatment selection).

Statistical analyses

Due to inexplicable malfunctioning of the digital compact discs used to record the diagnostic interviews, three of the diagnostic interviews and two of the stimulated recall interviews had to be removed from further analysis. This resulted in 24 participants for the diagnostic interview and 25 participants for the stimulated recall interview. To examine differences between the diagnostic interview and the stimulated recall interview in the mean percentage meaningful units of the four types of decisions, we used a Wilcoxon test. Also, for the diagnostic interview and the stimulated recall interview separately, we examined differences in the type and sequence of decisions within each interview. To do so, we performed Friedman tests on the mean percentages meaningful units of the categories. Wilcoxon tests were used to follow up on significant findings. A Bonferroni procedure was used to maintain an overall significance level of .05. Effect sizes were measured using *r*.

To examine the consistency of decisions from interview to report we compared the content of the types of decisions Diagnostic classification, Diagnostic formulation and Treatment selection of the stimulated recall with the content of these decisions from the written report. We calculated the number of participants that reflected on (stimulated recall interview) or described (report) a certain type of classification, potential stressor or predisposing experience, explanatory mechanism and treatment method or treatment goal. The percentage agreement between diagnostic classifications, diagnostic formulations and treatment methods and goals reflected on in the stimulated recall interviews and those described in the written reports was calculated as the number of classifications, formulations and treatments agreed on / (number of classifications, formulations and treatments disagreed on + number of classifications, formulations and treatments agreed on).

RESULTS

Due to technical problems while recording the diagnostic interviews (see Method), 24 of the 27 diagnostic interviews and 25 of the 27 stimulated recall interviews were used for further analyses. The mean total number of meaningful units is 195.2 (*SD* = 46.2, *range* = 109 – 290) for the diagnostic interview and 96.44 (*SD* = 40.76, *range* = 29 – 192) for the stimulated recall interview. The mean percentage meaningful units for each category and type of interview are described in Table 4.2.

Table 4.2. Mean Percentage Meaningful Units and 95% Confidence Interval (CI) for Each of the Coding Categories and Type of Interview.

Coding category	Diagnostic Interview (n = 24)			Stimulated Recall Interview (n = 25)		
	Mean	SD	CI	Mean	SD	CI
Complaints & symptoms	25.3	7.8	22.0 – 28.6	32.5	6.5	29.8 – 35.2
Diagnostic classification	1.6	1.6	0.8 – 2.2	4.8	4.3	3.0 – 6.6
Diagnosis	1.6	1.6	0.8 – 2.2	3.7	3.2	2.4 – 5.1
Differential diagnosis				1.1	1.6	0.4 – 1.7
Diagnostic formulation	25.9	7.1	22.9 – 28.9	11.4	6.4	8.7 – 14.0
Social history *	15.0	5.1	12.9 – 17.2			
Physical history *	5.6	3.2	4.3 – 7.0			
Family history *	2.5	2.8	1.3 – 3.7			
Biography *	1.6	2.8	0.4 – 2.8			
Psychiatric history *	0.8	0.9	0.4 – 1.2			
Personality *	0.3	0.5	0.1 – 0.5			
Potential stressors [+]				7.3	4.2	5.6 – 9.0
Protective factors [+]				2.2	2.5	1.1 – 3.2
Explanatory mechanism [+]				1.9	2.3	1.0 – 2.9
Treatment selection	4.1	3.4	2.7 – 5.5	7.1	5.2	4.9 – 9.2
Other	43.2	9.0	39.4 – 47.0	44.2	12.1	39.2 – 49.2

* Categories from the Diagnostic Interview
[+] Categories from the Stimulated Recall Interview

Differences in the types of decisions performed during the diagnostic interview and reflected on during the stimulated recall interview were examined. There were differences between the four types of decisions (Complaints & symptoms, Diagnostic classification, Diagnostic formulation and Treatment selection). A Wilcoxon test showed that for the Stimulated Recall Interview the percentages meaningful units were significantly higher for the categories Complaints & symptoms ($z = -3.29$, $p < .01$, $r = -.47$), Diagnostic classification ($z = -3.10$, $p < .01$, $r = -.45$), and Treatment selection ($z = -2.66$, $p = < .01$, $r = -.38$). The percentage meaningful units for the category Diagnostic formulation was significantly lower ($z = -4.29$, $p < .01$, $r = -.62$) for the Stimulated Recall Interview. The percentage meaningful units of the category Other (see Appendix C and D for a description) did not differ significantly ($r = -.06$).

Types of Decisions
Diagnostic interview
Complaints & symptoms, Diagnosis, Differential diagnosis, the six Diagnostic formulation categories and Treatment selection differed from each other in the mean percentage meaningful units ($\chi^2(9) = 165.51$, $N = 24$, $p < .001$). All participants asked or remarked at least once about Complaints & symptoms, about the Diagnostic formulation category Social history and about Treatment selection. Out of 24 participants, 17 discussed the classification with the client.

The mean percentages meaningful units of the six Diagnostic formulation categories differed from each other ($\chi^2(5) = 86.74$, $N = 24$, $p < .001$). Participants most often asked or remarked about Social history. Social history differed significantly from every other

Diagnostic formulation category (all *z*-values greater than -4.20, all *p*-values < .001, all *r* greater than -.61). Furthermore, all but one participant talked at least once about Physical history, 18 about Family history, 13 about Psychiatric history, 12 about Biography, and 8 about Personality.

Stimulated recall interview

Complaints & symptoms, Diagnosis, Differential diagnosis, the three Diagnostic formulation categories and Treatment selection differed from each other in the mean percentage meaningful units ($\chi^2(6) = 98.09$, $N = 25$, $p < .001$). All participants reflected at least once on Complaints & symptoms and the Diagnostic formulation category Potential stressors. Out of 25 participants, 21 and 22 participants reflected at least once on Diagnostic classification and Treatment selection respectively.

There were differences in the mean percentages meaningful units of the three Diagnostic formulation categories ($\chi^2(2) = 27.17$, $N = 25$, $p < .001$). Participants most often reflected on Potential stressors. Potential stressors differed significantly from Protective factors ($z = -4.10$, $p < .001$, $r = -.58$) and Explanatory mechanism ($z = -3.91$, $p < .001$, $r = -.55$). Out of 25 participants, 18 and 17 participants reflected at least once on Explanatory mechanism and Protective factors respectively.

Sequence of Decisions

Diagnostic interview

The mean percentages of questions or remarks for each of the four types of decisions and each part of the interview are displayed in Table 4.3.

Table 4.3. Mean Percentage Meaningful Units and 95% Confidence Interval (CI) for Each Type of Decision, Type of Interview and Part of the Interview.

Type of decision	*Part of Interview*								
	First part			Second part			Third part		
	Mean	SD	CI	Mean	SD	CI	Mean	SD	CI
	Diagnostic Interview (n =24)								
Complaints & symptoms	8.2	4.1	6.6 – 9.9	9.7	5.2	7.5 – 11.9	7.4	3.1	6.1 – 8.7
Diagnostic classification	0.3	0.7	0.0 – 0.6	0.4	0.7	0.1 – 0.7	0.8	1.0	0.4 – 1.2
Diagnostic formulation	7.9	3.4	6.4 – 9.3	10.7	5.2	8.5 – 12.9	7.3	4.2	5.6 – 9.1
Treatment selection	0.2	0.4	0.0 – 0.3	0.2	0.5	0.2 – 0.4	3.7	3.2	2.3 – 5.1
	Stimulated Recall Interview (n = 25)								
Complaints & symptoms	12.6	3.1	11.3 – 13.9	11.0	3.1	9.7 – 12.3	9.0	3.0	7.7 – 10.2
Diagnostic classification	1.8	2.2	0.9 – 2.7	2.0	1.9	1.2 – 2.8	1.0	1.4	0.4 – 1.6
Diagnostic formulation	3.2	2.3	2.3 – 4.2	4.2	3.7	2.6 – 5.7	4.0	4.1	2.3 – 5.7
Treatment selection	0.8	1.4	0.2 – 1.3	1.7	2.4	0.7 – 2.7	4.6	3.8	3.0 – 6.2

Significant differences were found between the four types of decisions in the first part of the diagnostic interview ($\chi^2(3) = 60.33$, $N = 24$, $p < .001$), the second part ($\chi^2(3) = 57.54$, $N = 24$, $p < .001$), and the third part ($\chi^2(3) = 42.93$, $N = 24$, $p < .001$). On average, in every part of the interview participants asked or remarked most about Complaints & symptoms and Diagnostic formulation. Complaints & symptoms and Diagnostic formulation did not differ significantly from each other for any part of the diagnostic interview (all $r < .07$). Complaints & symptoms and Diagnostic formulation differed significantly from every other category for the first two parts of the interview (all z-values > -4.11, all p-values $< .001$, all $r > -.59$). For the third part of the interview, Treatment selection did not differ significantly from Complaints & symptoms ($r = -.42$) or Diagnostic formulation ($r = -.39$). Furthermore, these three categories differed significantly from Diagnostic classification (all z-values > -4.11, all p-values $< .001$, all $r > -.59$).

Participants asked and remarked about all four types of decisions in every part of the diagnostic interview. Their remarks were mostly about Complaints & symptoms, though at the end of the interview questions and remarks about Treatment selection increased.

Stimulated recall interview

The mean percentages of reflections on each of the four types of decisions for each part of the interview are displayed in Table 4.3. Significant differences were found between the types of decisions in the first part of the stimulated recall interview ($\chi^2(3) = 56.22$, $N = 25$, $p < .001$), the second part ($\chi^2(3) = 43.66$, $N = 25$, $p < .001$), and the third part ($\chi^2(3) = 41.28$, $N = 25$, $p < .001$). On average, in every part of the interview participants most often reflected on Complaints & symptoms. Complaints & symptoms differed significantly from every other type of decision for each part of the interview (all z-values > -3.43, all p-values $< .001$, all $r > -.49$).

Participants reflected on all types of decisions in every part of the stimulated recall interview. They mostly reflected on Complaints & symptoms, though reflections on Treatment selection increased towards the end of the interview.

Consistency of Decisions: From Interview to Report.

First, we were interested in the content of the decisions Diagnostic classifications, Diagnostic formulation and Treatment selection that the participants reflected on during the stimulated recall interview and described in the written report. Second, we were interested to see whether the content of these decisions was related to the content of the decisions in the reports written afterwards.

Diagnostic classification

The participants reflected on eight different classifications for the client's problems during the stimulated recall interview and described 10 DSM-IV classifications on Axis I and II in the written report. In Table 4.4 the content of these classifications and number of participants reflecting on or describing each classification are displayed.

Most participants (84%) considered the same classification: depression. Seven participants (28%) reflected on a differential diagnosis (*range* = 0 – 4). Apart from depression, which was always stated as being present, other classifications (psychosis, anxiety disorder, manic episode, personality disorder, adjustment disorder, obsessive-compulsive disorder, and dysthymic disorder) were considered as differential diagnoses or were discarded during the interview. Percentage agreement between the classifications reflected on during the stimulated recall interviews and those described in the written reports was 38%.

Table 4.4. Content of Diagnostic Classifications Reflected on (Stimulated Recall Interview) and Described (Report).

Content of diagnostic classifications	Stimulated recall	Report
Depression, depressive episode or depressed mood	21	20
Psychosis or psychotic features	3	10
Anxiety disorder	3	
Manic episode	2	
(Avoidant) personality disorder or traits	2	5
Adjustment disorder	1	4
Obsessive-compulsive disorder or features	1	1
Dysthymic disorder	1	
Axis II diagnosis delayed		17
Problems at work		2
Loneliness		1
Life phase problems		1
Post-traumatic stress disorder		1

Diagnostic formulation

First, the content of the potential stressors reflected on and described was examined; next, the content of the explanatory mechanisms (see Method for a description of the diagnostic formulation categories). The participants reflected on 11 different potential stressors and 13 different explanatory mechanisms during the stimulated recall interviews. They described nine different potential stressors and seven different explanatory mechanisms in the diagnostic formulations. In Table 4.5 the potential stressors and explanatory mechanisms are displayed as well as the number of participants reflecting on or describing each stressor or mechanism.

Three participants did not reflect on the conflict at work or the client's physical illness. The majority of participants reflected on (stimulated recall interview) and described (report) more than one type of potential stressor (92%, *range* = 0 – 8, and 80%, *range* = 0 – 3, respectively). The minority of participants reflected on and described more than one type of explanatory mechanism (24%, *range* = 0 – 4, and 12%, *range* = 0 – 3, respectively). Percentage agreement between the types of Potential stressors reflected on during the simulated recall interviews and described in the written diagnostic formulations was 67%. Percentage agreement for the Explanatory mechanisms was 11%.

Treatment selection

The participants reflected on 13 different types of treatment methods and/or goals during the stimulated recall interviews and described 14 different types in the treatment proposals of the written reports. In Table 4.6 the different treatment methods and goals are displayed as well as the number of participants reflecting on or describing each method or goal.

Table 4.5. Content of Diagnostic Formulations Reflected on (Stimulated Recall Interview) and Described (Report).

Content of diagnostic formulation	Stimulated recall	Report
Potential stressors or predisposing experiences		
Conflict or problems at work	15	15
Physical illness	15	21
Personality traits	14	1
Genetic factors	11	7
Family of origin and/or affective neglect in childhood	9	2
Current role in own family	6	1
Experiences of loss	5	
Upcoming surgery	5	4
Cycling accident	3	2
Relationship problems	3	
Other neurological or physical conditions	2	
Lack of social support		1
Explanatory mechanisms		
Physical problems, head injury, or medication	9	
Inefficient coping with quicker, younger colleagues at work	3	
Life phase problems	3	
Losing social support or not accepting social support	2	
Dysfunctional thoughts	2	
Reactive to multiple current, but unidentified, stressors	2	2
Psychological conflict about father	1	
Psychological conflict about male identity	2	
Insecurity or fear about failing body	1	2
Burn-out	1	
Relationship problems	1	
Transgenerational problems	1	
Emotion regulation	1	
Setting high demands for oneself		3
(Avoidant) coping style		2
Reversal of caregiver's role		1
Psychological "hurt"		1
Fear of becoming like his mother		1

Nine participants (36%) reflected on more than one type of treatment method or treatment goal during the stimulated recall interview (*range* = 0 – 3). About half of the participants (56%) described more than one type of treatment method in their reports (*range* = 0 – 3), while only 22% described more than one treatment goal (*range* = 0 – 3). The majority of participants (70%) described antidepressant or antipsychotic medication in the written reports, while only two participants mentioned that topic during the stimulated recall interview. Also, participants more often described specific therapies (such as cognitive-behavioural therapy) in the written reports, while treatment goals were more often mentioned in the stimulated recall interviews. Percentage agreement between the treatment methods and goals reflected on during the stimulated recall interviews and those described in the treatment proposals was 37%.

Table 4.6. Content of Treatment Selection Reflected on (Stimulated Recall Interview) and Described (Report).

Content of treatment methods and goals	Stimulated recall	Report
Ask client about goals or expectations	6	3
Referral to psychiatrist	4	8
Antidepressant or antipsychotic medication	2	19
Strengthen client's own responsibility for solving problems	2	
Coping with sister's depression	2	
Strengthen role as a father	1	
Consult others and/or explore other topics	1	11
Suicide contract or attend to suicide risk	1	8
Hospitalization	1	
System or family therapy, (involve partner)	1	7
Crisis intervention	1	
Change daily routine	1	2
Coping with life phase problems	1	
Cognitive behavioural therapy		8
Interpersonal therapy		2
Supportive therapy		2
Psychotherapy not otherwise specified		1
Perform relaxation or activation exercises		4
Attend to negative thoughts or feelings		3

CONCLUSIONS & DISCUSSION

In the current study, we investigated the type, sequence and content of diagnostic decisions (i.e. complaint analysis, diagnostic classification, diagnostic formulation and treatment selection) that psychologists make and consider during an authentic diagnostic interview with time constraints. Furthermore, we investigated to what extent psychologists consider the same diagnostic classifications, diagnostic formulations and proposed treatments, and to what extent psychologists referred to the decisions made during the interview in reports written after the interview. The design of the study allowed investigation of the assessment process of the same psychologists in three ways, based on the diagnostic interview, stimulated recall interview and written report.

Psychologists' decisions and considerations could be traced from the first moment they met the client to a written report summarizing their findings.

Our results emphasize the adaptive nature of the diagnostic decision making process as a continuous and unstructured shifting between decisions based on the dynamics of the interview situation. Psychologists were considering the client's complaints and symptoms, possible diagnostic classifications, diagnostic formulations and treatment options right from the start of the interview and continued to address each of these decisions throughout the interview. Apparently, psychologists are not using any readily ascertainable decision strategy to structure their assessment processes. Their diagnostic processes seem to be highly unstructured. Also, they tended to agree highly with each other about descriptive information (such as classification), while agreement on more inferential information (such as explanatory mechanisms) seemed to be moderate to low. Diagnostic classifications, diagnostic formulations and treatment options considered during the interview were only partly related to the written reports.

It appeared that, during the entire diagnostic interview, psychologists mainly discussed and asked about the client's complaints and symptoms and about potential causal factors. On the other hand, afterwards the psychologists mainly reflected on the complaints and symptoms, throughout the stimulated recall interview, and much less on causal factors. These findings confirm to some extent previous findings from Groenier et al. (2008) showing that psychologists judge an analysis of the client's complaints and symptoms more important than diagnostic formulation. A possible explanation for the inconsistency between what psychologists do and what they reflect back on could be that they gather information for diagnostic formulation routinely but do not actually use it. Guidelines prescribe which information should be collected, including information relevant for diagnostic formulation. Psychologists may be merely paying lip-service to these guidelines. Another explanation for not explicitly analyzing information relevant to diagnostic formulation could be that psychologists are unable to integrate this information into complex, coherent causal models of their clients (cf. Eells et al., 1998) and thus are unable to explicitly reason about it.

Possible diagnostic classifications did not seem to be discussed or considered very often in any part of the interview. Apparently, almost all psychologists considered one possible diagnostic classification early on and reflected on few other diagnostic classifications. This finding is consistent with results from previous studies on psychiatric diagnoses (Garb, 1998; Haverkamp, 1993). One explanation would be that psychologists 'satisfice' (Simon, 1957): as soon as they find a classification that sufficiently describes the client's condition, they choose that classification and stop searching. Another possible explanation is that they focused on the client's depressive complaints as a reference point (anchoring; see for example Tversky & Kahneman, 1974) and did not sufficiently consider alternatives, in this case differential diagnostic classifications.

Psychologists in the current study considered a wide range of potential causal factors, explanatory mechanisms and treatment options for the same client. They agreed less with each other about this more inferential information than about descriptive information, such as the diagnostic classifications. These results are in line with results from a study by Kuyken et al. (2005) who found that reliability of diagnostic formulations decreased when more theory-driven inferences were made. It should be noted that psychologists apparently do not 'satisfice' when they think about potential causal factors and treatment options: multiple options are considered at the same time. In diagnostic classification, the goal is to categorize a client's problem(s) into, preferably, one disorder (APA, 2000), while in

diagnostic formulation multiple causes and their relations need to be taken into account to form a satisfactory explanation (Groth-Marnat, 2003; Haynes & Williams, 2003). Checklists and classification systems to support psychologists in classification, such as the DSM-IV (APA, 2000), are explicitly a-theoretical and not designed to identify possible causes and causal mechanisms responsible for the client's problems.

The low agreement among the psychologists about the kind of causal factors, mechanisms, and treatment options could result from the different theoretical orientations of the psychologists in the current study. Different theoretical orientations propose different causal factors and models to explain client problems (Eells, 2007). Differences in theoretical orientations could also explain the low agreement observed for treatment proposals in our study to some extent. Witteman and Koele (1999) found that psychologists' theoretical orientations partly explained their treatment decisions. Dirmaier, Harfst, Koch, and Schulz (2006) showed that therapy goals differed for cognitive-behavioural and psychodynamic psychologists. Thus, the diverse theoretical orientations of the psychologists in the current study might explain their low agreement on treatment methods and goals. Another explanation may be that psychologists' implicit, causal theories result in differences in causal factors and mechanisms considered (Kim & Ahn, 2002).

Psychologists' decisions described in the written reports were not clearly related to the decisions reflected on during the diagnostic interview. The number of diagnostic classifications, diagnostic formulations, and treatment options reflected on and referred to in the written report is low. What is striking is that the psychologists proposed diagnostic classifications, causal factors, causal mechanisms, treatment methods and treatment goals in the written report which they had not reflected on previously. Consistency of clinical judgement has been shown to be poor across different cases (Shanteau, Weiss, Thomas, & Pounds, 2002) and the current study suggests that this may be true even for the same case. A possible explanation for the differences in the content of decisions could be that the written report elicited a different task expectation compared to the stimulated recall interview. A written report is commonly used to communicate assessment findings to colleagues, which requires psychologists to make their considerations explicit in terms that are commonly shared (such as DSM-IV classifications) and easily understandable.

In the current study, an authentic assessment situation was used with a simulated client and time constraints to resemble clinical practice as much as possible. With a simulated client, realistic interaction is possible, non-verbal behaviour can be observed, and a therapeutic relationship can be established, in contrast to artificial task situations using written case descriptions. However, there are also some limitations to our methodology. Although stimulated recall is designed to measure psychologists' thoughts at the time of the actual diagnostic interview, thoughts after the fact may also be included. Therefore, the data are not an actual reflection of what happened, but are used to assess considerations. We assume that these considerations will not have changed much from the diagnostic interview and writing the report to the stimulated recall session. We did find differences between the considerations described in the reports and those expressed during the stimulated recall interview. This suggests that considerations might also have changed in between the diagnostic and stimulated recall interview. However, using alternative methods to assess considerations such as thinking aloud during the task would have disrupted the authenticity of the interview.

The use of only one case describing a single disorder limits generalization of the

results to other disorders. For future research, multiple cases presenting different disorders should be used.

Implications for Clinical Practice, Training and Research

The diagnostic decision process can be characterized as an adaptive decision process: psychologists analyze the client's complaints and symptoms, consider possible classifications, start a diagnostic formulation, and plan the treatment right from the start and continuously reconsider these decisions throughout the whole process. Agreement about inferential information seems to be low and psychologists' written reports are poorly related to the decisions considered while talking to a client. We additionally conclude that psychologists' decision processes deviate from the decision process described in prescriptive diagnostic decision models. Even though psychologists consider all major types of decisions prescribed in formal decision models, they do so in an apparently unstructured manner. Conclusions drawn from the outcome of one step should be used as input for the next step. Deviating from the prescribed sequence of decisions could result in a loss of information because the necessary input for the next step is missing and decisions are based on incomplete information.

To improve the transparency of psychological assessment process we suggest that psychologists more closely adhere to the types and sequence of decisions prescribed by theoretical diagnostic models. Thus far, the expected benefit of such a systematic and thorough diagnostic process has not been established (cf. Witteman, Harries, Bekker, & Van Aarle, 2007). However, a study by Sartorius et al. (1993) showed that following a structured method for classification, such as structured interviews based on DSM-IV or ICD-10, does lead to improved classification decisions. Therefore, a structured diagnostic process, including both classification and case formulation, could result in improved treatment decisions. Regular reflection, supervision, and peer-to-peer discussions at every level of experience are recommended to help psychologists identify and, if necessary, eliminate deviations from theoretical models. Also, more attention could be paid to designing educational aids that support psychologists in following the prescriptions of theoretical diagnostic models. For example, Kendjelic and Eells (2007) showed that the quality of diagnostic formulations could be improved with a short training using a structured diagnostic formulation method.

Future studies should focus on the effect of adhering to the prescribed sequence of decisions on the quality of the decisions made and on therapy outcome. Do psychologists who adhere to the prescriptions of diagnostic decision models make better treatment decisions than psychologists who do not? Is adherence to the prescribed types and sequence of decisions related to improved treatment outcomes? Also, the influence of theoretical orientation on the types and sequence of decisions made in the assessment process should be studied further.

CHAPTER 5

The Effect of Client Case Complexity on Clinical Decision Making

ABSTRACT

In mental health care, clinicians' treatment decisions are expected to be based on the formulation (that is: exploration of the causing and maintaining mechanisms) of the client's problems. Previous research showed two things: the quality of clinicians' case formulations is generally low, and it is unclear to what extent treatment decisions actually are based on case formulations. In this study, we investigated to what extent case formulations explain treatment decisions for simple and complex problems and we tested whether the complexity of clients' problems influences case formulation quality. Results show that case formulations are only weakly associated with treatment decisions, for both types of problems. Also, we found that case formulations for complex problems were of lower quality than those for simple problems. We conclude that clinicians give adequate case formulations when these are least needed, that is: for simple cases, for which an empirically supported treatment is available. Clinicians appear to base treatment decisions on descriptions of overt client symptoms rather than on case formulations.

Case formulation is an essential part of clinical decision making and "… aims to describe a person's presenting problems and to use theory to make explanatory inferences about causes and maintaining factors that can inform interventions" (Bieling & Kuyken, 2003). A case formulation is thought to be more informative for a treatment decision than the classification of the client's problem because it takes the unique situation and characteristics of the client into account (Haynes & Williams, 2003). It is proposed as an important aid to the design of treatments and a useful tool to organize complex and at times contradictory information from a client. However, constructing a reliable and adequate case formulation appears to be difficult. Previous research has shown that case formulations mostly contain descriptive rather than explanatory information (Eells, Kendjelic, & Lucas, 1998) and that reliability and quality is poor (Kuyken, Fothergill, Musa, & Chadwick, 2005; Persons, Mooney, & Padesky, 1995). A study by Witteman and Koele (1999) showed that clinicians' idiosyncratic explanations offered for client problems accounted for differences in the treatment decisions. Differences in treatment decisions found for the same client (Witteman & Kunst, 1997) might thus be explained by differences in content and quality of case formulations (De Kwaadsteniet, Hagmayer, Krol, & Witteman, 2010). In this paper we investigate two aspects of case formulation: their quality and the relation with treatment decisions, for both simple and complex problems.

Several models from different theoretical perspectives have been proposed prescribing what should be included in a case formulation (e.g. Curtis, Silberschatz, Sampson, & Weiss, 1994; Haynes & O'Brien, 1990; Persons & Tompkins, 2007). Though differences between these models have been reported (Eells, 2007), they also have several aspects in common. A case formulation should consist of a description of the client's overt problem(s), relevant developmental history, an explanatory mechanism linking causal and maintaining factors that explain the problem(s), coping strengths and weaknesses, and guides for intervention (cf. Bieling & Kuyken, 2003; Eells, 2007; Perry, Cooper, & Michels, 1987).

However, studies by Eells et al. (1998) and Kuyken et al. (2005) showed that clinicians' actual case formulations are of low quality. Quality is determined by the amount of causal information and level of integration of that information into a parsimonious, coherent, and meaningful account of a client's problems (cf. Haynes & Williams, 2003). These researches found that case formulations mainly contained descriptions of the client's overt symptoms and problems and that the participating clinicians insufficiently integrated this descriptive information into an explanatory mechanism. Furthermore, in the study by Kuyken et al. (2005) less than half of the case formulations met the requirements for adequacy of case formulation. Inadequate case formulations included much irrelevant information, lacked detail, and described few links of relevant information to problematic situations. Although thus far it has not yet been established whether case formulation indeed improves treatment outcome (cf. Nelson-Gray, 2003), Kuyken et al. (2005) convincingly argue that high quality case formulations enhance outcomes, especially with complex cases.

A possible reason for the low number of adequate case formulations could be that clinicians only construct a detailed case formulation when they think they need to. Persons and Tompkins (2007) and Haynes and Williams (2003) argue that in situations where a client does not sufficiently respond to the first choice treatment or when the client's problems are complex, clinicians should perform a more elaborate and specific case formulation. In case of less complex client problems, for example when an empirically supported treatment is available, a case formulation could be redundant and possibly

needlessly delay the start of treatment (Nelson-Gray, 2003). The poor quality of case formulations found in previous studies could have resulted from the use of case descriptions presenting clients with simple problems.

In the current study, we investigate the influence of complexity of client problems on case formulation. We define case formulation as a hypothesis about the causing and maintaining mechanisms explaining the clients' problems that is helpful for treatment planning. We examine whether a case description presenting a client with more complex problems leads to case formulations of higher quality than one presenting a client with less complex problems. We predict that for the more complex case clinicians' formulations will more often contain causal chains or models, describe relationships between causal factors, be relevant, consistent, specific, and testable, and contain modifiable factors and positive indicators for treatment. We further examine to what extent treatment decisions are actually based on explanatory mechanisms described in the case formulations, or whether other variables, such as the clinician's theoretical orientation or the classification, predict treatment decisions just as well or even better.

METHOD

Participants

In the study, 211 psychologists participated (67% female). Most of them were recruited by email via the Dutch Institute of Psychologists (NIP) to all members of the mental health care division. Others responded to a poster displayed at six local mental health care organizations. Participants who responded to the poster were sent a paper version of the questionnaire, those responding to the email were sent a reply with a hyperlink to the web-based questionnaire (see Materials). Upon returning a completely filled in questionnaire, participants received a gift certificate of €30.

On average, the participants were 45.0 years old ($SD = 10.9$, *range* = 23 – 74), had 16.1 years of experience ($SD = 9.3$, *range* = 1 – 44) and worked with clients 18.5 hours per week ($SD = 6.9$, *range* = 0 – 40). Almost all participants (96%) were certified mental health care psychologists. The theoretical orientation most often adhered to was cognitive-behavioural (35%; see Results for a more detailed description). Most participants (53%) were employed in a mental health care institute. Others had their own practice (27%), worked in a (psychiatric) hospital (8%), a forensic setting (7%), a nursing home (3%) or a rehabilitation centre (2%).

Materials

We developed a questionnaire with two vignettes of female clients presenting with problems of low or high complexity. The questionnaire started with a description of the study's purpose, the structure of the questionnaire, and instructions for completion. Then the first one of the two vignettes was presented.

The two vignettes were selected from five vignettes tested in a pilot study. The vignettes differed in how common, familiar and easy to treat (i.e. availability of an empirically supported treatment) the clients' problems were. These two vignettes perceived as least and most complex by the participants in the pilot study were chosen. The vignettes were based on actual clients; however, any identifying information was removed or altered so that the client's identity could not be inferred. In the least complex vignette a female client with panic complaints was presented, in the most complex vignette a female client with dissociative and depressive complaints was presented.

The vignettes were written in a standard psychological report format with the following sections: personal information, reason for referral, client complaints, psychiatric, somatic, and family history, current social context and mental status. Possible classifications and explanations for the problems were left out on purpose. The length of the two vignettes was approximately equal: 1178 words for the least complex vignette and 1334 for the most complex. The order in which the two vignettes were presented was randomized.

After each vignette, participants were asked to select the most likely classification from a list of 9 DSM-IV classifications (APA, 2000), to describe in their own words what they thought caused the client's problems, and to select a maximum of two from a list of 18 treatment methods. The lists of classifications and treatment methods were derived from participants' responses in a pilot study to open-ended questions about classification and treatment options. There was no limitation to the number of words participants could use to describe what they thought caused the client's problems. Then participants rated the complexity of each vignette and familiarity with the problems described on a scale from 1 (not very complex/familiar) to 10 (highly complex/familiar).

The last part of the questionnaire consisted of questions about the participants' personal background and job characteristics. These questions asked about gender, age, registration as a mental health care psychologist, theoretical orientation, work experience, setting, and average hours per week working with clients. The questionnaire ended with thanking the participants for their cooperation.

Analysis

To analyze the participants' responses, these responses were scored in categories. First, to be able to examine the relationship between classifications and treatment decisions, the 9 classifications and 18 treatment methods were clustered into fewer categories. We asked an expert (a practicing psychiatrist with over 25 years of experience) to make a categorization of the classifications and the treatment methods. The lists of classifications were different for the two vignettes, while the lists of treatment methods were the same. The results of these categorizations are described in the Results section. Second, the case formulations were coded for quality and content, see below.

Manipulation check

To verify the effect of manipulating the complexity of client problems, we calculated the mean complexity ratings for both vignettes and examined the number of participants who judged the vignette we intended to be most complex as indeed the most complex vignette.

Quality coding case formulation

To assess the quality of the case formulations we developed a coding schema based on the case formulation quality rating scale designed by Kuyken, Fothergill, Musa, and Chadwick (received in personal communication; see Kuyken et al., 2005) and work by Vermande (1995) and Vermande, Van den Bercken, and De Bruyn (1996). We distinguished eight quality categories: (1) form (simple/composite hypothesis, causal chain/model), (2) relations between causal factors, (3) relevance, (4) consistency, (5) specificity, (6) testability, (7) modifiability of causal factors (not modifiable, indirectly, directly), and (8) positive indicators for treatment. See Appendix E for a description of the categories.

The quality coding procedure consisted of two steps: first, the case formulations were divided into one or more explanatory hypotheses; second, these explanatory hypotheses

were coded for quality. An explanatory hypothesis consists of at least a sentence or part of an enumeration and is (part of) an argument (e.g. "One possible reason for the client's depression is his uncertainty at work."). Repetition of vignette information is not considered an explanatory hypothesis. The first and last author segmented the case formulations into explanatory hypotheses and then coded these explanatory hypotheses for quality. The inter-coder agreement of segmentation was calculated using the method proposed by Strijbos, Martens, Prins, and Jochems (2006). Percentages agreement ranged from 72% to 93% and were considered satisfactory. To assess inter-coder reliability of the quality coding of the explanatory hypotheses, we calculated Cohen's Kappas which ranged from .46 to 1.0 for the different categories. For the categories testability and specificity, Cohen's Kappas were unsatisfactory (.46 and .53 respectively). These categories were removed from further analyses. Cohen's Kappas of the other categories ranged from .62 to 1.0 and were considered satisfactory (cf. Landis & Koch, 1977). A copy of the coding schema is available from the first author.

Content coding case formulation

To examine the relation between case formulations and treatment decisions, we investigated the kind of explanatory mechanism described, if present. To assess these explanatory mechanisms we adopted part C of the Case Formulation Content Coding Method of Eells, Kendjelic, Lucas, and Lombart (received in personal communication, see Eells et al., 1998; Eells, Lombart, Kendjelic, Turner, & Lucas, 2005). This Part C distinguished nine categories: (1) psychological mechanism general, (2) problematic aspects/traits of the self, (3) problematic aspects of relatedness to others, (4) dysfunctional thoughts and/or beliefs, (5) problems to manage emotions, (6) defense mechanism or coping style, (7) skill, social learning or behavioural deficit, (8) biological mechanism and (9) social or cultural mechanism.

The content coding procedure consisted of two steps: first, the case formulations were divided into Content Units (CU); second, these CUs were coded for content. A content unit is a proposition on semantic grounds, that is a simple declarative sentence. The largest and most frequently occurring unit is a complete sentence. Sentences that contain several propositions can be divided further, depending on their meaning. The first and last author segmented the case formulations into CUs and then coded these CUs for content. The inter-coder agreement of segmentation was calculated using the method proposed by Strijbos et al. (2006). Percentages agreement ranged from 72 to 94% and were considered satisfactory. To assess inter-coder reliability of the content coding of the CUs, we calculated Cohen's Kappas which ranged from .57 to .60 for the different categories and were considered satisfactory (cf. Landis & Koch, 1977). A copy of the coding schema is available from the first author.

Statistical analysis

To test the effect of complexity on the quality of the case formulations, the mean proportion of explanatory hypotheses in which a quality category was judged to be present was calculated for each quality category and vignette. Paired sample *t*-tests were used to test for differences between the two cases in the mean proportion explanatory hypotheses that: (a) were described as simple/composite hypotheses or causal chains/models, (b) described relations between causal factors, (c) were relevant, (d) were consistent, (e) contained directly, indirectly or no modifiable causal factors, and (f) contained one or more or no positive treatment indicators. Effect sizes were measured using Cohen's *d*.

To test the relationship between the participants' theoretical orientations, the classifications, the case formulations, and the treatment decisions, Guttman's Lambdas

(Goodman & Kruskall, 1954) were calculated. Guttman's Lambda is an association index for nominal data that ranges from 0 to 1 and represents the extent to which one variable is of influence in the prediction of another variable. The value 0 means that the one variable is not at all of predictive influence, the value 1 that a perfect prediction can be made.

RESULTS

The vignette we had intended to be the most complex was indeed perceived as most complex by 93% of the participants (*mean complexity rating* = 7.9, *SD* = 1.2). Seven percent of the participants judged the two vignettes equally complex (less than 2 points difference on a scale from 1 to 10), and one participant judged the vignette we intended to be most complex as least complex. The mean complexity rating of the least complex vignette was 3.5 (*SD* = 1.6).

Categorization of Classifications and Treatment Methods

The classifications and treatment methods were categorized into fewer categories to facilitate analysis of the responses. This resulted in the following three classifications for the least complex vignette: (a) anxiety disorder (containing panic disorder with or without agoraphobia, specific phobia and generalized anxiety disorder), (b) depressive episode and (c) anxiety disorder with depressive episode (containing panic disorder with or without agoraphobia, specific phobia and generalized anxiety disorder, all in combination with depressive episode). Categorization resulted in the following five classifications for the most complex vignette: (a) post-traumatic stress disorder, (b) depressive episode, (c) dissociative disorder (containing depersonalization disorder, dissociative identity disorder and dissociative disorder NOS), (d) dissociative disorder with depressive episode (containing depersonalization disorder, dissociative disorder NOS and dissociative identity disorder, all in combination with depressive episode) and (e) post-traumatic stress disorder with depressive episode.

Categorization resulted in the following five treatment methods for both vignettes: (a) cognitive-behavioural therapies (containing cognitive therapy, mindfulness based cognitive therapy, behaviour therapy and cognitive-behavioural therapy), (b) psychodynamic therapies (containing psychodynamic therapy and supportive therapy), (c) gestalt therapy, (d) client centred therapy and (e) other therapies (containing creative therapy, body-oriented therapy, system therapy, family therapy, hypnosis therapy, medication, trauma therapy, problem-solving therapy, psychodrama and psycho-education).

Quality of Case Formulation

For the least complex vignette, a total number of 341 explanatory hypotheses were identified for the 211 participants, which equals an average of 1.6 hypothesis per participant. For the most complex vignette, a total number of 426 explanatory hypotheses were identified, which equals an average of 2.0 hypothesis per participant. Across both vignettes, 29% of the explanatory hypotheses were simple hypotheses, 17% composite hypotheses, 28% causal chains, and 26% causal models. Of the total number of hypotheses, 28% of the hypotheses contained a description of the relationships between causal factors, 75% was relevant and 92% was consistent. In 40% of the hypotheses no modifiable factors were described, in 24% indirectly modifiable factors, and in 36% directly modifiable factors. In 97% of the hypotheses no positive indicators for treatment were described. Table 5.1 summarizes the mean proportions of explanatory hypotheses

across participants for both vignettes in which each quality category was judged to be present by the coders.

Table 5.1. Mean Proportion Explanatory Hypotheses (Standard Deviations) for the Least and Most Complex Vignette and 95% Confidence Interval (CI) of the Difference in Mean Proportions.

| | *Vignette* | | | | | |
| | Least complex | | Most complex | | | |
Quality Category	Mean	(SD)	Mean	(SD)	n	CI
Form						
Simple hypothesis	.18	(.30)	.21	(.30)	208	-.08 - .03
Composite hypothesis	.13	(.26)	.15	(.27)	208	-.07 - .02
Causal chain	.29	(.39)	.25	(.34)	208	-.03 - .11
Causal model	.34	(.43)	.27	(.38)	208	-.00 - .14
Relations	.47	(.42)	.22	(.36)	196	.17 - .32
Relevance	.86	(.28)	.74	(.35)	197	.06 - .17
Consistency	.94	(.20)	.93	(.20)	196	-.02 - .05
Modifiability						
No modifiable factors	.27	(.38)	.43	(.38)	208	-.22 - -.09
Indirectly modifiable	.17	(.31)	.25	(.34)	208	-.14 - -.02
Directly modifiable	.49	(.44)	.20	(.32)	208	.22 - .35
Positive indicators						
No positive indicators	.90	(.26)	.85	(.28)	208	.01 - .09
One or more indicators	.04	(.17)	.03	(.12)	208	-.02 - .04

For the least complex vignette in comparison to the most complex vignette, the mean proportion of explanatory hypotheses that contained descriptions of the relationships between the causal factors was higher ($t(195) = 6.55$, $p < .001$, $d = .47$), the mean proportion of relevant hypotheses was higher ($t(196) = 3.77$, $p < .001$, $d = .28$), the mean proportion of hypotheses containing not modifiable causal factors was lower ($t(207) = -4.51$, $p < .001$, $d = -.31$) and the mean proportion of hypotheses containing directly modifiable causal factors was higher ($t(207) = 8.88$, $p < .001$, $d = .62$). There were no significant differences between the two vignettes in the mean proportions of explanatory hypotheses for the categories form, consistency, indirectly modifiable causal factors and positive indicators for treatment.

Treatment Decisions

The strength of the relationship between theoretical orientations, classifications, explanatory mechanisms, and treatment decisions was investigated using Guttman's Lambda (λ).

Theoretical orientation

The participants could indicate to which one of eight theoretical orientations they adhered: 35% of the participants had a cognitive-behavioural orientation, 31% an eclectic orientation, 14% a cognitive orientation, 9% a psychodynamic orientation, 5% a solution focused orientation, 2% a system-theoretical orientation, 2% a humanistic orientation and 2% a behavioural orientation. The theoretical orientation of the participants did not appear to be related to the treatment method they proposed for both vignettes (least complex: $\lambda =$

.015; most complex: $\lambda = .016$).

Classifications

For the least complex vignette, 80% of the participants selected anxiety disorder and 20% selected anxiety disorder with depressive episode. For the most complex vignette, 55.5% of the participants selected dissociative disorder with depressive episode, 30% selected dissociative disorder, 7% selected post-traumatic stress disorder, 7% selected post-traumatic stress disorder with depressive episode and 0.5% selected depressive episode. The classifications did not appear to be related to the treatment methods proposed, for either vignette. For both vignettes, λ was 0. However, this does not mean that the treatment methods were completely unrelated to the classifications. For the least complex vignette, 94% of the participants chose some form of cognitive-behavioural therapy (either cognitive, behavioural, mindfulness based or cognitive-behavioural therapy). For the most complex vignette, 76% of the participants chose other therapies (such as trauma therapy), irrespective of the classification they selected.

Explanatory mechanism

At least one explanatory mechanism was described by 77% and 69% of the participants for the least complex and most complex case respectively. The same results on the relation between explanatory mechanisms and treatment methods for the least complex vignette were found as those on the relationship between classifications and treatment decisions ($\lambda = 0$). Here too, almost all participants chose cognitive-behavioural therapy, irrespective of the explanatory mechanisms they had described. For the most complex vignette, the association between explanatory mechanisms and treatment methods is very weak ($\lambda = .012$). The type of vignette appeared to be associated with the treatment methods more than theoretical orientation, classification, or explanatory mechanism, although still weakly ($\lambda = .23$).

DISCUSSION

We investigated the influence of the complexity of client problems on the quality of psychologists' case formulations of these problems. Furthermore, we examined to what extent these case formulations were associated with psychologists' treatment decisions. Our results showed that the psychologists generated higher quality case formulations for a client presenting with less complex problems than for a client with more complex problems. Also, the case formulations were not associated more with the treatment decisions than the psychologists' theoretical orientations or the classifications they provided for the clients' problems.

Complexity of client problems appears to affect the quality of case formulations. Case formulations for complex client problems were less often relevant, and contained descriptions of the relationships between causal factors less often, not modifiable causal factors more often, and directly modifiable causal factors less often. It seems that psychologists provide adequate case formulations when it is least necessary: in less complex cases, for which an empirically supported treatment is available. The majority of psychologists indeed selected the recommended empirically supported treatment for the client with less complex problems.

In addition, these treatment decisions were in neither of the two cases associated with the kind of explanatory mechanisms the psychologists judged to be causing the clients' problems. Strikingly, individualized formulations were not associated with the psychologists' treatment decisions more than the classifications or the psychologists'

theoretical orientations. The clients' overt complaints were most highly associated with the treatment decisions, although still weakly. The results suggest that other factors, such as guidelines or institutional policy, might also have influenced treatment choice, especially for the least complex case.

The hypothesis of Kuyken et al. (2005) that especially for complex cases psychologists would give higher quality case formulations is not supported by our results. It seems that psychologists could construct an adequate explanation for the less complex case more easily. For this case, a client presenting with an anxiety disorder, a firm theoretical and evidence based framework describing underlying mechanisms of the disorder is available, whereas this is lacking for the more complex case, a client presenting with a dissociative disorder. Psychologists should thus be able to more easily retrieve and apply information from memory for the less complex case. This would suggest that psychologists' case formulations are schema driven.

Furthermore, the relationship between theoretically driven explanations and treatment decisions described by Witteman and Koele (1999) was not found in the current study. Psychologists rather appear to decide on a treatment based on the clients' patterns of complaints. This again supports the hypothesis that psychologists' case formulations are schema driven.

There are some limitations to our methodology. First, the two case descriptions used presented clients with different disorders. Differences in quality of case formulations could have resulted from aspects related to the specific disorders, not only from a difference in complexity. For future research, case descriptions presenting clients with the same disorder should be used, varying only in complexity.

A second limitation is using the treatment method as the dependent variable to examine the relationship between case formulation and treatment decision. The treatment method might not be detailed enough to reflect the differences in the explanatory mechanisms described in the case formulations. There might be a stronger relationship between case formulation and a specific treatment goal (cf. Persons, 2006), and the same goal can be reached with different treatment methods. We plan to examine this relationship in a follow up study.

A third limitation is the framing of the question. Different question formats have been used to investigate clinicians' case formulations. In the study by Eells et al. (1998), a document analysis was performed and it is unclear what the instructions to the staff were at the time of writing their reports. In the study by Persons et al. (1995), participants were asked to rate predefined underlying cognitive mechanisms, instead of describing them in their own words. In the study by Kuyken et al. (2005), participants provided case formulations using a specific diagram as part of a workshop on this particular topic. The use of these different question formats might explain differences in the results.

A fourth limitation is the use of only one expert to group the classifications and treatment methods. Using only one expert might have biased the resulting categorization. For future studies, we suggest that a panel of experts is used.

Finally, self-selection of the participants may have resulted in a sample that is not representative of the target population. Despite an estimated low response rate, the sample size of the current study was large enough for statistical analysis. Generalization of results to the population of interest should be done with caution.

Implications for Clinical Practice

The quality of psychologists' case formulations for a client's problems depends on the

complexity of these problems. We found that case formulations for a client with simple problems were of higher quality than those for a client with complex problems. Furthermore, these case formulations appear not to be related to psychologists' treatment decisions. We conclude that classification of the problems suffices for an initial treatment plan and a swift start of the treatment. In-depth analysis of possible causes explaining the client's problems might be more beneficial for determining the specific focus of the treatment later on.

CHAPTER 6

Summary of Findings and General Discussion

In this thesis psychologists' diagnostic decision making processes and their relationship with treatment decisions were investigated. Theoretical diagnostic models prescribe that the diagnostic process consists of two decision processes: classification and case formulation (Witteman, Harries, Bekker, & Van Aarle, 2007). Classification includes a description of the client's problems, their severity and categorization of the client's problems into one or more mental disorders (De Bruyn, Ruijssenaars, Pameijer, & Van Aarle, 2003; Krol, De Bruyn, & Van den Bercken, 1992). Case formulation consists of a causal explanation, relating the client's problems to factors that cause and sustain them (Hayens & Williams, 2003; Kuyken, Fothergill, Musa, & Chadwick, 2005). These diagnostic decisions are used to form an integrated client model on which psychologists base their treatment decisions (Gough, 1971; Haynes, 1993). However, in clinical practice, the complex and dynamic nature of the diagnostic tasks hampers an optimal performance (Gambrill, 2005). Limited time and exceeding the psychologists' cognitive capabilities constrain psychologists' performance (Garb, 1998). Thus far, it has remained unclear whether the prescribed diagnostic decisions are performed in clinical practice and determine treatment decisions.

Two research questions were investigated in this thesis:

1. *What characterizes the diagnostic decision making process in clinical practice?*
2. *What is the role of the diagnostic decision making processes in designing treatments?*

The studies in this thesis show that psychologists' diagnostic processes are adaptive and characterized by a focus on the client's complaints and symptoms and on treatment selection while they continuously shift between all diagnostic decisions and reconsider previous decisions. The decisions made in the diagnostic process contribute only slightly to treatment design.

Psychologists' diagnostic processes were examined with different methods to triangulate results and overcome limitations of using only one method. Written case descriptions, an authentic diagnostic interview and stimulated recall were used to study the diagnostic process and the role of diagnostic decisions in designing treatments. The stimulus material and tasks presented to the participants in the four studies varied from highly unstructured, an intake interview with a simulated client, to highly structured, a written case description with pre-structured response options.

First, the methods and main findings are summarized and the limitations and advantages of the methods used are discussed for each chapter. Next, the main findings are discussed. After that, the applicability of theoretical diagnostic decision models in clinical practice is evaluated. Finally, I describe implications for clinical practice and training, propose a decision support tool and discuss opportunities for further research.

METHODS AND FINDINGS

In chapters 2 and 3 the type of diagnostic decisions considered in the diagnostic process (chapter 2) and the sequence of decisions (chapter 3) was explored. An artificial task was used which consisted of a questionnaire with a written case description and pre-structured response options. The aim of these studies was to gain insight into the characteristics of psychologists' diagnostic processes.

Psychologists' Judgements of Diagnostic Activities

In chapter 2 the analysis is reported of the type and number of decision activities that psychologists judge necessary or would perform. One group of psychologists judged the necessity of diagnostic activities; another group indicated whether they would perform these diagnostic activities. Psychologists first read a case description and then filled in the questionnaire keeping this particular client in mind. Both questionnaires consisted of 63 diagnostic activities grouped into six main decisions, i.e. registration, complaint analysis, problem analysis, explanation analysis, indication analysis and diagnostic scenario. The diagnostic activities were derived from De Bruyn et al.'s (2003) Diagnostic Cycle (DC) because this provides a complete and comprehensive inventory of the decision activities psychologists could consider. To facilitate analysis, we aggregated the number of activities judged necessary or that would be performed across the main decisions.

Results showed that psychologists judged diagnostic activities to be necessary more often than that they would actually perform them. More specifically, more psychologists judged diagnostic activities related to client registration, analysis of complaints, symptoms and problems and treatment selection to be necessary than there were psychologists who intended to actually perform them. Furthermore, psychologists judged diagnostic activities related to generating and testing explanations for the client's problem behaviour (i.e. case formulation) least necessary and would not perform most of these activities.

Some of the participants reported feeling a certain discomfort when asked to respond to a written case description, especially with pre-structured response options. They argued that their responses are not a valid representation of their behaviour in actual practice because they were not able to talk to and observe the client themselves. Even though responding to a written case description is different from actually interviewing a client, psychologists are familiar with the written case description format. For example, in clinical case conferences decision making is based on a description of the client by one or more team members (see e.g. Pijnenburg, 1996). Also, the written case descriptions used in the studies reported in this thesis were based on actual clinical records and on information provided by psychologists who had assessed or were treating the clients.

In chapter 2 a task situation was used in which psychologists responded to a written case description. The task situation used is characterized by low time pressure and cognitive effort. Psychologists have ample time to study the presented information and they can work at their own pace. Cognitive effort is low because the relevant information is usually presented all at once. Psychologists do not have to think about how to gather relevant information or monitor how the client responds to questions asked.

Structuring Decisions

In chapter 3 the analysis is reported of the sequence in which psychologists perform their diagnostic decisions. In addition, psychologists' agreement with the sequence of decisions from De Bruyn et al.'s (2003) DC, agreement among each other about the sequence of decisions and the influence of experience on psychologists' agreement was investigated. Psychologists rank ordered the six main decisions, derived from De Bruyn et al.'s (2003) DC described above, in the order in which they would perform them.

Results showed that psychologists would first identify and analyze the client's complaints. Next, they would classify the client's problem into a mental disorder and finally they would select a treatment method. The position of case formulation in the diagnostic process is less clear: some psychologists would perform this as one of the first decisions, others as one of the last. Psychologists' agreement with the prescribed sequence

of decisions as well as agreement among psychologists about the sequence of decisions was low. We found a trend that as experienced increased agreement with the sequence of decisions from a theoretical model as well as agreement among psychologists decreased.

In chapter 3, the sequence of decisions was analysed in a highly structured task with pre-structured response options and psychologists could rank order decisions only once. The limitation of this approach was that it is a restricted representation of actual practice. In clinical practice, psychologists can easily and quickly shift between decisions while talking to a client.

Decisions During a Diagnostic Interview

In chapter 4 the third study is described in which an authentic diagnostic task was used. The analysis is reported of the type and sequence of psychologists' diagnostic decisions and considerations in a diagnostic interview with a simulated client. In addition, we examined to what extent these considerations are connected to psychologists' assessment reports based on their diagnostic interviews. Psychologists first interviewed the same simulated client for at most thirty minutes. After this interview, they wrote a report summarizing their findings in a DSM-IV classification, a case formulation and a treatment proposal. After that, a stimulated recall interview was held during which psychologists watched their own diagnostic interview and reflected on their own performance. We analyzed psychologists' decisions in three different ways, based on the diagnostic interview, stimulated recall interview and written report. The type and sequence of the decisions *complaint analysis*, *classification*, *case formulation* and *treatment selection* were inspected and compared for the diagnostic and stimulated recall interview. Consistency of the content of the decisions was reviewed between the stimulated recall and the written report.

To analyze the type and sequence of decisions made during the diagnostic interview, we developed a coding schema based on the Dutch Guideline for Psychiatric Assessment with Adults by Sno, Hengeveld and Beekman (2004). We examined to what degree psychologists asked questions or made remarks related to *complaint analysis*, *classification*, *case formulation* and *treatment selection* decisions. We found that psychologists asked questions or made remarks about every type of decision throughout the entire diagnostic interview. They mostly asked and remarked about *complaint analysis* and *case formulation*. Questions and remarks about *treatment selection* increased towards the end of the interview.

To analyze the type and sequence of decisions made during the stimulated recall interview, we developed a coding schema based on the Diagnostic Cycle by De Bruyn et al. (2003) because it provides a complete and comprehensive inventory of the decision activities psychologists could consider. Again, we studied the occurrence of the four types of decisions, *complaint analysis*, *classification*, *case formulation* and *treatment selection*. Similar to the sequence of decisions seen during the diagnostic interview, we found that psychologists reflected on every type of decision throughout the entire stimulated recall interview. They mostly reflected on *complaint analysis*. There was an increase in reflections on *treatment selection* towards the end of the interview.

To analyze the consistency of the content of decisions considered during the stimulated recall and those described in the written reports, the content of these decisions was categorized. This was done for the *classification*, *case formulation* and *treatment selection* decisions. *Case formulation* was further divided into potential stressors (and predisposing experiences for the written report) and explanatory mechanisms. *Treatment selection* was further divided into treatment methods and treatment goals. We found reasonable

agreement between the content of the potential stressors considered during the stimulated recall interview and those described in the written reports. Consistency of content between classifications, explanatory mechanisms, treatment methods and treatment goals was moderate to low. Inter-clinician agreement about classification was high, while agreement about causal factors, explanatory mechanisms, treatment methods and treatment goals was moderate to low.

The diagnostic interview with a simulated client was designed to reflect actual practice as much as possible. Beforehand, participants were told they would either see a real client or an actor portraying a client, they were instructed to perform the interview as if they were in their own practices and they were able to take notes. This was done to enhance the authenticity of the diagnostic situation.

This task situation is characterized by profound time pressure and high cognitive effort. Psychologists have limited time to conduct an interview. Furthermore, cognitive resources need to be allocated to different mental tasks during the interview. Psychologists have to think about how to structure the interview effectively and gather relevant information while at the same time maintaining a therapeutic relationship with the client.

Authenticity of the situation is essential in an interview with a simulated client (Melluish, Crossley, & Tweed, 2007). To closely represent clinical practice the simulated client must have "faith validity": the extent to which a psychologist actually believes he is talking to a real client. Faith or face validity is reduced when there is an inconsistency between the performance of the simulated client and what psychologists would expect from a real client. These inconsistencies can result, for example, from an incorrect representation of complaints or symptoms (e.g. the simulated client reports feeling exhausted but comes across rested), a failure to respond with a genuine and appropriate emotion (e.g. not being able to break into tears when a sensitive topic is discussed) or being too verbose or rigid in the responses. Another obstacle to an authentic situation is using a camera, or other equipment, to record the interview. Only some participants reported inconsistencies in the simulated client's performance and some also mentioned their discomfort about the presence of the camera. For them, this might have reduced the authenticity of the diagnostic situation, but for all others the situation closely reflected clinical practice.

In chapter 4, the sequence of decisions was investigated as well as the relationship between diagnostic decisions and treatment plans. The sequence of decisions both during the diagnostic interview and the stimulated recall interview was analysed. The diagnostic interview is a highly unstructured task and also highly representative of actual practice. Some data reduction was necessary to facilitate analysis of the decision sequence. The sequence of decisions was verified at three moments in the diagnostic process, i.e. the beginning, middle and end of the diagnostic and the stimulated recall interview.

To analyse the relationship between diagnostic decisions and treatment plans, the considerations reported by the psychologists during the stimulated recall interview were related to the treatment plans described in the written reports. Psychologists might have had other considerations, corresponding to the ones described in the written reports, but not reported them during the stimulated recall interview. The stimulated recall procedure of watching the video is likely to have prompted a subset of all reflections on the diagnostic interview. The participants could have only reported reflections about salient or atypical incidents that occurred during the diagnostic interview.

Case Complexity and Clinical Decision Making

In chapter 5 the analysis is reported of the influence of client case complexity on the

quality of case formulations. In addition, we investigated the relationship between psychologists' diagnostic decisions (*classification* and *case formulation*), their theoretical orientations and the treatment decisions. We constructed a questionnaire with two case descriptions which varied in the complexity of client problems (where complexity refers to familiarity with client problems and availability of an empirically supported treatment). For each case description, psychologists selected a DSM-IV classification, provided a case formulation and selected a treatment method.

To analyze the influence of client case complexity on the quality of case formulations, the quality of the case formulations was evaluated. We developed a coding schema based on the case formulation quality rating scale designed by Kuyken et al. (2005) and work by Vermande (1995) and Vermande, Van den Bercken and De Bruyn (1996). Case formulations were divided into one or more explanatory hypotheses. Each explanatory hypothesis was evaluated on six dimensions: (1) form (simple/composite hypothesis, causal chain/model), (2) relations between causal factors, (3) relevance, (4) consistency, (5) modifiability of causal factors (not, indirectly and directly modifiable), and (6) positive indicators for treatment (present/absent). Psychologists' explanatory hypotheses for the least complex case were more often relevant, they more often contained relationships between causal factors and directly modifiable factors and less often not modifiable factors.

To analyze the relationship between psychologists' case formulations and their treatment decisions, the content of the case formulations was evaluated. We developed a coding schema based on the Case Formulation Content Coding Method of Eells, Kendjelic and Lucas (1998). More specifically, we analyzed the content of the explanatory mechanisms described in psychologists' case formulations (if present). In addition, we reviewed the relationship between classifications and psychologists' theoretical orientations and their treatment decisions. Psychologists' diagnostic decisions, either classifications or explanatory mechanisms, or their theoretical orientations were not or only very weakly related to the treatment decisions. The client's overt complaints (i.e. anxiety complaints versus dissociative and depressive complaints) were most highly associated with the treatment decisions, although still weakly.

In chapter 5, the relationship between diagnostic decisions and treatment plans was also analysed. Associations between explanatory mechanisms described in the case formulations and treatment decisions were determined. Not all participants described explanatory mechanisms, therefore only a subset of the participants' responses could be used. On a group level, we found weak associations between explanatory mechanisms and treatment decisions. Witteman and Koele (1999) also did not find a strong relationship between psychologists' considerations and their treatment decisions on a group level. However, De Kwaadsteniet, Hagmayer, Krol and Witteman (2010) did find a relationship on an individual level between psychologists' causal models of a client and their treatment decisions. Examining the relationship between causal factors and mechanisms and treatment decisions on an individual level could help explain the lack of associations on a group level.

DISCUSSION OF FINDINGS

Diagnostic Decision Process

Results from the studies reported in chapters 2, 3 and 4 indicate that psychologists' diagnostic processes are adaptive and characterized by a focus on the client's complaints and symptoms and on treatment selection. Furthermore, a continuous shifting between

decisions related to *complaint analysis, classification, case formulation* and *treatment selection* could be seen throughout the entire diagnostic process. This continuous shifting between decisions was also found in studies by Pijnenburg (1996) and Bartolo, Dockrell and Lunt (2001) who analyzed diagnostic decision making in clinical case conferences. The diagnostic decision processes in their studies showed various recursive patterns. These results suggest that psychologists cope with the restricting circumstances of clinical practice in a solution-oriented manner: they aim for an initial treatment plan based on information they consider minimally needed, mainly consisting of client complaints and symptoms.

The artificial task situations (the written case descriptions and the stimulated recall interview) revealed the focus on the client's complaints and symptoms and on treatment selection much more clearly than the authentic task situation (the diagnostic interview with a simulated client). Psychologists asked as many questions about complaints and symptoms as they did about case formulation in the authentic situation. There are two possible explanations for this finding. First, psychologists routinely gather case formulation information without any explicit goal and they do not use this information in their diagnostic processes. Second, psychologists do consider and use case formulation information they have gathered, but fail to report it. Why would psychologists not report considering case formulation information when in fact they do? It could be that psychologists do not elaborate on this kind of information; general, unspecific explanations are sufficient for further treatment planning. This explanation is supported by the result from chapter 4 that about three-fourth of the psychologists did mention an explanatory mechanism to account for the client's problems. Furthermore, when explicitly asked to consider case formulation information, as was done in chapter 5, psychologists were able to use that information to construct relevant and consistent explanations.

There are several possible explanations why psychologists would not explicitly perform case formulation. First, a possible explanation could be that focusing on the client's complaints and symptoms has advantages in situations with considerable time pressure and when decisions need to be made quickly. An analysis of the client's complaints is useful because it helps establish the therapeutic relationship and is a sign of professional respect towards a client (De Bruyn et al., 2003). For example, one participant remarked during the stimulated recall interview: "I always ask the client about the information from the referral letter for two reasons: first, to check the information and second, to have a kind of an equal start with the client." An analysis of the client's symptoms and their severity resulting in classification is useful because scientific knowledge about similar cases can be applied to the current case (e.g. about successful treatments) and classification guides treatment selection in cases where an empirically supported treatment is available (Chambless & Ollendick, 2001; De Bruyn et al., 2003). An analysis of the client's complaints, symptoms and problems can result in enough information to start initial treatment, e.g. when client problems are straightforward and an empirically supported treatment is available. De Kwaadsteniet (2009) argued that psychologists better not perform case formulation because in general reliability and validity of case formulations are poor. Invalid case formulations might do more harm than good.

Second, a possible explanation (see the discussions of chapters 2 and 4) could be that case formulation is unnecessary because psychologists rely on implicit, schema-based causal theories (cf. Brammer, 1997; Kim & Ahn, 2002). Psychologists use pattern recognition to see whether the pattern of complaints and problems of a specific client fits their implicit, causal theory. If they do, explicit case formulation would be redundant.

A third explanation, related to the second one, could be that psychologists are unable to construct causal representations because knowledge about the aetiology of disorders is lacking in most cases (Cicchetti & Sroufe, 2000; Shapiro, 1985; Stricker & Trierweiler, 1995). Case formulation is theory-driven (Kuyken et al., 2005), therefore it requires firm theoretical and evidence based frameworks describing the underlying mechanisms of a disorder. If such a framework is available, psychologists are of course more inclined to use that knowledge. This explanation is supported by the results from chapter 5. Psychologists gave higher quality case formulations for a client presenting with less complex problems, i.e. familiar problems for which an evidence based framework explaining the disorder is available, than for a client presenting with more complex problems. For example, one participant stated in an explanatory hypothesis for the client with anxiety problems: "Classic conditioning (closed in by trucks on the highway), followed by operant conditioning (by avoiding situations)."

Finally, a possible explanation could be that psychologists assume that case formulation is laborious and time-consuming (Perry, Cooper, & Michels, 1987; Virúes-Ortega & Haynes, 2005). Attempting to construct a complete and comprehensive case formulation taking all relevant information into account can become overly complex (Eells, 2007). Also, for less complex client problems for which an empirically supported treatment is available, case formulation would be redundant and needlessly delay treatment (Nelson-Gray, 2003).

Treatment Utility of Diagnostic Decisions

Results from the studies described in chapters 4 and 5 showed that decisions made in the diagnostic process were only slightly related to treatment decisions. To select an appropriate treatment, psychologists should identify and analyze the client's problems (classification) and identify factors and mechanisms causing these problems (case formulation) which can be modified during treatment (cf. De Bruyn et al., 2003; Nezu & Nezu, 1995). However, the results from chapter 4 showed that psychologists' treatment proposals described in their assessment reports were at best moderately related to the decisions they had considered during the diagnostic interview. Also, the results from chapter 5 showed that neither classifications nor explanatory mechanisms described in case formulations were related to the proposed treatment methods.

Other studies have also shown that psychologists' treatment decisions were only weakly related to their diagnostic decisions (e.g. Bus & Kruizenga, 1989; Witteman & Koele, 1999; Witteman & Kunst, 1997). In these studies, psychologists seemed to rely on pattern recognition heuristics shaped by a combination of personal experience, theoretical orientation and clients previously seen. Psychologists appear to develop cognitive schema's in which particular client complaints and symptoms are linked to specific treatment methods that can be used to treat a client with those complaints and symptoms (Mayfield, Kardash, & Kivlighan, 1999; Witteman & Koele, 1999). Psychologists weight the same symptoms differently (Kim & Ahn, 2002; De Kwaadsteniet et al., 2010). Depending on the weights of particular symptoms in psychologists' schemas, different symptoms of the same client could trigger different schemas.

De Kwaadsteniet et al. (2010) did find, on an individual level, a relationship between psychologists' causal models of a client and their treatment decisions. They suggest that the weakness of the relationship between psychologists' diagnostic and treatment decisions could be explained by differences in the causal relations that psychologists' describe between the same factors. Treatments have different effects depending on how these factors

are thought to be related. Diversity in treatment plans might be caused by individual differences in psychologists' perceived effectiveness of treatments and by differences in the explanations offered for client problems. In chapter 4, we observed large variation in explanations offered as well as in treatment methods and goals proposed for the same client. Some of the explanatory mechanisms the psychologists considered can be matched directly to specific treatment goals, for example the explanatory mechanism 'life phase problems' can be matched to the treatment goal 'coping with life phase problems'. However, the relation between specific explanatory mechanisms and treatment methods seems less clear. There were more psychologists who proposed multiple treatment methods than there were psychologists who proposed multiple treatment goals. Multiple treatment methods seem to be proposed to target and change the same causal factors and mechanisms. Nezu and Nezu (1995) even advocate using multiple treatment methods to increase generalization of the effects of different treatments and maintenance of the outcomes achieved.

Another possible explanation for the weak associations could be that treatment plans were influenced by the unstructured information gathering processes of the psychologists. Psychologists continuously shifted between decisions and reconsidered previous decisions. Depending on the moment that specific information is gathered, psychologists might reach different conclusions. Availability of the information in memory could then influence the final decisions that psychologists make and describe in their assessment reports.

Finally, an explanation for the lack of a clear relationship could be that psychologists base their treatment decisions on other factors, unrelated to their diagnostic decision processes, such as institutional policy, personal preferences, or evidence based guidelines (e.g. see Nelson & Steele, 2008; O'Donohue, Fisher, Plaud, & Curtis, 1990). For example, in the study reported in chapter 5, one participant remarked about the treatment for one of the clients: "Irrespective of the kind of classification, cognitive-behavioural treatments will be similar." In chapter 5, psychologists' treatment decisions for the case where an empirically supported treatment was available were consistent with this recommended treatment, even though psychologists differed in the classifications and case formulations proposed for this client. This could suggest that evidence based guidelines determined treatment selection more than classification and case formulation did even though in general there is a strong normative relationship between classification and treatment selection.

Inter-clinician Agreement

In general, reliability of diagnostic decisions tends to be low (Garb, 1998). The results from the studies described in this thesis showed that under certain circumstances agreement among psychologists can be high. In chapter 4 psychologists agreed highly with each other about the classification of depressive disorder. For classification, psychologists can rely on diagnostic criteria, which reduces the complexity of the task. For example, one participant in chapter 4 described classification as "going through the list of criteria". Also, in this case the client's problems could be classified into a single, common disorder, i.e. depressive disorder. It seems that inter-clinician agreement can be high for commonly encountered, overt problems.

A possible explanation for the higher agreement on classification than on causal information and treatment methods is that a standardized manual, the DSM, is available to help psychologists classify clients' problems. The psychologists in the third study could consult the DSM manual during the diagnostic interview and they were all familiar and

well trained with this manual. They presumably relied on prior and possibly commonly shared, knowledge from working with the manual in the past to assess and classify the current client's symptoms and problems. There are no such manuals to assess causal factors and mechanisms or to select treatments and neither are there firm evidence based frameworks describing a disorder's aetiology. Thus, psychologists could not rely on commonly shared knowledge concerning case formulation and treatment selection.

A possible explanation for the low agreement about decisions related to case formulation and treatment selection and the low agreement about the sequence of decisions, could be that psychologists used different heuristics to arrive at their decisions (Garb, 1998). Heuristics are based on previous experiences in certain environments. Psychologists' experiences and environments are likely to be different and highly individual, e.g. depending on training, setting, theoretical orientation and clients previously seen, leading to the development of different and highly individualized heuristics (Gigerenzer, 2000; Kahneman & Klein, 2009).

Experience

Experience is one of the most commonly investigated as well as hotly debated topics in clinical decision making. In general, research showed that experience is not associated with better clinical decisions (Dawes, 1996; Faust, 1986; Garb, 1998; Lichtenberg, 1997). However, a recent meta-analysis on the effect of experience on clinical decision making showed a small but reliable and positive effect of experience on diagnostic accuracy (Spengler et al., 2009). This effect was found across task situations, such as problem type or treatment decision, type of experience (clinical or educational) and breadth of experience (general or specific). Spengler et al. (2009) state that it is still unclear which factors or processes, such as decision strategies, contribute to the positive effect of experience. The same can be argued for the effect of experience on psychologists' decision strategies. In chapters 2 and 3 of this thesis, we examined the effect of experience on psychologists' decision strategies.

In chapter 2, the regression analysis showed a significant influence of background characteristics, including experience, on the diagnostic decisions that psychologists judge important and would perform. Psychologists' background characteristics seem to determine the diagnostic decision process to some extent. Unfortunately, the individual contribution of experience to the decision process was not determined because of the heterogeneity of the predictors used and the presence of suppressor variables.

In chapter 3, the effect of experience on the sequence of decisions was analyzed. Clinical experience helps to structure the diagnostic task (Brammer, 2002) and adapting the decision process to the task and situation at hand is considered an important characteristic of expertise (Ericsson & Lehmann, 1996). A trend was found showing that agreement with a prescriptive model's sequence of decisions and inter-clinician agreement both decrease as clinical experience increases, suggesting that the diagnostic decision process becomes more divergent. With increasing experience, psychologists seem to adapt their diagnostic decision processes to the task demands and to the specific situations they encounter in their clinical practices.

To conclude, these results seem to indicate that experience influences psychologists' diagnostic decision processes. Considering the small effect of experience on diagnostic accuracy described in the meta-analysis by Spengler et al. (2009), the effect of experience has not been significant in our studies because of methodological limitations, as discussed in chapters 2 and 3. Careful selection of novice, experienced and expert psychologists (cf.

Witteman & Van den Bercken, 2007) and investigating the relationship between experience, decision strategies and diagnostic accuracy could provide further insight into the role of experience in clinical decision making.

APPLICABILITY OF THEORETICAL DIAGNOSTIC DECISION MODELS

Although we did not explicitly intend to evaluate the applicability of diagnostic decision models, the studies in this thesis provide insight into their use in clinical practice. From the four studies described in this dissertation it can be concluded that psychologists do not always follow all of the decisions prescribed in diagnostic decision models. Diagnostic decision models are assumed to improve diagnostic decision performance because they rely on decision strategies to counter the unwanted effects of heuristics and biases and on increasing the likelihood of receiving corrective feedback by applying the scientific method (De Bruyn et al., 2003; Nezu & Nezu, 1995). These models were developed to aid psychologists in structuring the complex and difficult diagnostic task and thus improve the quality of the diagnostic process. Apparently, these models are insufficiently supporting psychologists' diagnostic processes.

There are several possible explanations for psychologists' lack of adherence to prescriptions of diagnostic decision models. First, most prescriptive models, including the diagnostic decision models, are rather time-consuming. They propose strict and lengthy procedures which require a lot of mental effort (Van Aarle & Van den Bercken, 1999; Vertommen, Ter Laak, & Bijttebier, 2005). For example, the guidelines for the assessment process (GAP) proposed by Fernández-Ballesteros et al. (2001) describe 40 decision steps psychologists should consider in psychological assessment. Time and resources are limited in clinical practice, constraining the amount and kind of information that can be gathered (Gambrill, 2005). Furthermore, psychologists' cognitive capacities limit the amount of information that can be processed (Newell & Simon, 1972). Therefore, psychologists do not perform all prescribed decisions from diagnostic decision models or do not perform these decisions as thoroughly as recommended.

Second, because of a lack of immediate and appropriate feedback from clients, psychologists do not receive accurate information about the effectiveness of their decision strategies (Dawes 1996; Garb, 1998). Psychologists can even be reinforced to use their personal, possibly ineffective, strategies because of judgement biases, such as the confirmation and hindsight bias (cf. Gambrill, 2005; Lichtenberg, 1997). Also, research suggests that psychologists have little awareness of their own decision making processes (Arkes, 1981; Aspel, Willis, & Faust, 1998; Dhami & Harries, 2001). Psychologists' own reported use of information differs from objective measures of their use of information. Thus, psychologists are not aware of the need to adjust their decision strategies and to follow prescriptions from diagnostic decision models.

Third and perhaps most important, so far, there are no empirical studies investigating whether following all the prescriptive decisions of diagnostic decision models improves the outcome of the diagnostic process. Research has shown that models of psychologists' diagnostic decision processes based on all information available, such as Bayesian models or regression analyses, fit psychologists' actual decision processes just as well as models based on limited information, such as fast and frugal heuristics (cf. Katsikopoulos, Pachur, Machery, & Wallin, 2008; Witteman et al., 2007). Therefore, the question remains whether scientifically based decision strategies actually lead to better decisions. Studies by Sartorius et al. (1993) and Kendjelic and Eells (2007) showed that following structured methods to

gather and process clinical information can improve parts of the diagnostic process, i.e. classification and case formulation respectively. Along the same lines, a structured and thorough diagnostic process could result in improved treatment decisions. Although this has not been empirically verified. Therefore, further research is needed to establish whether following the prescriptions of diagnostic decision models in clinical practice actually improves diagnostic and treatment decisions and therapy outcomes as well.

To conclude, despite efforts to disseminate and implement diagnostic decision models in clinical training and practice, they are not easily applied. Psychologists might deliberately use heuristics to reduce the number of decisions in the diagnostic process. The results reported in chapters 2, 3 and 4 and previous studies reported in the literature strongly suggest that psychologists use individualized heuristic decision strategies. These strategies seem to lead to more efficient diagnostic processes and might also lead to effective treatment plans. Establishing the efficiency and success of heuristics would shed a new light on the concept of professional experience. Professional experience is often defined in terms of a vast amount of declarative knowledge and applying the scientific method. Experienced psychologists' use of intuitive knowledge, captured in heuristic decision strategies, might resolve the complex relationship between professional experience and effective decision strategies.

Heuristic decision strategies have the advantage that they are easily applied, reduce the amount of information needed for a decision and often lead to similar outcomes as formal strategies based on all information available (Gigerenzer, Todd, & The ABC Research group, 1999). The psychological plausibility of heuristic decision strategies is indeed higher than that of expert systems that combine and integrate all information available (Dhami & Harries, 2001) and therefore, psychologists are more likely to accept and use them. Heuristic decision strategies are especially useful in situations with limited time to gather and process information. When psychologists deliberately use experience-based short-cuts in their diagnostic processes they could benefit from decision aids that make use of heuristic decision principles when available, to increase efficient use of information. Decision support based on psychologically plausible, heuristics-based knowledge instead of formal, explicit and scientifically based knowledge may convincingly be introduced into clinical practice. Further research into the characteristics of effective heuristic decision strategies would help the development of such decision aids.

IMPLICATIONS

Reflecting on the studies described in this thesis, it seems that psychologists' diagnostic processes are adaptive with a focus on identifying and analyzing the client's complaints and symptoms and on treatment selection while they continuously shift between decisions and reconsider previous decisions. Furthermore, classification and case formulation decisions appear to be only weakly related to treatment decisions.

Clinical Practice and Training
Deviating from the prescribed sequence of decisions could result in a loss of information because necessary input for the next step is missing and decisions are based on incomplete information (cf. Nezu & Nezu, 1995). Additionally, psychologists do not gather sufficient information for some decisions (such as case formulation) because they judge diagnostic activities related to those decisions less important. The effect of missing information on the outcome of the decision process depends on (i) the kind of decision that

is not, or not thoroughly, performed and (ii) the moment in the diagnostic process at which the information gathered is processed and integrated into a client model. Missing information about registration, for example, could be considered less harmful to further treatment planning than missing information about symptoms or causal factors. Also, gathering information in an unstructured manner does not imply that information is also processed and integrated in an unstructured manner. Psychologists gather as much information as possible before integrating this information into a client model and evaluating whether more information is needed to complete the model. This model consists of a description of the client's problems and includes an explanation representing how possible causal factors and explanatory mechanisms can account for the client's problems. In chapter 4, about three-fourth of the psychologists described an explanatory mechanism, suggesting that they were able to integrate the information gathered coherently.

Contrary to what is prescribed by diagnostic decision models, treatment decisions are made before the diagnostic process is completed and they are only very weakly related to decisions made in this process. Psychologists were found to consider treatment options right from the start of a diagnostic interview. Doing so, psychologists could risk 'premature closure' because they do not thoroughly explore alternative options (Arkes, 1981). However, psychologists' treatment decisions appear to change over time depending on the purpose of the task; the consistency of treatment decisions from interview to report was low. In the dynamic diagnostic situation, psychologists shift between decisions and reconsider previous decisions (see also Bartolo et al., 2001). This is not necessarily a harmful practice; it can help psychologists quickly adjust courses of action when the situation changes.

Our studies provided a method for analyzing decision making processes in a clinical setting, using the Diagnostic Cycle as a tool. We were able to capture the overall structure of the diagnostic process, rather than isolating separate parts of the process (e.g. Eells et al., 1998; Krol et al., 1992). This method can be used to increase psychologists' awareness of their deviations from a diagnostic decision model and monitor their own diagnostic processes. Canon-Bowers and Bell (1997) suggest that increasing monitoring and meta-cognitive skills is an effective training method to improve decision making processes. Barrows (2000) describes how the stimulated recall procedure can be used to identify "possible problems in the subject's reasoning process as well as deficiencies in clinical performance" (p. v). The coding schema, concentrating on the *complaint analysis*, *classification*, *case formulation* and *treatment selection* decisions, can be very useful in combination with stimulated recall for in-service training of the diagnostic task.

Decision Aids

Decision aids should be developed to support and improve psychologists' diagnostic processes in this complex and dynamic situation. Decision aids that reduce the time and cognitive effort needed to follow diagnostic decision models seem to improve adherence to the prescriptions of these models (e.g. see Witteman & Kunst, 1999). These aids should focus on case formulation because, compared with other diagnostic decisions, psychologists (i) judged diagnostic activities related to this decision less important, (ii) considered this decision less often during a diagnostic interview, (iii) agreed less with each other about its contents and (iv) constructed case formulations of lower quality for cases in which it is most needed, i.e. when client problems are complex and an empirically supported treatment is unavailable. Case formulation is supposed to be an essential part of the diagnostic process (Eells, 2007; Haynes, 1993; Tarrier & Calam, 2002). It can help

psychologists decide on a treatment when there are different treatments available for the same disorder and classification alone is insufficient to make a choice, the first choice treatment fails or client problems are complex (Haynes & Williams, 2003; Persons, 2006; Persons & Mikami, 2002).

Representation of information in structural causal models is expected to be more helpful as a decision support tool than an expert knowledge-based system. The body of knowledge about the aetiology of disorders in clinical psychology is continuously changing and such systems would need to be continuously updated (Pijnenburg, 1996). For a structural causal model to be a useful tool it should aim at visualizing the relevant causal factors and their relations (Haynes, Spain, & Oliveira, 1993; Morton, 2004). Such a model would provide psychologists with insight into the essential parts of a case formulation that should be considered. Visualization would show which additional factors and relations need to be examined and described to arrive at a meaningful explanation.

Several decision support tools have been developed for case formulation. However, they are insufficient because they are either time-consuming and require complex computations (see e.g. Virúes-Ortega & Haynes, 2005) or they are developed for specific theoretical orientations (see Eells, 2007, for an overview of case formulation models). In clinical practice, most psychologists work from more than one theoretical orientation or within multidisciplinary teams. A structural causal model that allows visualization and integration of biological, cognitive, behavioural and environmental factors from different theoretical orientations is John Morton's 'causal modelling' method (2004; see also Krol, Morton, & De Bruyn, 2004; Morton & Frith, 1993). A case study by Krol & Kuijpers (2007) showed that Morton's causal modeling method can be applied in-session allowing psychologists to develop a case formulation together with the client. This would limit the amount of time needed to construct a case formulation. Also, co-constructing an explanation for the client's problems could foster a shared understanding (cf. Gijlers, 2005) and promote client agreement with the treatment plan. We suggest that further development of decision support tools should focus on the application of the causal modelling method for in-session use with a client. Additionally, the quality of case formulations and treatment decisions made with the causal modelling tool should be compared with those made with traditional models and methods.

Further Research

The studies described in this thesis did not completely reveal (i) how psychologists use the information about client complaints and symptoms in the diagnostic process and for their treatment decisions, (ii) whether experience influences the type, sequence, content and quality of decisions made in the diagnostic process and (iii) to what extent a better diagnostic decision process leads to better treatment decisions and improved treatment outcomes.

First, the results showing a focus on client complaints and symptoms and on treatment selection can be followed by investigating how psychologists use this kind of information in the diagnostic process. Psychologists' treatment decisions seem to be schema driven, based on links between client complaints and symptoms and possible treatments. Few studies have explored how clinicians' knowledge about client complaints and symptoms is represented in these schemas and how this is linked to other concepts (see e.g. Mayfield et al., 1999). The knowledge structure of these schema's as well as the inferences made from them should be investigated further.

Second, a recent meta-analysis examining the influence of experience on diagnostic accuracy (Spengler et al., 2009) suggest that experience also influences psychologists' diagnostic decision processes. Previous studies showed that as experience increases, psychologists approach the diagnostic process in a more flexible way, based on the clinical knowledge they have acquired in practice (Brammer, 1997; Hillerbrand & Claiborn, 1990). Our results showed a trend that as experience increases, agreement about the sequence of decisions in the diagnostic process decreases, resulting in a more divergent diagnostic process. Careful design of studies investigating the influence of experience on the diagnostic decision process could help further explain the diversity of psychologists' diagnostic processes and to discern different decision strategies used.

Finally, research investigating the treatment utility of a thorough diagnostic process is scarce (Nelson-Gray, 2003). Though the focus of the current thesis is strictly process-oriented, the need for research establishing that a systematic and thorough diagnostic process leads to better decisions is acknowledged. In other domains, it has been established that decisions based on heuristics using limited information are not necessarily less accurate and less successful than decisions based on all available information (Gigerenzer, 2009). Determining the situations in which heuristics are successful and in which a more thorough diagnostic process is called for, should be the focus of further research. However, in the domain of clinical psychology, research on the accuracy of decision strategies and their outcomes is obstructed because no generally acknowledged outcome measures exist. Psychologists disagree about the quality criteria of diagnostic decisions. To advance research on the effectiveness of different decision strategies, the daunting task of reaching agreement about the criteria for decision accuracy needs to be accomplished first.

REFERENCES

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders.* (4th revised ed.). Washington, DC: Author

Ajzen, I. (2001). Nature and operation of attitudes. *Annual Review of Psychology, 52,* 27-58. doi: 10.1146/annurev.psych.52.1.27

Ajzen, I. & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin, 84,* 888-918. doi: 10.1037/0033-2909.84.5.888

Arkes, H. R. (1981). Impediments to accurate clinical judgment and possible ways to minimize their impact. *Journal of Consulting and Clinical Psychology, 49,* 323 - 330. doi: 10.1037/0022-006X.49.3.323

Aspel, A. D., Willis, W. G., & Faust, D. (1998). School psychologists' diagnostic decision-making processes: Objective-subjective discrepancies. *Journal of School Psychology, 36,* 137 - 149. doi:10.1016/S0022-4405(98)00002-8

Barrows, H. S. (2000). *Stimulated recall. Personalized assessment of clinical reasoning.* Springfield: Southern Illinois University School of Medicine.

Bartolo, P. A., Dockrell, J., & Lunt, I. (2001). Naturalistic decision-making task processes in multiprofessional assessment of disability. *Journal of School Psychology, 39,* 499 - 519. doi: 10.1016/S0022-4405(01)00086-3

Beutler, L. E. (1991). Have all won and must all have prizes? Revisiting Luborsky et al.'s verdict. *Journal of Consulting and Clinical Psychology, 59,* 226 - 232. doi: 10.1037/0022-006X.59.2.226

Beutler, L. E., Williams, R. E., & Wakefield, P. J. (1993). Obstacles to disseminating applied psychological science. *Applied & Preventive Psychology, 2,* 53 - 58. DOI: 10.1016/S0962-1849(05)80111-8

Beutler, L. E., Williams, R. E., Wakefield, P. J., Entwistle, S. R. (1995). Bridging scientist and practitioner perspectives in clinical psychology. *American Psychologist, 50,* 984 - 994. doi: 10.1037/0003-066X.50.12.984

Bieling, P. J. & Kuyken, W. (2003). Is cognitive case formulation science or fiction? *Clinical Psychology: Science and Practice, 10,* 52-69. doi: 10.1093/clipsy.10.1.52

Blashfield, R. K., & Draguns, J. G. (1976). Evaluative criteria for psychiatric classification. *Journal of Abnormal Psychology, 85,* 140 - 150. doi: 10.1037/0021-843X.85.2.140

Borsboom, D. (2008). Psychometric perspectives on diagnostic systems. *Journal of Clinical Psychology, 64,* 1089 - 1108. doi: 10.1002/jclp.20503

Boshuizen, H. P. & Schmidt, H. G. (1992). On the role of biomedical knowledge in clinical reasoning by experts, intermediates, and novices. *Cognitive Science, 16,* 153-184. doi: 10.1207/s15516709cog1602_1

Brammer, R. (1997). Case conceptualization strategies: The relationship between psychologists' experience levels, academic training, and mode of clinical inquiry. *Educational Psychology Review, 9*, 333-351. doi: 10.1023/A:1024794522386

Brammer, R. (2002). Effects of experience and training on diagnostic accuracy. *Psychological Assessment, 14,* 110 - 113. doi: 10.1037/1040-3590.14.1.110

Bus, A. G., & Kruizenga, T. H. (1989). Diagnostic problem-solving behavior of expert practitioners in the field of learning disabilities. *Journal of School Psychology, 27*, 277-287. doi: 10.1016/0022-4405(89)90042-3

Canon-Bowers, J. A. & Bell, H. H. (1997). Training decision-makers for complex environments: Implications of the naturalistic decision making perspective. In C. E. Zsambok & G. Klein (Eds.), *Naturalistic decision making* (pp. 99 - 110). Hillsdale, NJ: Erlbaum.

Caspar, F. (1997). What goes on in a psychotherapist's mind? *Psychotherapy Research, 7,* 105 - 125. doi: 10.1080/10503309712331913

Chambless, D. L. & Ollendick, T. H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology, 52,* 685 - 716. doi: 10.1146/annurev.psych.52.1.685

Cicchetti, D. & Sroufe, A. L. (2000). Editorial: The past as prologue to the future: The times, the've been a-changin'. *Development and Psychopathology, 12,* 255 - 264. doi: 10.1017/S0954579400003011

Coderre, S., Mandin, H., Harasym, P. H., & Fick, G. H. (2003). Diagnostic reasoning strategies and diagnostic success. *Medical Education, 37*, 695-703. doi: 10.1046/j.1365-2923.2003.01577.x

Committee on Training in Clinical Psychology (1947). Recommended graduate training program in clinical psychology. *American Psychologist, 2,* 539 - 558. doi: 10.1037/h0058236

Cooper, R. (2004). What is wrong with the DSM? *History of Psychiatry, 15,* 5 - 25. doi: 10.1177/0957154X04039343

Curtis, J., Silberschatz, G., Sampson, H., & Weiss, J. (1994). The plan formulation method. *Psychotherapy Research, 4,* 197-207. doi: 10.1080/10503309412331334032

Dhami, M. K. & Harries, C. (2001). Fast and frugal versus regression models of human judgement. *Thinking & Reasoning, 7,* 5 - 27. doi: 10.1080/13546780042000019

Dawes, R. M. (1996). *House of cards: Psychology and psychotherapy built on myth.* New York: Free Press.

De Bruyn, E. E. J., Ruijssenaars, A., Pameijer, N. K., & Van Aarle, E. J. M. (2003). *De diagnostische cyclus. Een praktijkleer. [The diagnostic cycle. Learning from practice].* Leuven: Acco.

De Kwaadsteniet, L. (2009). *Clinicians as mechanics? Causal reasoning in clinical judgment and decision making.* Unpublished doctoral dissertation, Radboud University, Nijmegen.

De Kwaadsteniet, L., Hagmayer, Y., Krol, N. P. C. M., & Witteman, C. L. M. (2010). *Causal client models in selecting effective interventions: A cognitive mapping study.* Manuscript submitted for publication.

De Kwaadsteniet, L., Krol, N. P. C. M., & Witteman, C. L. M. (2008). Causality in psychodiagnosticians' mental representations of client information: A think aloud study. *Manuscript submitted for publication.*

Dijkstra, W. (2002). Sequence Viewer 3.0, Department of Social Research Methods, Vrije Universiteit Amsterdam.

Dijkstra, W. & Taris, T. (1995). Measuring the agreement between sequences. *Sociological Methods & Research, 24,* 214 - 231. doi: 10.1177/0049124195024002004

Dirmaier, J., Harfst, T., Koch, U., & Schulz, H. (2006). Therapy goals in inpatient psychotherapy: Differences between diagnostic groups and psychotherapeutic orientations. *Clinical Psychology and Psychotherapy, 13,* 34 - 46. doi: 10.1002/cpp.470

Dougherty, M. R. P., & Hunter, J. E. (2003). Hypothesis generation, probability judgment, and individual differences in working memory capacity. *Acta Psychologica, 113,* 263 - 282. doi: 10.1016/S0001-6918(03)00033-7

Dumont, F. (1993). Inferential heuristics in clinical problem formulation: Selective review of their strengths and weaknesses. *Professional Psychology: Research and Practice, 24,* 196 - 205. doi: 10.1037/0735-7028.25.3.195

Eells, T. D. (2007). *Handbook of psychotherapy case formulation* (2nd edition). New York: The Guilford Press.

Eells, T. D., Kendjelic, E. M., & Lucas, C. P. (1998). What's in a case formulation? Development and use of a content coding manual. *Journal of Psychotherapy Practice and Research, 7,* 144 - 153. Retrieved from http://jppr.psychiatryonline.org/

Eells, T. D., & Lombart, K. G. (2003). Case formulation and treatment concepts among novice, experienced, and expert cognitive-behavioral and psychodynamic therapists. *Psychotherapy Research, 13*, 187-204. doi: 10.1093/ptr/kpg018

Eells, T. D., Lombart, K. G., Kendjelic, E. M., Turner, L. C., & Lucas, C. P. (2005). The quality of psychotherapy case formulations: A comparison of expert, experienced, and novice cognitive-behavioral and psychodynamic therapists. *Journal of Consulting and Clinical Psychology, 73,* 579 - 589. doi: 10.1037/0022-006X.73.4.579

Elstein, A. S., & Schwarz, A. (2002). Clinical problem solving and diagnostic decision making: Selective review of the cognitive literature. *British Medical Journal, 324*, 729-732. doi: 10.1136/bmj.324.7339.729

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology, 47,* 273 - 305. doi: 10.1146/annurev.psych.47.1.273

Faust, D. (1986). Research on human judgment and its application to clinical practice. *Professional Psychology: Research and practice, 17,* 420 - 430. doi: 10.1037/0735-7028.17.5.420

Fernández-Ballesteros, R. (1999). Psychological assessment: Future challenges and progresses. *European Psychologist, 4,* 248 - 262. doi: 10.1027//1016-9040.4.4.248

Fernández-Ballesteros, R., De Bruyn, E. E. J., Godoy, A., Hornke, L. F., Ter Laak, J., Vizcarro, C., Westhoff, K., Westmeyer, H., Zaccagnini, J. L. (2001). Guidelines for the assessment process (GAP): A proposal for discussion. *European Journal of Psychological Assessment, 17,* 187 - 200. doi: 10.1027//1015-5759.17.3.187

Field, A. (2000). *Discovering statistics using SPSS for windows*. Thousand Oaks, CA: Sage Publications Ltd.

Follette, W. C., & Houts, A. C. (1996). Models of scientific progress and the role of theory in taxonomy development: A case study of the DSM. *Journal of Consulting and Clinical Psychology, 64,* 1120 - 1132. doi: 10.1037/0022-006X.64.6.1120

Gambrill, E. (2005). *Critical thinking in clinical practice* (2nd ed.). New York: Wiley.

Garb, H. N. (1989). Clinical judgment, clinical training, and professional experience. *Psychological Bulletin, 105,* 387 - 396. doi: 10.1037/0033-2909.105.3.387

Garb, H. N. (1996). The representativeness and past-behavior heuristics in clinical judgment. *Professional Psychology: Research and Practice, 27,* 272 - 277. doi: 10.1037/0735-7028.26.5.526.

Garb, H. N. (1998). *Studying the clinician: Judgment research and psychological assessment.* Washington, DC: American Psychological Association.

Garb, H. N. (2004). Clinical judgment and decision making. *Annual Review of Clinical Psychology, 1,* 67 - 89. doi: 10.1146/annurev.clinpsy.1.102803.143810

Garb, H. N., & Boyle, P. A. (2003). Understanding why some clinicians use pseudoscientific methods: Findings from research on clinical judgment. In S. O. Lillienfeld, S. J. Lynn, & J. M. Lohr (Eds.), *Science and pseudoscience in clinical psychology* (pp. 17 - 38). New York: Guilford Press.

Gigerenzer, G. (2000). *Adaptive thinking: Rationality in the real world.* New York: Oxford University Press.

Gigerenzer, G., & Brighton, H. (2009). Homo heuristicus: Why biased minds make better decisions. *Topics in Cognitive Science, 1,* 107 - 143. doi: 10.1111/j.1756-8765.2008.01006.x

Gigerenzer, G., Todd, P. M., & the ABC Research Group (1999). *Simple heuristics that make us smart.* New York: Oxford University Press.

Gijlers, A. H. (2005). *Confrontation and co-construction: Exploring and supporting collaborative scientific discovery learning with computer simulations.* Unpublished doctoral dissertation, University of Twente, Enschede.

Glasman, L. R. & Albarracín, D. (2006). Forming attitudes that predict future behavior: A meta-analysis of the attitude-behavior relation. *Psychological Bulletin, 132,* 778-822. doi: 10.1037/0033-2909.132.5.778

Glöckner, A., & Witteman, C. L. M. (2010). Beyond dual-process models: A categorisation of processes underlying intuitive judgement and decision making. *Thinking & Reasoning, 16,* 1 - 25. doi: 10.1080/13546780903395748

Goodman, L. A. & Kruskall, W. H. (1954). Measures of association for cross-classifications. *Journal of the American Statistical Association, 49,* 732-764. Retrieved from http://www.amstat.org/

Gough, H. (1971). Some reflections on the meaning of psychodiagnosis. *Journal of Personality Assessment, 59,* 410 - 423. doi: 10.1037/h0030855

Green, L., & Mehr, D. R. (1997). What alters physicians' decisions to admit to the coronary care unit? *Journal of Family Practice, 45,* 219 - 226. Retrieved from: www.jfponline.com

Groenier, M., Pieters, J. M., Hulshof, C. D., Wilhelm, P., & Witteman, C. L. M. (2008). Psychologists' judgments of diagnostic activities: Deviations from a theoretical model. *Clinical Psychology and Psychotherapy, 15,* 256 - 265. doi: 10.1002/cpp.587

Groth-Marnat, G. (2003). *Handbook of Psychological Assessment.* (4th ed.). New York: John Wiley & Sons, Inc.

Haverkamp, B. E. (1993). Confirmatory bias in hypothesis testing for client-identified and counselor self-generated hypotheses. *Journal of Counseling Psychology, 40,* 303 - 315. doi: 10.1037/0022-0167.40.3.303

Haynes, S. N. (1993). Treatment implications of psychological assessment. *Psychological Assessment, 5,* 251 - 253. doi: 10.1037/1040-3590.5.3.281

Haynes, S. N., & O'Brien, W. H. (1990). Functional analysis in behavior therapy. *Clinical Psychology Review, 10,* 649 - 668. doi: 10.1016/0272-7358(90)90074-K

Haynes, S. N., Spain, E. H., & Oliveira, J. (1993). Identifying causal relationships in clinical assessment. *Psychological Assessment, 5,* 281 - 291. doi: 10.1037/1040-3590.5.2.281

Haynes, S. N., & Williams, A. E. (2003). Case formulation and design of behavioral treatment programs: Matching treatment mechanisms to causal variables for behavior problems. *European Journal of Psychological Assessment, 19,* 164 - 174. doi: 10.1027//1015-5759.19.3.164

Hillerbrand, E. T., & Claiborn, C. D. (1990). Examining reasoning skill differences between expert and novice counselors. *Journal of Counseling & Development, 68,* 684 - 691. Retrieved from: www.counseling.org

Kagan, N., Krathwohl, D. R., & Miller, R. (1963). Stimulated recall in therapy using video tape: A case study. *Journal of Counseling Psychology, 10,* 237 - 243. doi: 10.1037/h0045497

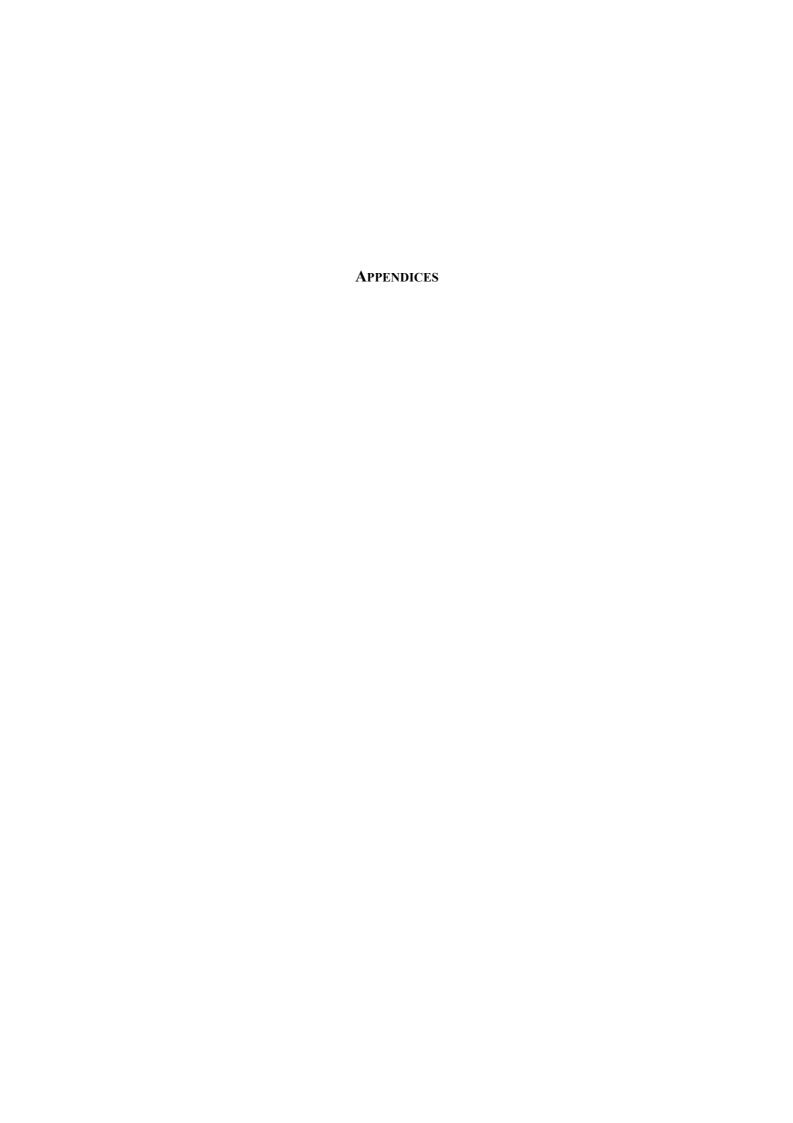Kahneman, D. & Klein, G. A. (2009). Conditions for intuitive expertise: A failure to

disagree. *American Psychologist, 64,* 515 - 526. doi: 10.1037/a0016755

Katsikopoulos, K. V., Pachur, T., Machery, E., Wallin, A. (2008). From Meehl to fast and frugal heuristics (and back). *Theory & Psychology, 18,* 443 - 464. doi: 10.1177/0959354308091824

Kazdin, A. E. (1983). Psychiatric diagnosis, dimensions of dysfunction, and child behavior therapy. *Behavior Therapy, 14,* 73 - 99. doi: 10.1016/S0005-7894(83)80088-4

Kendjelic, E. M., & Eells, T. D. (2007). Generic psychotherapy case formulation training improves formulation quality. *Psychotherapy: Theory, Research, Practice, Training, 44,* 66-77. doi: 10.1037/0033-3204.44.1.66

Kessing, L. V. (2003). Subtypes of depressive episodes according to ICD-10: Prediction of risk of relapse and suicide. *Psychopathology, 36,* 285 - 291. doi: 10.1159/000075186

Kim, N. S., & Ahn, W-k. (2002). Clinical psychologists' theory-based representations of mental disorders predict their diagnostic reasoning and memory. *Journal of experimental psychology: General, 131,* 451 - 476. doi: 10.1037//0096-3445.131.4.451

Klayman, J., & Ha, Y-w. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review, 94,* 211 - 228. doi: 10.1037/0033-295X.94.2.211

Klein, G. A., Orasanu, J., Calderwood, R., & Zsambok, C. E. (1993). *Decision making in action: Models and methods.* Westport, CT: Ablex Publishing.

Krol, N. P., De Bruyn, E. E. J., & Van den Bercken, J. H. (1992). Diagnostic classification by experts and novices. *Acta Psychologica, 81,* 23 - 37. doi: 10.1016/0001-6918(92)90009-3

Krol, N. P. and Kuijpers, R. (2007, April). *Mind the GAP: The application of assessment guidelines in clinical practice.* Presentation at the 2nd European Workshop on Causal Reasoning in Clinical Decision Making, Leuven.

Krol, N., Morton, J., & De Bruyn, E. (2004). Theories of conduct disorder: A causal modelling analysis. *Journal of Child Psychology and Psychiatry, 45,* 727 - 742. doi: 10.1111/j.1469-7610.2004.00267.x

Kuyken, W., Fothergill, C. D., Musa, M. & Chadwick, P. (2005). The reliability and quality of cognitive case formulation. *Behavior Research and Therapy, 43,* 1187 - 1201. doi: 10.1016/j.brat.2004.08.007

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33,* 159 - 174. Retrieved from: http://www.jstor.org/

Lesgold, A., Glaser, R., Rubinson, H., Klopfer, D., Feltovich, P., & Wang, Y. (1988). Expertise in a complex skill: Diagnosing x-ray pictures. In M. T. Chi, R. Glaser & K. L. Farr (Eds.), *The nature of expertise* (pp. 311-342). Hillsdale, NJ: Lawrence Erlbaum Associates.

Lichtenberg, J. W. (1997). Expertise in counseling psychology: A concept in search of support. *Educational Psychology Review, 9,* 221 - 238. doi: 10.1023/A:1024783107643

Lilienfeld, S. O., Lynn, S. J., & Lohr, J. M. (2003). Science and pseudoscience in clinical psychology: Initial thoughts, reflections, and considerations. In: S. O. Lilienfeld, S. J. Lynn, & J. M. Lohr (Eds.). *Science and pseudoscience in clinical psychology* (pp. 1 - 14). New York: Guilford Press.

Mayfield, W. A., Kardash, C. M., & Kivlighan, D. M., Jr. (1999). Differences in experienced and novice counselors' knowledge structures about clients: Implications for case conceptualization. *Journal of Counseling Psychology, 46,* 504-514. doi:

10.1037/0022-0167.46.4.504

Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence.* Minneapolis, MN: University of Minnesota Press.

Meehl, P. E. (1960). The cognitive activity of the clinician. *American Psychologist, 15,* 19 - 27. doi: 10.1037/h0041744

Melluish, S., Crossley, J. & Tweed, A. (2007). An evaluation of the use of simulated patient role-plays in the teaching and assessment of clinical consultation skills in clinical psychologists' training. *Psychology Learning and Teaching, 6,* 104-113. Retrieved from: http://www.psychology.heacademy.ac.uk/

Morton, J. (2004). *Understanding developmental disorders: A causal modelling approach.* Malden, MA: Blackwell Publishing.

Morton, J., & Frith, U. (1993). Causal modelling: A structural approach to developmental psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology. Volume 1. Theory and methods* (pp. 357 - 390). New York: Wiley.

Musewicz, J., Marczyk, G., Knauss, L., & York, D. (2009). Current assessment practice, personality measurement, and Rorschach usage by psychologists'. *Journal of Personality Assessment, 91*, 453 - 461. doi: 10.1080/00223890903087976

Münsterberg, H. (1899). Psychology and history. *Psychological Review, 6,* 1 - 31. doi: 10.1037/h0071306

Münsterberg, H. (1913). *Psychology and industrial efficiency.* Boston: Houghton Mifflin.

Münsterberg, H. (1914). *Psychology and social sanity*. New York: Doubleday, Page, & Company.

Nelson, T. D. & Steele, R. G. (2008). Influences on practitioner treatment selection: Best research evidence and other considerations. *The Journal of Behavioral Health Services & Research, 35,* 170 - 178. doi: 10.1007/s11414-007-9089-8

Nelson-Gray, R. O. (2003). Treatment utility of psychological assessment. *Psychological Assessment, 15,* 521 - 531. doi: 10.1037/1040-3590.15.4.521

Newell, A., & Simon, H. A. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice-Hall.

Newnham, E. A., & Page, A. C. (2009). Bridging the gap between best evidence and best practice in mental health. *Clinical Psychology Review, 30,* 127 - 142. doi: 10.1016/j.cpr.2009.10.004

Nezu, C. M., & Nezu A. M. (1995). Clinical decision making in everyday practice: The science in the art. *Cognitive and Behavioral Practice, 2,* 5 - 25. doi: 10.1016/S1077-7229(05)80003-3

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2,* 175 - 220. doi: 10.1037/1089-2680.2.2.175

O'Donohue, W., Fisher, J. E., Plaud, J. J., & Curtis, S. D. (1990). Treatment decisions: Their nature and their justification. *Psychotherapy: Theory, Research, Practice, Training, 27,* 421 - 427. doi: 10.1037/0033-3204.27.3.421

Patel, V. L., Arocha, J. F., & Zhang, J. (2005). Thinking and reasoning in medicine. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning.* (p. 727). New York: Cambridge University Press.

Paul, G. L. (1967). Strategy of outcome research in psychotherapy. *Journal of Consulting Psychology, 31,* 109 - 118. doi: 10.1037/h0024436

Perry, S., Cooper, A. M., & Michels, R. (1987). The psychodynamic formulation: Its purpose, structure, and clinical application. *American Journal of Psychiatry, 144,* 543-550. Retrieved from http://ajp.psychiatryonline.org/

Persons, J. B. (1991). Psychotherapy outcome studies do not accurately represent current models of psychotherapy. *American Psychologist, 46,* 99 - 106. doi: 10.1037/0003-066X.46.2.99

Persons, J. B. (2006). Case formulation-driven psychotherapy. *Clinical Psychology: Science and Practice, 13,* 167 - 170. doi: 10.1111/j.1468-2850.2006.00019.x

Persons, J. B. & Bertagnolli, A. (1995). Inter-rater reliability of cognitive-behavioral case formulations of depression: A replication. *Cognitive Therapy and Research, 23,* 271 - 283. doi: 10.1023/A:1018791531158

Persons, J. B. & Mikami, A. Y. (2002). Strategies for handling treatment failure successfully. *Psychotherapy: Theory, Research, Practice, Training, 39,* 139 - 151. doi: 10.1037//0033-3204.39.2.139

Persons, J. B., Mooney, K. A., & Padesky, C. A. (1995). Interrater reliability of cognitive-behavioral case formulations. *Cognitive Therapy and Research, 19,* 21-34. doi: 10.1007/BF02229674

Persons, J. B., & Silberschatz, G. (1998). Are results of randomized controlled trials useful to psychotherapists? *Journal of Consulting and Clinical Psychology, 66*, 126 - 135. doi: 10.1037/0022-006X.66.1.126

Persons, J. B. & Tompkins, M. A. (2007). Cognitive-behavioral case formulation. In T. D. Eells (Ed.). *Handbook of psychotherapy case formulation* (2nd edition). New York: The Guilford Press.

Pijnenburg, H. M. (1996). *Psychodiagnostic decision-making within clinical case conferences: Exploring a domain.* Unpublished doctoral dissertation, Radboud University, Nijmegen.

Pliske, R., & Klein, G. (2003). The naturalistic decision-making perspective. In S. L. Schneider & J. Shanteau (Eds.), *Emerging perspectives on judgment and decision research.* (pp. 165-200). Cambridge: Cambridge University Press.

Plous, S. (1993). *The psychology of judgment and decision making.* New York: McGraw-Hill.

Quaestio Survey Manager. (1993). Version 5.0 rev 9475. Leiden: Cope BV.

Ruiter, C. de., & Hildebrand, M. (Eds.) (2006). *Handboek psychodiagnostiek: Van testmethode naar toepassing [Handbook psychodiagnostic practice: From testmethod to application]*. Amsterdam: Harcourt.

Sartorius, N., Kaelber, C. T., Cooper, J. E., Roper, M. T., Rae, D. S., Gulbinat, W., Ustun, T. B., & Regier, D. A. (1993). Progress toward achieving a common language in psychiatry. Results from the field trial of the clinical guidelines accompanying the WHO classification of mental and behavioral disorders in ICD-10. *Archives of General Psychiatry, 50,* 115 - 124. Retrieved from: http://archpsyc.ama-assn.org

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7*, 147-177. doi: 10.1037//1082-989X.7.2.147

Shanteau, J. (1988). Psychological characteristics and strategies of expert decision makers. *Acta Psychologica, 68*, 203-215. doi:10.1016/0001-6918(88)90056-X

Shanteau, J., Weiss, D. J., Thomas, R. P., & Pounds, J. C. (2002). Performance-based assessment of expertise: How to decide if someone is an expert or not. *European Journal of Operational Research, 136,* 253 - 263. doi: 10.1016/S0377-2217(01)00113-8

Shapiro, D. (2002). Renewing the scientist-practitioner model. *The Psychologist, 15,* 232 - 234. Retrieved from: www.thepsychologist.org.uk

Shapiro, M. B. (1967). Clinical psychology as an applied science. *British Journal of*

Psychiatry, 113, 1039 - 1042. doi: 10.1192/bjp.113.502.1039

Shapiro, M. B. (1985). A reassessment of clinical psychology as an applied science. *British Journal of Clinical Psychology, 24,*1 - 11. Retrieved from: www.bpsjournals.co.uk

Simon, H. A. (1957). *Models of man: Social and rational.* New York: Wiley.

Sloves, R. E., Doherty, E. M., & Schneider K. C. (1979). A scientific problem-solving model of psychological assessment. *Professional Psychology, 10,* 28–35. doi: 10.1037/0735-7028.10.1.28

Smith, E. E., & Medin, D. L. (1981). *Categories and concepts.* Cambridge, MA: Harvard University Press.

Sno, H. N., Hengeveld, M. W., & Beekman, A. T. F. (2004). *Richtlijn psychiatrisch onderzoek bij volwassenen [Guideline for psychiatric assessment with adults].* Amsterdam: Boom.

Spengler, P. M., Strohmer, D. C., Dixon, D. N., Shivy, V. A. (1995). A scientist-practitioner model of psychological assessment. *The Counseling Psychologist, 23,* 506 - 534. doi: 10.1177/0011000095233009

Spengler, P. M., White, M. J., Ægisdóttir, S., Maugherman, A. S., Anderson, L. A., Cook, R. S., Nichols, C. N., Lampropoulos, G. K., Walker, B. S., Cohen, G. R., & Rush, J. D. (2009). The meta-analysis of clinical judgment project: Effects of experience on judgment accuracy. *The Counseling Psychologist, 37,* 350 - 399. doi: 10.1177/0011000006295149

Stewart, R. E., & Chambless, D. L. (2007). Does psychotherapy research inform treatment decisions in private practice? *Journal of Clinical Psychology, 63,* 267 - 281. doi: 10.1002/jclp.20347

Stricker, G., & Trierweiler, S. J. (1995). The local clinical scientist: A bridge between science and practice. *American Psychologist, 50,* 995 - 1002. doi: 10.1037/0003-066X.50.12.995

Strijbos, J. -W., Martens, R. L., Prins, F. J., & Jochems, W. M. G. (2006). Content analysis: What are they talking about? *Computers & Education, 46,* 29 - 48. doi: 10.1016/j.compedu.2005.04.002

Tarrier, N., & Calam, R. (2002). New developments in cognitive-behavioural case formulation. Epidemiological, systemic and social context: An integrative approach. *Behavioural and Cognitive Psychotherapy, 30,* 311 - 328. doi: 10.1017/S1352465802003065

Thomas, R. P., Dougherty, M. R., Sprenger, A. M., & Harbison, J. I. (2008). Diagnostic hypothesis generation and human judgment. *Psychological Review, 115,* 155 - 185. doi: 10.1037/0033-295X.115.1.155

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124 - 1131. doi: 10.1126/science.185.4157.1124

Van Aarle, E., & Van den Bercken, J. (1992). The development of a knowledge-based system supporting the diagnosis of reading and spelling problems. *Computers in Human Behavior, 8,* 183 - 201. doi:10.1016/0747-5632(92)90003-W

Van Aarle, E. J. M., & Van den Bercken, J. H. L. (1999). The development of a knowledge-base system supporting the diagnosis of reading and spelling problems (ii). *Computers in Human Behavior, 15*, 693-712. doi: 10.1016/S0747-5632(99)00041-2

Van Merriënboer, J. G. J., & Sweller, J. (2009). Cognitive load theory in health profession education: Design principles and strategies. *Medical Education, 44,* 85 - 93. doi: 10.1111/j.1365-2923.2009.03498.x

Van Strien, P. (1997). Towards a methodology of psychological practice. *Theory and*

Psychology, 7, 683 - 700. doi: 10.1177/0959354397075006

Vermande, M. M. (1995). *Psychodiagnostics: Classification and hypothesis generation.* Unpublished Doctoral Thesis, Catholic University of Nijmegen.

Vermande, M. M., Van den Bercken, J. H., & De Bruyn, E. E. J. (1996). Effects of diagnostic classification systems on clinical hypothesis-generation. *Journal of Psychopathology and Behavioral Assessment, 18,* 49 - 70. doi: 10.1007/BF02229102

Vertommen, H., Ter Laak, J. J. F., & Bijttebier, P. (2005). Het diagnostisch proces [The diagnostic process]. In F. Luteijn, W. A. Arrindell, B. G. Deelman, J. H. Kamphuis & H. Vertommen (Eds.), *Psychologische diagnostiek in de gezondheidszorg [Psychological diagnosis in health care]* (pp. 15-34). Utrecht: Lemma.

Virúes-Ortega, J. & Haynes, S. N. (2005). Functional analysis in behavior therapy: Behavioral foundations and clinical application. *International Journal of Clinical and Health Psychology, 5,* 567 - 587. Retrieved from: http://www.aepc.es/ijchp/

Watkins, C. E., Campbell, V. L., Nieberding, R., & Hallmark, R. (1995). Contemporary practice of psychological assessment by clinical psychologists. *Professional Psychology: Research and Practice, 26,* 54 - 60. doi: 10.1037/0735-7028.26.1.54

Westhoff, K., Hagemeister, C., & Strobel, A. (2007). Decision-aiding in the process of psychological assessment. *Psychology Science, 49,* 271 - 285. Retrieved from: http://www.doaj.org/

Westmeyer, H. (1975). The diagnostic process as a statistical-causal analysis. *Theory and Decision, 6,* 57 - 86. doi: 10.1007/BF00139821

Widiger, T. A., & Samuel, D. B. (2005). Diagnostic categories or dimensions? A question for the diagnostic and statistical manual of mental disorders - fifth edition. *Journal of Abnormal Psychology, 114,* 494 - 504. doi: 10.1037/0021-843X.114.4.494

Witteman, C. L. M., Harries, C., Bekker, H. L., & Van Aarle, E. J. M. (2007). Evaluating psychodiagnostic decisions. *Journal of Evaluation in Clinical Practice, 13,* 10 - 15. doi: 10.1111/j.1365-2753.2006.00689.x

Witteman, C. L. M., & Koele, P. (1999). Explaining treatment decisions. *Psychotherapy Research, 9,* 100 - 114. doi: 10.1080/10503309912331332621

Witteman, C. L. M., & Kunst, H. (1997). Planning the treatment of a depressed patient. *Clinical Psychology and Psychotherapy, 4,* 157 - 171. doi: 10.1002/(SICI)1099-0879(199709)4:3<157::AID-CPP127>3.0.CO;2-8

Witteman, C. L. M. & Kunst, H. (1999). SelectCare - in aid of psychotherapists' treatment decisions. *Computers in Human Behavior, 15,* 143 - 159. doi: 10.1016/S747-5632(99)0014-X

Witteman, C. L. M. & Van den Bercken, J. H. L. (2007). Intermediate effects in psychodiagnostic classification. *European Journal of Psychological Assessment, 23,* 56 - 61. doi: 10.1027/1015-5759.23.1.56

World Health Organization (1993). *The tenth revision of the international statistical classification of diseases and related health problems (ICD-10).* Geneva: Author.

APPENDICES

APPENDIX A

Diagnostic Activities (in randomized order)

Registration	
a01	establish whether people involved are prepared to make arrangements about their contribution
a02	determine the follow up procedure in writing
a03	establish whether diagnostic examination is necessary
a04	decide whether the registration procedure can be continued
a05	check the demands for the length of the treatment
a06	establish whether the diagnostician can perform the examination
a07	establish whether the formal positions are in conflict with legal provisions
a08	determine client's motives and expectations
a09	make arrangements about the follow up procedure
a10	determine who are involved at registration
a11	determine the follow up procedure orally

Complaint analysis	
k01	check whether complaints and diagnostic questions are complete
k02	formulate the goals of the complaint analysis
k03	order complaints and diagnostic questions in importance
k04	explain the importance of the complaint analysis to the client
k05	go over the arrangements from registration
k06	record the order of the complaints and diagnostic questions in writing
k07	check interpretation of the complaints against the client's interpretation
k08	convert client's experience of the complaints into diagnostic questions
k09	check that complaints and diagnostic questions are consistent
k10	formulate the complaints
k11	explain the methods of the complaint analysis

Problem analysis	
p01	compare the clients' behaviour to dysfunctional behaviour categories in the literature (e.g. DSM)
p02	make an inventory of problem behaviours and the situations in which they occur
p03	discuss the problem analysis with colleagues
p04	establish the risk factors of the client's behaviour
p05	assign disorders to a category with the help of a classification system
p06	assess the severity of the problems
p07	describe the problem behaviour
p08	explain the classification system to the client
p09	weigh the positive and negative behaviours
p10	order the disorders

Explanation analysis	
v01	operationalise the hypotheses about the problem's explanation into testable predictions
v02	evaluate the results of testing the diagnostic explanations
v03	test the diagnostic explanations
v04	split up the diagnostic reasoning schema into testable statements
v05	check whether there is knowledge that allows the testing of the diagnostic explanations
v06	determine the degree of certainty about the results of testing the diagnostic

explanations

v07   order causal relations between problems and conditions into a preliminary diagnostic reasoning schema

v08   process the results of testing the diagnostic explanations into an integrated model

v09   perform a literature search on the causal relationships between problems and conditions

v10   establish the criteria for the testable predictions

v11   analyze the hypotheses about the explanations of the client's problem

Indication analysis

i01   select the most appropriate treatment(s)

i02   formulate concrete and specific treatment goals

i03   weigh the costs and benefits of a possible treatment

i04   check whether treatment is possible

i05   check the requirements for the length of a treatment

i06   consult the literature on treatment instruments and techniques

i07   ask client's appreciation about treatment proposals

i08   check whether treatment is necessary

i09   make an inventory of treatment instruments and techniques

i10   select type of setting for treatment

i11   formulate final global treatment goals

i12   check whether treatment is desirable

i13   weigh chance of success and failure of a possible treatment

i14   select a theoretical framework

i15   choose between a direct or indirect treatment

Diagnostic Scenario

d01   explain the follow up procedure to the client

d02   rewrite (partial) diagnostic questions briefly

d03   check sub-questions in diagnostic questions

d04   identify types of queries and types of diagnostic examinations

d05   formulate diagnostic sub-questions

APPENDIX B

Case Description

The case description below is based on the first conversation between a client and the psychologist who is treating her.

Case description Mrs. W.
Mrs. W. says that she feels she has not gotten over the death of her mother. Her mother passed away two years ago. Since a year ago, Mrs. W. often cries suddenly, she talks to her mother in her thoughts, and she often visits her mother's grave. Her mother's death also keeps her preoccupied in other situations. At work, Mrs. W. finds it hard to distance herself from the stories she hears about accidents. She notices that she has become more sensitive. She feels that she has lost her joy of living. Furthermore, she talks about the strains of taking care of her sister who has multiple sclerosis, and about the burden of her husband's alcohol addiction. They separated eight years ago, but never got a divorce. She also feels weighed down by having to take care of her father after her mother's death.
Three years ago she was hospitalized in a psychosomatic clinic for three weeks because of her problems with her husband. This did not lead to the expected relief. Mrs. W. did not open herself up to the therapeutic possibilities. Mrs. W. is still married, but she doesn't live together with her husband, although they have three sons together. The eldest was born in 1972. Two years ago she started a new relationship but she became less involved with her new partner the past few weeks. The past years she has been taking Oxazepam in stressful situations because of her restlessness and sleeping disorders. There is no regular intake.
The past 15 years Mrs. W. has been working in the administration department of a physical therapy practice. At the moment she works 28 hours a week.

Diagnostic Interview Coding Schema

Category	Description	Examples
Dutch Guideline Psychiatric Assessment (Sno et al., 2004)		
Complaints and symptoms	utterances about complaints reported by the client, complaints inferred by the psychologist, symptoms of depression, libido, suicidal thoughts, manic episodes, psychotic features, anxiety	"So you feel like you are an old man already?" "You are going through some intense experiences." "Do you feel sad all day long?" "You are completely convinced that your life is worthless."
Classification	utterances about type and severity of the disorder, the client's awareness of illness or a differential disorder	"Your depression can also be seen as an illness." "I am sure you are suffering from a depression."
Psychiatric history	utterances about previous psychiatric treatment	"Did you ever experience a depressive episode before?"
Family history	utterances about the dates of birth or death of family members, the psychiatric or physical history of family members	"Does depression run in the family?" "Is the arteriosclerosis a hereditary condition?"
Physical history	utterances about previous or current physical illnesses	"Do you know the prognosis of the operation on your legs?"
Social history	utterances about the client's current living situation, relationship with life partner, children or family, social support or functioning, work, hobbies	"How is your daughter doing at the moment?" "Do you have any serious debts?" "You mentioned sports means a lot to you."
Biography	utterances about client's course of life, family origins, life history, stressful life events	"Do you have any siblings?" "What kind of education have you had?"
Personality	utterances about client's personality traits, coping skills, defense mechanisms	"Do you recognize yourself as such a person?" "Are you usually an optimistic person?"
Treatment	utterances about the client's expectations or the psychologist's expectations about treatment options	"Would you be willing to take antidepressants?" "Depression can be treated."
Meta-cognitive & Miscellaneous		
Other	utterances to foster or maintain the therapeutic relationship; utterances about personal details; reason for referral; symptoms of other psychiatric disorders; substance abuse; medication; non-verbal behaviour; duration of the interview; awareness of the test situation; supportive remarks; non-classifiable remarks.	

APPENDIX D

Stimulated Recall Interview Coding Schema

Category	Description	Examples
Diagnostic Cycle (De Bruyn et al., 2003)		
Complaints and symptoms	thoughts about identifying or evaluating the client's verbal and non-verbal behaviour, the client's complaints and symptoms and suicidal tendencies	"I noticed his date of birth." "His posture, his presentation." "He really looks depressed." "I should ask about his suicidal thoughts."
Classification	thoughts about identifying the disorder or evaluating its severity; thoughts about identifying a differential disorder	"He has been depressed for about two months now." "Possibly a personality disorder as well."
Potential stressors	thoughts about recent and previous stressful events, personality traits or biological/genetic dispositions that may have contributed to the development of the disorder	"I thought about the extent to which his illness played a role." "To what extent did he get support from his father when he was young?" "He seems to have a hereditary defect."
Protective factors	thoughts about social support, personality traits or biological/genetic dispositions that may have protected the client from developing the disorder	"I try to find out more about the support he gets from his family." "He does have a good job."
Explanatory mechanism	thoughts about a psychological, biological or socio-cultural mechanism that explains the (development of) the client's disorder	"His sense of worthlessness was increased by his own thoughts about the conflict at work." "I wondered whether his cycling accident involved a head injury which might have caused his mood change."
Treatment	thoughts about identifying or evaluating a future therapeutic intervention, suicide prevention, the therapeutic intervention's intensity, medication, the client's motivation for treatment, performing therapeutic interventions during the diagnostic interview	"His condition can certainly be treated." "Do you let him go back home or do you hospitalize him?" "He needs to have some future prospects or hope." "I thought: I should offer him some future prospects."
Meta-cognitive & Miscellaneous		
Other	thoughts about fostering or maintaining the therapeutic relationship; further assessment; own method of working; (counter)transference; the duration of the interview; awareness of the test situation; non-classifiable remarks.	

Descriptions of the Quality Coding Categories.

Category	Description
Form	
Simple hypothesis	A hypothesis containing only one direct factor. A direct factor immediately precedes or follows the part to be explained (explanandum).
Composite hypothesis	A hypothesis containing two or more direct factors (no indirect factors).
Causal chain	A hypothesis containing one or more direct factors and at least one indirect factor linked together in a linear and unidirectional way.
Causal model	A hypothesis containing several direct and indirect factors linked together in a non-linear and bidirectional way.
Relations	Specification of the kind of relationship between two indirect or direct factors or between a direct and an indirect factor. For example: 'strengthen', 'maintaining'. Statements such as 'leads to' or 'causes' are not sufficient.
Relevance	Information in the hypothesis should be linked to either information from the vignette or the client's complaint(s).
Consistency	A hypothesis should be an actual explanation for the problem, i.e. it should not be circular, contradictory or only a restatement of the problem.
Specificity	A hypothesis should have sufficient depth and does not require an explanation of the explanatory factor itself.
Testability	A testable hypothesis is falsifiable, i.e. it is clear what would constitute a counterexample to the hypothesis, and observation of a counterexample would have to be practically feasible.
Modifiability	
Not modifiable	A factor that is remote in time and that the client cannot influence him or herself.
Indirectly modifiable	A factor that is not remote in time and concerns someone from the client's direct environment or some part of the living circumstances (e.g. work environment) that the client can influence partly.
Directly modifiable	A factor that lies in the present and can be influenced by the client him or herself through a behavioural change (e.g. a client's exaggerated perfectionism).
Positive indicators	
No positive indicators	Absence of positive indicators for treatment such as adaptive skills, traits of the self, perceptions of self or others, whishes, goals and hopes, positive motivation for treatment, good social support or progress already achieved by the client.
1 or more positive indicators	Presence of at least one of the abovementioned positive indicators for treatment.

NEDERLANDSE SAMENVATTING

HET BESLISSENDE MOMENT:
DIAGNOSTISCHE BESLISSINGEN EN HET ONTWERPEN VAN BEHANDELINGEN

In de dagelijkse praktijk van de klinisch psycholoog wordt een optimale uitvoering van de diagnostische taak bemoeilijkt door de complexe en dynamische situatie (Gambrill, 2005). Beperkte tijd en het overschrijden van de cognitieve capaciteiten van de psycholoog dragen bij aan het inperken van een optimale taakuitvoering (Garb, 1998). Daarom is het begrijpelijk dat psychologen ook op andere bronnen vertrouwen dan het uitgebreid opstellen en toetsen van verschillende hypotheses om het diagnostisch proces goed te laten verlopen. Zij kunnen daarbij bijvoorbeeld gebruik maken van hun eigen theorieën over mentale stoornissen en hun oorzaken (Kim & Ahn, 2002), de theoretische oriëntatie waarin ze opgeleid zijn (Witteman & Koele, 1999) en eerdere ervaringen met vergelijkbare cliënten (Garb, 1996). Echter, prescriptieve diagnostische modellen schrijven voor dat psychologen via objectieve observatie en het genereren en toetsen van hypotheses, opgesteld op basis van wetenschappelijke kennis, tot een geïntegreerd beeld van de cliënt komen (cf. Nezu & Nezu, 1995). Tot nu toe is het nog steeds onduidelijk in hoeverre psychologen in de praktijk richtlijnen volgen van prescriptieve diagnostische modellen, gebaseerd op de wetenschappelijke methode, die zij in hun opleiding geleerd hebben.

Volgens deze modellen moeten psychologen de problematiek van een cliënt classificeren en een verklaring opstellen voor de problemen om op basis van deze twee besluitvormingsprocessen tot een behandelvoorstel te komen. Onder classificeren wordt het beschrijven van de soort en de ernst van de problemen van een cliënt verstaan en het onderbrengen van deze problemen in één of meerdere psychische stoornissen (Krol, De Bruyn, & Van den Bercken, 1992). Onder het opstellen van een verklaring wordt het beschrijven van de relaties tussen factoren die de problemen veroorzaakt hebben of in stand houden verstaan (Haynes & Williams, 2003).

Prescriptieve diagnostische modellen zijn gebaseerd op de veronderstelling dat het volgen van de wetenschappelijke methode tot betere beslissingen leidt. Deze modellen schrijven voor dat het diagnostisch proces grondig en compleet uitgevoerd wordt: de problematiek wordt in kaart gebracht, oorzaken voor de problemen worden achterhaald en op basis daarvan wordt een behandelplan opgesteld (De Bruyn, Ruijssenaars, Pameijer, & Van Aarle, 2003; Nezu & Nezu, 1995). Deze veronderstellingen tezamen impliceren dat het behandelplan afhangt van de uitkomst van het diagnostisch proces en dat deze uitkomst wederom afhangt van de soort beslissingen die genomen wordt in het proces. In dit proefschrift wordt de rol van diagnostische besluitvormingsprocessen bij het opstellen van een behandelplan onderzocht. Het doel van dit onderzoek is het beantwoorden van de volgende twee onderzoeksvragen die afgeleid zijn van deze veronderstellingen:

1. *Wat kenmerkt het diagnostisch besluitvormingsproces in de klinische praktijk?*
2. *Wat is de rol van diagnostische besluitvormingsprocessen bij het opstellen van een behandelplan?*

In de vier studies beschreven in hoofdstukken 2 tot en met 5 van dit proefschrift zijn de besluitvormingsprocessen van psychologen met verschillende methoden onderzocht. Door verschillende methoden te gebruiken kunnen de resultaten verkregen met ene methode gecontroleerd worden met de andere methode en worden de beperkingen van elke methode afzonderlijk gecompenseerd. In de studies zijn gevalsbeschrijvingen, een diagnostisch intake interview met een simulatiepatiënt en *stimulated recall* gebruikt om het diagnostisch

besluitvormingsproces en de rol van diagnostische beslissingen bij het opstellen van een behandelplan te bestuderen.

METHODEN EN BEVINDINGEN

In hoofdstukken 2 en 3 zijn de soort beslissingen die psychologen in het diagnostisch proces overwegen onderzocht (hoofdstuk 2) en de volgorde waarin deze beslissingen overwogen worden (hoofdstuk 3). In deze twee studies werd gebruik gemaakt van een vragenlijst met een gevalsbeschrijving en voorgestructureerde antwoordmogelijkheden. Het doel van deze studies was de kenmerken van de besluitvormingsprocessen van psychologen te verkennen.

Beoordeling van diagnostische besluitvormingsactiviteiten

In hoofdstuk 2 wordt de analyse beschreven van de soort en aantal besluitvormingsactiviteiten die psychologen nodig achten of uit zouden voeren. Eén groep psychologen beoordeelde hoe nodig de activiteiten waren, de andere groep of ze deze zouden uitvoeren. Psychologen vulden aan de hand van een gevalsbeschrijving een vragenlijst in. Beide vragenlijsten bestonden uit besluitvormingsactiviteiten die gegroepeerd waren in zes types beslissingen: registratie, klachtenanalyse, probleemanalyse, verklaringsanalyse, indicatieanalyse en diagnostisch scenario. De activiteiten zijn afgeleid van de Diagnostische Cyclus (DC) van De Bruyn en anderen (2003).

Psychologen vonden besluitvormingsactiviteiten vaker nodig dan dat ze deze zouden uitvoeren. Besluitvormingsactiviteiten van de beslissingen registratie, klachtenanalyse, probleemanalyse en indicatieanalyse werden vaker nodig gevonden dan dat deze uitgevoerd zouden worden. Verder bleek dat de besluitvormingsactiviteiten van de verklaringsanalyse het minst vaak nodig gevonden werden en ook het minst vaak uitgevoerd zouden worden.

Enkele deelnemers van deze studie gaven aan dat ze het beoordelen van een gevalsbeschrijving geen valide afspiegeling vinden van hun dagelijkse praktijk. Echter, psychologen zijn niet onbekend met het werken met gevalsbeschrijvingen. In teambesprekingen worden ook vaak besluiten genomen op basis van gevalsbeschrijvingen. Daarnaast is de gevalsbeschrijving die gebruikt is in deze studie gebaseerd op een bestaande cliënt en werd de beschrijving opgesteld door een psycholoog die deze cliënt beoordeeld en behandeld heeft.

Opbouw van het diagnostisch besluitvormingsproces

In hoofdstuk 3 wordt de analyse beschreven van de volgorde van diagnostische beslissingen die psychologen nemen. Tevens wordt de overeenstemming van psychologen met de voorgeschreven volgorde van beslissingen volgens de DC bekeken en de overeenstemming tussen psychologen onderling over de volgorde van beslissingen. Psychologen gaven aan in welke volgorde ze de zes types beslissingen van de DC zouden uitvoeren.

Psychologen zouden eerst de klachten van de cliënt analyseren. Vervolgens zouden ze de problemen van de cliënt classificeren, oftewel de symptomen bij een stoornis onderbrengen. Als laatste zouden zij een passende behandelvorm selecteren. De plaats van de verklaringsanalyse in het diagnostisch proces is minder duidelijk: sommige psychologen zouden deze beslissing als één van de eerste uitvoeren, anderen als één van de laatste. Verder was de overeenstemming tussen de voorgeschreven volgorde van

108

beslissingen en de volgordes aangegeven door de psychologen laag evenals de overeenstemming tussen psychologen onderling over de volgorde van de beslissingen.

Besluitvorming tijdens een diagnostisch interview

In hoofdstuk 4 wordt de analyse beschreven van de soort en de volgorde van beslissingen die psychologen nemen en overwegen tijdens een interview met een simulatiepatiënt. Daarnaast hebben we onderzocht in hoeverre deze overwegingen gerelateerd zijn aan de beslissingen die de psychologen beschreven in hun verslagen van de interviews. De psychologen interviewden allemaal maximaal een half uur dezelfde simulatiepatiënt. Na dit interview vatten ze hun bevindingen samen in de vorm van een DSM-IV classificatie, een verklarende diagnose en een behandelplan. Daarna werd een *stimulated recall* interview gehouden waarbij de psychologen hun eigen diagnostisch interview terugkeken en reflecteerden op de uitvoering van het interview. De aanwezigheid van de beslissingen klachtenanalyse, classificatie, verklarende diagnose en behandelplan en de volgorde van deze beslissingen werd voor het diagnostisch interview en *stimulated recall* interview bekeken en met elkaar vergeleken. Tevens werd de overeenstemming van de inhoud van deze beslissingen tussen het *stimulated recall* interview en het verslag onderzocht.

Om de soort beslissingen en de volgorde ervan tijdens het diagnostisch interview te analyseren is een codeerschema opgesteld, gebaseerd op de Nederlandse Richtlijn voor Psychiatrisch Onderzoek met Volwassenen van Sno, Hengeveld en Beekman (2004). Psychologen stelden gedurende het gehele interview vragen over elke soort beslissing. Het vaakst stelden de psychologen vragen over klachtenanalyse en verklarende diagnose. Vragen over het behandelplan namen toe aan het einde van het interview.

Om de soort beslissingen en de volgorde ervan tijdens het *stimulated recall* interview te analyseren is een codeerschema opgesteld gebaseerd op de DC van De Bruyn en anderen (2003). Ook hier bleek dat psychologen gedurende het gehele interview reflecteerden op elke soort beslissing. Het vaakst reflecteerden ze op de beslissing klachtenanalyse. Naarmate het interview vorderde namen reflecties op het behandelplan toe.

Om de consistentie van de inhoud van de beslissingen classificatie, verklarende diagnose (stressoren en verklarende mechanismen) en behandelplan (behandeldoelen en -methoden) tussen het *stimulated recall* interview en de verslagen te bestuderen is de inhoud van deze beslissingen onderzocht en gecategoriseerd. Er was een redelijke overeenstemming tussen de inhoud van de mogelijke stressoren die overwogen waren tijdens het *stimulated recall* interview en die beschreven waren in de verslagen. De consistentie van de inhoud van de classificaties, verklarende mechanismen, behandeldoelen en -methoden was matig tot laag. De overeenstemming tussen psychologen onderling over classificatie was hoog, terwijl die voor mogelijke stressoren, verklarende mechanismen behandeldoelen en -methoden matig tot laag was.

Het interview met de simulatiepatiënt was zo opgezet dat het zoveel mogelijk op de dagelijkse praktijk zou lijken. Zo werd de psychologen vooraf niet verteld of ze een acteur of een echte cliënt zouden spreken. Verder werd hen gevraagd het interview te houden zoals ze dat in hun eigen praktijk gewend waren en konden ze aantekeningen maken tijdens het interview. Hiermee werd de geloofwaardigheid van de taaksituatie vergroot.

Enkele deelnemers twijfelden aan de geloofwaardigheid van de situatie: zij waren er niet helemaal van overtuigd dat ze mogelijkerwijs met een echte cliënt te maken hadden. Daarnaast gaven enkele deelnemers aan dat ze zich onprettig voelden door de aanwezigheid

van een camera voor de opname van het interview. De overige deelnemers beoordeelden het interview een betrouwbare weergave van de dagelijkse praktijk.

Complexiteit van de problematiek en klinische besluitvorming

In hoofdstuk 5 wordt de analyse beschreven van het effect van de complexiteit van de problematiek van een cliënt op de kwaliteit van de verklarende diagnoses. Tevens hebben we de relatie onderzocht van de diagnostische beslissingen classificatie en verklarende diagnose en van de theoretische oriëntatie van de psychologen met de behandelbeslissingen. Om dit te onderzoeken hebben we een vragenlijst opgesteld met twee gevalsbeschrijvingen die varieerden in complexiteit van de problematiek (waarbij complexiteit verwijst naar de bekendheid met de problematiek en de beschikbaarheid van een empirisch gevalideerde behandelmethode). Voor elke gevalsbeschrijving selecteerden de psychologen uit een lijst een classificatie, gaven ze een verklarende diagnose en selecteerden ze vervolgens uit een lijst een behandelmethode.

Om de invloed van de complexiteit van de problematiek op de kwaliteit van de verklarende diagnoses te onderzoeken hebben we eerst de kwaliteit van de diagnoses bekeken. Daarvoor hebben we een codeerschema ontwikkeld gebaseerd op het schema van Kuyken, Fothergill, Musa en Chadwick (2005) en op werk door Vermande (1995) en Vermande, Van den Bercken en De Bruyn (1996). De verklarende diagnoses werden opgedeeld in één of meer verklarende hypotheses die vervolgens beoordeeld werden op zes dimensies: (1) vorm (enkelvoudige/samengestelde hypothese, causale keten/causaal model), (2) relaties tussen causale factoren, (3) relevantie, (4) consistentie, (5) veranderbaarheid van causale factoren (niet, indirect, direct) en (6) positieve indicatoren voor behandeling (aanwezig, afwezig). De verklarende hypotheses voor de minst complexe casus waren vaker relevant, ze bevatten vaker beschrijvingen van relaties tussen causale factoren en van direct veranderbare factoren en minder vaak van niet veranderbare factoren.

Om de relatie van de verklarende diagnoses met de behandelbeslissingen te onderzoeken, hebben we de inhoud van de verklarende diagnoses bekeken. We ontwikkelden een codeerschema gebaseerd op de *Case Formulation Content Coding Method* van Eells, Kendjelic en Lucas (1998) waarmee de inhoud van de verklarende mechanismen, indien aanwezig in de verklarende diagnose, onderzocht werd. De diagnostische beslissingen, classificaties dan wel verklarende mechanismen, en de theoretische oriëntaties waren niet of nauwelijks gerelateerd aan de behandelbeslissingen. De observeerbare klachten van de cliënt waren het sterkst geassocieerd met de behandelbeslissing, hoewel de sterkte van deze associatie matig was.

CONCLUSIES & DISCUSSIE VAN DE BEVINDINGEN

Psychologen lijken hun diagnostische besluitvormingsprocessen aan te passen aan de situatie waarbij psychologen zich met name richten op het analyseren van de klachten en symptomen van de cliënt en het opstellen van het behandelplan. Ze wisselen alle soorten beslissingen continu af en heroverwegen voorgaande beslissingen (cf. Bartolo, Dockrell, & Lunt, 2001; Pijnenburg, 1996). Classificaties en verklarende diagnoses zijn nauwelijks gerelateerd aan de behandelplannen (cf. Bus & Kruizenga, 1989; Witteman & Koele, 1999).

De meest opvallende bevinding uit de eerste drie studies is dat psychologen zich met name richten op de klachtenanalyse en classificatie en minder op de verklarende diagnose. Een mogelijke verklaring is dat een analyse van klachten en symptomen duidelijke

positieve effecten heeft. Zo bevordert een klachtenanalyse de therapeutische relatie (De Bruyn et al., 2003) en classificatie maakt het mogelijk om wetenschappelijke informatie over vergelijkbare gevallen toe te passen. Daarnaast zou een uitgebreide verklarende diagnose onnodig zijn omdat psychologen gebruik maken van impliciete, causale theorieën in de vorm van cognitieve schema's. Psychologen toetsen klachten en symptomen van een bepaalde cliënt aan hun causale theorieën (Kim & Ahn, 2002). Komen klachten en symptomen overeen met een causale theorie dan is expliciet zoeken naar mogelijke verklarende diagnoses overbodig. Een andere verklaring voor de bevinding is dat psychologen niet in staat zijn om informatie in een causaal model te representeren omdat de benodigde etiologische informatie daarvoor ontbreekt (Stricker & Trierweiler, 1995; Cicchetti & Sroufe, 2000). Als een empirisch gevalideerde theorie over de oorzaken van een stoornis bekend is dan wel dat een empirisch gevalideerde behandelmethode voor handen is, dan zouden psychologen eerder geneigd zijn om deze informatie te gebruiken voor een verklarende diagnose. Deze verklaring wordt ondersteund door de bevindingen in hoofdstuk 5 waarin psychologen kwalitatief betere verklarende diagnoses opstellen voor de problematiek waarvoor een empirisch gevalideerde theorie beschikbaar was. Een laatste mogelijke verklaring voor de bevinding dat psychologen minder aandacht besteden aan verklarende diagnoses is dat psychologen het opstellen van een verklarende diagnose een veeleisende en tijdrovende bezigheid vinden (cf. Virúes-Ortega & Haynes, 2005).

De diagnostische beslissingen die psychologen overwegen tijdens het diagnostisch proces, of het nu classificaties of verklarende diagnoses zijn, blijken nauwelijks samen te hangen met de behandelplannen. De keuze voor een behandelmethode lijkt gebaseerd te zijn op de cognitieve schema's van psychologen. In deze schema's zijn bepaalde klachten en symptomen gekoppeld aan effectief gebleken behandelmethoden bij eerdere cliënten (cf. Witteman & Koele, 1999). Een andere mogelijke verklaring is dat psychologen gebruik maken van andere bronnen, zoals het beleid van de instelling waar ze werken, persoonlijke overtuigingen of *evidence based* richtlijnen (cf. Nelson & Steele, 2008).

Over het algemeen is de overeenstemming over diagnostische beslissingen tussen psychologen laag (Garb, 1998). In dit proefschrift was de overeenstemming tussen psychologen alleen voor de classificatie van een vaak voorkomende stoornis met duidelijk observeerbare symptomen (namelijk depressieve stoornis) hoog. Psychologen konden bij het classificeren mogelijk terugvallen op de kennis en ervaring met de diagnostische criteria van de DSM-IV die ze in hun werk en opleiding opgedaan hebben. Dergelijke handboeken ontbreken voor het opstellen van verklarende diagnoses en behandelplannen. Daardoor hebben psychologen voor die besluitvormingsprocessen hun eigen persoonlijke strategieën ontwikkeld (heuristieken) om beslissingen te kunnen nemen. Verschillen in deze heuristieken, ontstaan door een mix van ervaring, theoretische oriëntatie en eerdere cliënten, kunnen de lage overeenstemming over verklarende diagnoses en behandelplannen verklaren.

Uit de hoofdstukken 2 en 3 komt naar voren dat de ervaring van een psycholoog mogelijk de soort beslissingen die overwogen wordt en de volgorde van deze beslissingen beïnvloedt. Het effect was zwak in deze studies, wat bevestigd wordt door de studie van Spengler e.a. (2009) naar de invloed van ervaring op het diagnostisch proces en de methodologische beperkingen van de studies (zie hiervoor de discussies van hoofdstukken 2 en 3).

IMPLICATIES VOOR PRAKTIJK, BESLISSINGSONDERSTEUNING EN VERDER ONDERZOEK

Psychologen die afwijken van de voorgeschreven besluitvormingsprocessen van diagnostische modellen lopen het risico dat ze informatie over het hoofd zien of teveel waarde hechten aan irrelevante informatie en daardoor een onjuiste beslissing nemen (cf. Nezu & Nezu, 1995). Daarnaast kan het overwegen van verschillende behandelmogelijkheden aan het begin van het diagnostisch proces ertoe leiden dat psychologen vroegtijdig en op basis van onvolledige informatie al tot definitieve conclusies komen. Echter, psychologen heroverwegen ook veelvuldig hun eigen beslissingen waardoor deze zogenoemde 'premature closure' voorkomen kan worden.

De methode die in dit proefschrift gebruikt is om het diagnostisch besluitvormingsproces te bestuderen kan ook in de praktijk ingezet worden. Het codeerschema met de onderverdeling in klachtenanalyse, classificatie, verklarende diagnose en behandelplan kan gebruikt worden om het eigen diagnostisch besluitvormingsproces onder de loep te nemen en de meta-cognitieve vaardigheden te verbeteren. Volgens Barrows (2000) is de *stimulated recall* procedure geschikt om mogelijke problemen in het redeneerproces en verminderde uitvoering van het diagnostisch proces aan het licht te brengen. De combinatie van het codeerschema en *stimulated recall* bieden mogelijkheden voor de opleiding en verdere training van psychologen.

Om psychologen te ondersteunen bij het uitvoeren van de complexe en dynamische diagnostische taak moet beslissingsondersteuning ontwikkeld worden. Deze beslissingsondersteuning zou het besluitvormingsproces van de verklarende diagnose moeten ondersteunen omdat (i) psychologen besluitvormingsactiviteiten van deze beslissing minder belangrijk vinden, (ii) psychologen minder vaak reflecteren op deze beslissing, (iii) de betrouwbaarheid van verklarende diagnoses laag is (cf. Eells, Kendjelic, & Lucas, 1998) en (iv) de kwaliteit van verklarende diagnoses onder bepaalde omstandigheden laag is (cf. Kuyken, Fothergill, Musa, & Chadwick, 2005). De kennisbasis over de oorzaken van stoornissen is beperkt en daarom zou beslissingsondersteuning het beste kunnen bestaan uit visualisatie van de relevante causale factoren en hun relaties in een causaal model (cf. Morton, 2004). De 'causaal modeleren' methode van Morton (2004) is daarvoor zeer geschikt omdat dit model opgesteld kan worden samen met de cliënt. Daardoor kan een verbeterd begrip van de problematiek ontstaan bij zowel psycholoog als de cliënt en bevordert het de instemming van de cliënt met het behandelplan.

De studies in dit proefschrift geven slechts gedeeltelijk inzicht in de kenmerken van het diagnostisch besluitvormingsproces en de rol van dit proces bij het opstellen van behandelplannen. Ten eerste is onduidelijk gebleven hoe psychologen informatie over klachten en symptomen organiseren in cognitieve schema's en in hoeverre deze schema's de besluitvorming bepalen. Ten tweede zijn zorgvuldig opgezette onderzoeken naar de invloed van ervaring op het diagnostisch besluitvormingsproces van belang om verschillen in de besluitvormingsprocessen te verklaren. Als laatste is de waarde van diagnostische beslissingen voor het opstellen van behandelplannen onbeslist. Het is nog maar de vraag of een grondig en compleet diagnostisch besluitvormingsproces daadwerkelijk tot betere behandelplannen leidt. Onderzoek naar de effectiviteiten van heuristieken (cf. Gigerenzer & Brighton, 2009) laat zien dat mentale 'short-cuts' tot goede uitkomsten kunnen leiden. Maar voordat de waarde van verschillende besluitvormingsstrategieën bepaald kan worden wacht de zware taak om overeenstemming te bereiken over wat een goede uitkomst eigenlijk is.

DANKWOORD

Dit proefschrift was niet mogelijk geweest zonder de hulp en wijze raad van velen. Een aantal van deze personen wil ik hier daarvoor bedanken.

Jules, bedankt dat je mij de mogelijkheid hebt geboden om bij jou te promoveren. Ik kan bij jou altijd terecht met mijn vragen, twijfels en ideeën. Jouw eindeloze enthousiasme en steun voor alles wat ooit ik wil(de) ondernemen en dat wat ik uiteindelijk ondernomen heb is ongekend.

Cilia, bedankt dat je me onder je hoede genomen hebt bij de Diagnostic Decision Making groep. Hoewel Nijmegen echt vreselijk ver weg is, voelde het altijd alsof je heel dichtbij was. Je betrokkenheid en inzet maar vooral ook je kritische vragen bij mijn plannen zijn erg stimulerend.

Pascal en Casper, jullie bedank ik samen, hoewel jullie allebei op geheel eigen wijze mij begeleid hebben. Jullie hebben mij geholpen bij het verbeteren van mijn werk door mij continu uit te dagen de lat hoger te leggen. Daarnaast hebben jullie mij ook veel geleerd over de ongeschreven regels van de academische wereld.

Vos, samenwerken met jou aan de meest omvangrijke studie van mijn proefschrift was inspirerend. Ik ben dankbaar dat jij jouw veelzijdige kennis en ervaring met mij wil delen en op die manier mijn werk verdieping geeft.

Leontien, via Cilia leerde ik jou en je werk kennen. De discussies en "vragenuurtjes" met jou hebben me elke keer weer vele nieuwe inzichten opgeleverd.

Guido & Elian, bedankt dat jullie mijn paranimfen willen zijn, ik kan jullie steun meer dan ooit gebruiken. Dank voor alle momenten gevuld met koffie leuten, het meedenken over onderzoek en bespreken van (het werk van) studenten. Het heeft mijn academische werk zeker dragelijker gemaakt.

Sonja Lehmann, Anja Kuijer, Lotte Croese, Hilde Mees, Milou Kievik, Marjon Stijntjes, Ludwig Fritzsch, Petra Hagens, Helena Boering & Gery Zoer: dank voor het immense werk dat jullie verzet hebben bij het verzamelen en verwerken van de onderzoeksgegevens.

De studies beschreven in dit proefschrift waren natuurlijk nooit mogelijk geweest zonder de medewerking van vele psychologen. Met name de deelnemers die persoonlijk contact met mij opnamen voor opmerkingen en/of aanmoedigingen wil ik bedanken voor hun moeite. Daarnaast wil ik René van Zinnicq Bergman bedanken voor zijn bijzondere inzet als simulatiepatiënt tijdens één van de studies.

Ook ben ik dank verschuldigd aan de promovendi van de ProIST bijeenkomsten, de GW-PhD vergaderingen/borrels/etentjes en die van de ICO cursussen voor het meedenken over mijn onderzoek. Daarnaast natuurlijk ook dank daarvoor aan mijn gewaardeerde collega's van IST en de DDM groep uit Nijmegen. Speciale dank voor Rianto Purbojo en José Kooken die het met mij als kamergenote wisten uit te houden.

CURRICULUM VITAE

Marleen Groenier werd op 29 september 1979 te Groningen geboren. In 1997 behaalde zij haar VWO diploma aan het Praedinius Gymnasium te Groningen, waarna zij een jaar Engelse literatuur en Psychologie heeft gestudeerd aan de Bangor University te Wales, Groot-Brittannië. De studie Psychologie heeft zij vervolgd aan de RijksUniversiteit Groningen en deze na zes jaar met goed gevolg afgesloten. In 2005 begon zij aan haar promotieonderzoek aan de Universiteit Twente, Faculteit Gedragswetenschappen, bij de vakgroep Instructietechnologie. Tijdens haar promotieonderzoek heeft zij haar werk gepresenteerd op internationale congressen en workshops (EGPROC, workshops Causal Reasoning in Clinical Decision Making, European Conference on Psychological Assessment). Daarnaast heeft zij als werkcollege- en hoorcollegedocente onderwijs gegeven aan eerstejaars psychologiestudenten alsmede tweedejaars, bachelor en master studenten begeleid bij het uitvoeren en verslagleggen van hun onderzoeksopdrachten. Op dit moment werkt zij nog steeds bij de vakgroep Instructietechnologie aan dezelfde faculteit waar zij haar taken als docente voortzet.