The development and function of illness scripts: studies on the structure of medical diagnostic knowledge

Citation for published version (APA):

Document status and date:
Published: 01/01/1995

DOI:
10.26481/dis.19950601ec

Document Version:
Publisher's PDF, also known as Version of record

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.umlib.nl/taverne-license

Take down policy
If you believe that this document breaches copyright please contact us at: repository@maastrichtuniversity.nl providing details and we will investigate your claim.
CHAPTER 6

SUMMARY OF THE RESULTS AND GENERAL DISCUSSION

INTRODUCTION

We will start this chapter by presenting a concise summary of the main results of the studies described in Chapters 2 through 5. In the subsequent sections, the implications of the present findings for the representation and structure of medical knowledge and medical problem solving, for the development of this knowledge, and for medical education will be discussed. Finally, some suggestions for future research will be presented.

SUMMARY OF THE PRESENT RESULTS

In the previous chapters, four studies were described that aimed at investigating the development and structure of clinical medical knowledge, with a particular emphasis on diagnostic knowledge. In the introductory chapter, it was argued that this kind of knowledge might be represented as illness scripts, i.e., script- or schema-like structures, that are activated as a consequence of the information available in the clinical diagnostic situation, and that are instantiated with respect to information about the specific patient at hand. The expression "illness script" was first expanded upon by Feltovich and Barrows (1984), with the concept being relatively independently defined, i.e., as an elaboration of a template structure, without reference to the work of Schank and Abelson (1977) on scripts, nor to recent studies on schemas (e.g., Graesser & Nakamura, 1982). Feltovich and Barrows (1984) were apparently inspired by recent efforts to implement medical expertise in an expert system, like the causal-rule based expert system MYCIN (cf. Clancey, 1983); hence, they conceived of illness scripts more as causal mental models than as "true" scripts. However, Feltovich and Barrows (1984) acknowledged the importance of Enabling Conditions (contextual factors and patient characteristics) causing a certain Fault (abnormality or malfunction in the human body), which in turn results in particular Consequences (complaints, signs and symptoms). Though this distinction between three script components has proven to be a highly fruitful idea, the causal model view of the illness script has not received that much empirical support: causal, biomedical reasoning turned out to be in general not a salient feature of expert problem solving, at least not in routine cases (cf. Boshuizen & Schmidt, 1992; Clancey, 1983; Norman et al., 1989; Patel & Groen, 1986a; Schmidt & Boshuizen, 1993a). However, the fact that Feltovich and Barrows (1984) opted for the expression "illness scripts" has not remained without consequences; in the years following the introduction of the concept, a number of notions associated with the original Schank and Abelson' (1977) scripts, have been transferred to the medical domain, for example the idea that an illness script can be selected and instantiated in a diagnostic situation (cf. Schmidt et al., 1990).

The work of Hobus and his colleagues (Hobus et al., 1987, 1989, 1990) has revealed an important aspect concerning the development of the presumed illness scripts during the course of medical education. These authors report two important findings: First, that experienced physicians are much more able to take En-
abling Conditions into account in activating early diagnostic hypotheses than beginning physicians, and second, that experienced physicians are generally inclined to mention relatively more Enabling Conditions than their less experienced colleagues when asked to describe a prototypical patient with a specific disease. Based on these results and related work on medical expertise (e.g., Allen, Brooks, & Norman, 1988; Patel et al., 1989), a theory of the development of clinical medical knowledge was constructed (cf. Schmidt et al., 1990). This theory states that in the preclinical phase of the study, disease information is accumulated, and an enormous network of medical knowledge is constructed. During the clerkship phase, in which students walk the wards, this knowledge is tuned to use in practical situations, and gradually full-fledged illness scripts develop. However, knowledge of Enabling Conditions is thought to show a considerable lag or delay compared to knowledge of the Consequences, as manifested by the relative inability of beginning physicians to exploit knowledge of patient contextual factors, and the relatively smaller contribution of Enabling Conditions in their descriptions of prototypical patients (Hobus et al., 1987, 1989). It is not quite clear why this developmental lag actually occurs, but the assumption is that students may be inclined to focus especially on the complaints, signs and symptoms associated with a disease, because these aspects are generally used to confirm diagnostic hypotheses and to alleviate diagnostic uncertainty, while Enabling Conditions are less conclusive in this respect. The tendency of medical textbooks to emphasize Consequences may also contribute to this developmental divergence of Enabling Conditions and Consequences. Furthermore, Schmidt and Boshuizen (1993a) and Schmidt et al. (1990) allow for a further specialization of clinical knowledge structures into "instance scripts", i.e., memories of individual patients that are thought to play an important role in expert future diagnoses of similar cases.

The studies described in this thesis were designed with two objectives in mind: First, to assess the "scripted" qualities of illness scripts, and second, to further investigate the proposed developmental course. As far as the first issue concerns, simply naming a hypothesized knowledge structure a script is obviously not sufficient for it to "be" a script; it is necessary to show that the features characteristic of "classical" scripts also apply to illness scripts. If experimental evidence can be gathered that is consistent with script-based predictions, this can be interpreted as showing the adequacy of the script representation for that particular knowledge. With respect to the second issue, if the development of knowledge of Consequences indeed precedes development of knowledge of Enabling Conditions, then this might be reflected by differences in a number of both process and performance measures, apart from the ones already investigated by Hobus and his colleagues.

In the study described in Chapter 2, particularly this second, developmental, aspect of illness scripts was addressed. On basis of the Hobus et al. (1989) study, in which subjects at different levels of expertise were asked to describe a prototypical patient with a particular disease, it can not definitely be concluded that less expert subjects actually lack the relevant knowledge; they may simply, for whatever reason, be less inclined to volunteer it in a recall task like the one employed by Hobus et al. (1989). To investigate whether this is the case, i.e., whether less experienced subjects are only less inclined to produce Enabling Conditions, rather than actually ignorant on this illness script component, every subject was presented with a scrambled case embedded in a pool of "noise" (i.e., irrelevant and case-inconsistent information), and asked to reconstruct a plausible case. It was hypothesized that if less experienced subjects lacked
knowledge of Enabling Conditions, they would be less inclined to include Enabling Conditions into their case reconstructions; however, if they did select patient contextual factors, this might be taken as support for the assumption that they do possess this knowledge, and consequently that the Hobus et al. (1989) results should be contributed to access or activation deficits, rather than to the sheer absence of knowledge. In addition, to increase the sensitivity of the study, a third level of expertise, i.e., fourth-year students, was included, while in the Hobus et al. (1989) study only beginning and expert physicians participated.

The results of the study failed to show any differences in case reconstruction between the three expertise levels. Subjects at all three levels selected on the average approximately 13 statements; about one-third of these were Enabling Conditions. Additional manipulations, like asking subjects to add new statements or to discard a fixed number of selected ones, also did not yield any differences between the two groups. Moreover, $\chi^2$-analyses, applied to individual case statements, showed that subjects at different levels of expertise agreed to a large extent as to the information that was included in the case reconstruction. Some circumstantial evidence for the script representation was also gathered: information typical for the "hidden" disease was selected frequently, while irrelevant or inconsistent information was not included in most subjects' reconstruction. In summary, the general conclusion from these results was that it is highly probable that there are no differences between expert physicians and advanced preclinical students, as far as their ability to recognize case descriptions in general, and patient contextual factors in particular, is concerned. Nevertheless, this conclusion should be stated somewhat cautious, because only one disease was presented and hence, a case-specificity effect can not be excluded.

In order to investigate the development of illness scripts and the proposed differential role of Enabling Conditions and Consequences into greater detail, we decided to replicate and extend the Hobus et al. (1989) study. This extended replication study was described in Chapter 3. Four different levels of expertise were included: fourth-year students, sixth-year students, interns (i.e., postgraduates in training as family physicians), and experienced family physicians. It was hypothesized that the number and/or proportion of Enabling Conditions produced in a free production task would increase with amount of experience, while simultaneously the number of biomedical elements mentioned would decrease. Furthermore, it was also predicted that experts' descriptions of a particular disease would be highly similar, regardless of the way they were probed, while especially subjects at intermediate levels of expertise would be sensitive to differences in instruction. Such sensitivity would suggest that intermediates are featured by a lack of integration of diagnostic knowledge into illness script structures. Finally, it was also investigated whether the number of patients seen with a specific disease, apart from expertise level in general, influenced the number and/or proportion of Enabling Conditions mentioned.

In this study, two different probes were used: subjects were asked to describe either a prototypical patient with, or the clinical picture of, each of 20 diseases, differing in frequency, seriousness, and organ system involved. The prediction was that expert physicians would always, i.e., regardless of the nature of the probe, be inclined to include some Enabling Conditions into their descriptions, while less expert subjects would only do so if probed to describe a prototypical patient, but not if asked to describe the clinical picture of a disease. Furthermore, it was hypothesized that especially the preclinical students would mention relatively much Fault-related, i.e., biomedical, information, while the importance of this illness script component would decline with increasing (practical) experience.
To a large extent, these expectations were borne out. In general, up to the level of interns, more experienced subjects were inclined to mention more Enabling Conditions, but beyond, no effect of expertise level was found. However, there was a large difference between the two experimental probes: If asked to describe a prototypical patient, an inverted-U shaped relationship between expertise level and number (and proportion) of Enabling Conditions mentioned was found, while a request to describe the clinical picture of a disease resulted in a monotonically increasing relationship between these variables. This result was interpreted as in support of the idea that with increasing expertise level, Enabling Conditions become a more integral part of illness scripts; accessible also if the subject is not directly probed into this direction. Thus, it is not knowledge availability, but knowledge accessibility that can account for the superior performances of expert subjects reported in previous studies. Also in accordance with the predictions, the number of Fault-related items decreased with increasing experience, for both types of probes alike, and also up to the level of interns. Furthermore, it appeared to be practical experience with diseases, rather than expertise level in general, that was most strongly related to number of Enabling Conditions mentioned (and hence, to illness script development).

The aim of the studies described in the Chapters 4 and 5 was to zero in on the script aspects, rather than the developmental features, of illness scripts. The design of these studies was of a more quantitative nature, focusing at a relatively fine-grained level of analysis. The study described in Chapter 4 investigated the influence of expertise level, case typicality, and illness script component on case reading times and case probability estimates. Short computerized case descriptions, in which a patient was described with Enabling Conditions and Consequences that were either atypical or prototypical for the disease he or she was suffering from, were presented to sixth-year students and experienced family physicians. Subjects were instructed to read the cases as fast as possible, to decide whether or not the patient described in the case suffered from the presumed disease (which was announced prior to the case presentation), and to provide a probability estimate, expressed as a percentage between 0 and 100, that the patient described in the case indeed had this disease. In addition, some NO-cases were included, in which a patient was described with a disease completely different from the one announced at the beginning of the case. Reading times of individual case statements were separately recorded, and average reading times of Enabling Conditions, Consequences, and complete cases were computed. It was hypothesized that experienced subjects would be faster than less experienced subjects, that prototypical case information would be processed faster than atypical case information, and that experienced physicians would be more sensitive to typicality of Enabling Conditions than advanced students. Furthermore, typicality would also contribute to the probability estimates, with the experts taking typicality of both Enabling Conditions and Consequences into account, while the non-experts would focus mainly on Consequences in determining a probability estimate. However, though the expected effects of expertise level and case typicality on case reading times were indeed found, the influence of illness script component turned out to be additive, rather than interactive: the contribution of typicality of Enabling Conditions was similar to that of Consequences, at both expertise levels. As far as the probability estimates are concerned, indeed the expected interaction between expertise level and illness script component materialized, with the expert subjects showing somewhat greater sensitivity for typicality of Enabling Conditions than sixth-year students, though these latter subjects generally were also influenced by this factor. The overall conclusion was that...
expert knowledge structures can be adequately described as illness scripts, and that the development of these scripts involves both Enabling Conditions and Consequences.

The study presented in Chapter 5 was in essence a replication, modification, and extension of previous script based research, performed by Yekovich and Walker (Walker & Yekovich, 1984, 1987; Yekovich & Walker, 1986) and Graesser and his colleagues (Graesser, 1981; Graesser et al., 1979; Graesser et al., 1980; Graesser & Nakamura, 1982; Smith & Graesser, 1981). In a recognition study, the influence of expertise level, actual textual presence, typicality, and illness script component on the accuracy of responses and the concomitant reaction times was investigated. In a study phase, fourth-year students, sixth-year students, and experienced family physicians were presented computerized case descriptions; the diagnosis was announced prior to the presentation, and all the subjects had to do was to read and understand the case. In a subsequent test phase, case statements were shown, and subjects had to decide, as quickly and accurately as possible, whether these statements had been literally presented in the associated case description, or not. These test statements differed in typicality, actual presence, and illness script component involved. Decision times and quality of the response (hit, false alarm, correct rejection, miss) were recorded. Concerning the accuracy of responding, it was hypothesized that memory discrimination would be better for atypical than for prototypical statements, and that this difference would increase with expertise level; expert subjects would be prone to falsely recognize prototypical, but unstated items. Concerning the speed of responding, it was predicted that experienced subjects would be faster than inexperienced subjects, that atypical statements would be responded to more quickly than prototypical statements, and that especially prototypical unstated items would show slow reaction times. An interaction with expertise level was also expected, with the expert physicians, though being generally fast, experiencing difficulty with prototypical unstated items. In addition, the reaction time data were compared to those of Yekovich and Walker (1986), in order to corroborate a proposed stage model of responding to scripted information.

The results of this study showed that subjects, and especially the more experienced ones, indeed made more errors of commission on prototypical unstated items than on atypical unstated items. It was also found that this effect could be attributed more to Enabling Conditions than to Consequences, at all three levels of expertise alike. As far as reaction times were concerned, all subjects showed script-consistent behavior, i.e., relatively long reaction times for unstated, prototypical information. This effect was also mainly due to the Enabling Conditions, at all three levels of expertise. Comparison of the present data with those of Yekovich and Walker (1986) showed similar results for hits and correct rejections, but whereas false alarms in this latter study apparently were a consequence of quick "recognitions", in the present study they seemed to be more the outcome of an extended deliberation, ultimately resulting in an inaccurate decision. A relevant additional result was that experienced physicians were strongly inclined to falsely recognize paraphrases, but not omitted information: Apparently, their memory for surface features of presented information quickly fades, but they are, in contrast with reports by other authors (e.g., Arkes & Harkness, 1980), not more inclined than students to infer unstated signs, symptoms, or complaints.

In summary, the results indicate that illness scripts share the important behavioral properties (e.g., processing speed, recognition memory performance) of scripts and schemas, as defined and investigated in the relevant literature. Furthermore, it can also be concluded that, with increasing disease-relevant
experience, subjects' illness scripts become more and more integrated and consolidated. The differential role of Enabling Conditions and Consequences in script development materialized mainly as differences in accessibility of scripted structures in a recall task as a consequence of probing; in an illness script recognition task, Enabling Conditions and Consequences seem to behave in a highly similar fashion for both students and experts.

**IMPLICATIONS FOR THE REPRESENTATION OF MEDICAL KNOWLEDGE AND MEDICAL PROBLEM SOLVING**

How can the representation of medical clinical knowledge best be described? In the previous chapters, four options were addressed, either theoretically or in experiments, or both. These options were: mental models of medical knowledge, prototypical patients, illness scripts, and representations of individual cases. First, it should be emphasized that physicians probably dispose of medical knowledge corresponding to all of these four types of representation: They are able to reason about diseases and cases in order to explain complaints, symptoms, and signs (cf. Feltovich et al., 1992; Hassebrock & Prietula, 1992); they have, for many diseases, a relatively clear-cut image of the prototypical patients associated with these diseases (cf. Chapter 3); their knowledge shows script-like properties (cf. Chapter 4 and Chapter 5), and they have memories for previous cases, even of many years' standing (cf. Hassebrock & Prietula, 1990; Van Rossum et al., 1990).

In this thesis, we have tried to make a case for an illness script representation of medical (clinical) knowledge. First, many previous studies have shown that elaborate reasoning is generally not a salient feature of experienced physicians' diagnostic performances, at least not in routine cases (e.g., Norman, Brooks, & Allen, 1989; Patel & Groen, 1986a; Schmidt & Boshuizen, 1993a). Experts' elaborate reasoning is associated with difficult diagnoses, and a relatively high proportion of errors (Norman, Brooks, & Allen, 1989). Second, though expert physicians in general have no difficulty in describing a prototypical patient with a particular disease, prototype representations are not adequately equipped to deal with the underlying time dimension of medical cases (however, for a different opinion, see Barsalou & Sewell, 1985). The finding that subjects tend to recall scrambled case information in accordance with the standard patient presentation order (Coughlin & Patel, 1987) does not point to a knowledge structure based on prototypicality; the features of a prototype do not have that kind of natural order. Prototypes also do not offer opportunities for representing the differential role of Enabling Conditions and Consequences. Third, as yet there is also no conclusive evidence that experienced physicians' diagnosis relies heavily on a "chorus of individual previous instances," stored in memory. It seems likely that repeated experience with similar cases results in knowledge structures from which the irrelevant aspects of the patients are filtered out; otherwise, physicians would be easily induced to activate a large number of highly unlikely diagnostic hypotheses. We venture the hypothesis that the use of memories for previous cases reported in the literature may sometimes have been an artefact of reinstating the original context in which the previous case was presented (cf. Godden & Baddeley, 1975). For example, physicians participating in an experiment might be reminded of a previous session, in which the prior case was presented, when again faced with the experimental equipment; seeing a corresponding case in daily practice probably may not lead to renewed activation. However, it is feasible that for rare diseases, for which (initially) no appropriate
illness script is present, physicians will use memories of concrete prior cases. Similarly, particularly highly salient Enabling Conditions may also result in a lasting memory representation of a case, even if the actual fault is relatively common; for example, Van Rossum et al. (1990) found that — the terminology is ours — "curious Enabling Conditions" indeed led to reactivation of the prior diagnostic hypothesis when a new case with the same background was presented, but not to the point of completely overruling alternative hypotheses. In summary, though patient memories may play an important role in non-routine situations and for some remarkable patients, it seems reasonable to assume that after having seen two or more cases with a specific disease, a more general structure, i.e., an illness script, will be formed.

Findings that causal models, prototypes, and instances all have drawbacks when it comes to representing expert medical knowledge, do not in itself constitute evidence for the psychological viability of illness scripts. On purely theoretical grounds, illness scripts have to be different from "classical" scripts, for they involve neither goals nor actors\(^{20}\). Nevertheless, it might be argued that the underlying time dimension provides a basis for describing diseases as illness scripts, with the illness script components (Enabling Conditions, Consequences, and Course/Therapy) playing the role of script scenes, and the individual context factors, signs, symptoms, and complaints playing the role of script actions and concepts. The results of the experiments described in the Chapters 4 and 5 have shown that the presumed illness scripts indeed show similar behavior as classical scripts, such as slow reading times for atypical information and a high proportion of false recognitions for paraphrased information. However, one important finding could not be replicated: the tendency of subjects to falsely recognize unstated, i.e., omitted, but typical, script information. In this respect, illness scripts appear to be truly different from classical scripts. In the discussion section of Chapter 5, some reasons were mentioned why illness scripts are different from classical scripts in this respect, the most important of these being that disease complaints and symptoms do not necessarily imply each other, while actions in "classical" scripts often do. How detrimental is this finding to the idea of illness scripts? If scripts are relatively broadly defined, as general event sequences, then there is no reason to deny illness scripts their script status. The alternative would be to assume a category structure of disease knowledge based on prototypicality; as we have already argued, prototype representations are unable to cover a number of relevant findings with respect to disease knowledge. However, in recent years, category theories are no longer restricted to perceptually based categories, but allow for causal underlying structures (cf. Medin & Ortony, 1989); as such, they may gradually be able to incorporate script-like properties (and vice-versa, cf. Barsalou & Sewell, 1985). In summary, based on the present results we think that illness scripts should be placed somewhere between prototypes and "classical" scripts, be it somewhat closer to the latter ones.

As far as the representation of medical knowledge is concerned, some more basic issues have to be dealt with here. The results of the studies presented in Chapters 2 and 3, in combination with previous data of Hobus et al. (1987), show that subjects at intermediate levels of medical expertise (sixth-year students, interns) do possess knowledge of Enabling Conditions, but are still unable to use

\(^{20}\) In our conceptualization of the illness script, neither the physician, nor the patient can be considered a "true" actor in the sense of Schank and Abelson (1977). Similarly, the "goal" of an illness script (i.e., correct diagnosis or treatment) can also not be put on a par with the socially defined goals of the "classical" scripts.
this knowledge in a diagnostic situation. It was concluded that in subjects at these levels of expertise, Enabling Conditions are not yet integrated into illness script structures. But what does this conclusion mean? Throughout this thesis we have implicitly — and in Chapter 5 even explicitly — adhered to a network based view of knowledge representation, in which scripts are embedded as specific sub-networks (cf. also Yekovich & Walker, 1987). In such a network, concepts are represented by nodes, and relations between concepts by links between nodes. These links enable a process of spreading activation, by which nodes may activate or inhibit each other. Two important features of such links are association strength and directionality. The connections between two nodes may be strong or weak, in either direction. Poor integration of Enabling Conditions into illness scripts can possibly be modeled as unbalanced links between nodes: the script header (diagnosis) may be strongly connected with several important Enabling Conditions and Consequences, while in the opposite direction, especially the Enabling Conditions are only weakly connected with the script header, for example, as a consequence of a "fan effect" (cf. Anderson, 1983). This situation may apply to subjects at intermediate levels of expertise: If they are presented with a diagnosis, they have no difficulty in mentioning the associated Enabling Conditions; however, the diagnosis itself will be activated mainly by the Consequences; the presence of Enabling Conditions may add little activation to that provided by complaints, signs and symptoms. Thus, if the diagnosis does not become part of working memory (activated long-term memory) as a result of activation of Consequences, Enabling Conditions may be of little help.

The gradual integration, with increasing expertise level, of Enabling Conditions into illness scripts might also be modeled by a production system architecture (cf. Anderson, 1983, 1993). In fact, the inherent asymmetry of production systems, provided by the directionality of production rules, may be particularly apt to model this situation. For example, with the diagnosis on the condition side of a production (i.e., explicitly mentioned), it may be easy to activate both relevant Enabling Conditions and Consequences. On the other hand, if the purpose of the production rule is to activate the diagnosis, inclusion of Enabling Conditions into an already complex condition consisting of Consequences may not increase the probability that this activation is achieved by firing of the production rule. This may be the case for intermediates; when they gradually become experts, the Enabling Conditions included in the condition side of the productions may tune the production rules to activation of more specific diagnoses than on basis of the Consequences alone.

However, presently we are especially concerned with a more psychological description of the diagnostic process. The core, and in our view also the most interesting stage of this process consists of the quick activation of a limited number of illness scripts as a consequence of the information available in the initial stage of the consultation or visit. This process of illness script activation proceeds automatically, for experts and non-experts alike. The critical difference between experienced and less experienced subjects concerns the appropriateness of the activated illness scripts; experts are more likely to activate the correct script than novices, because they are able to use information about Enabling Conditions, available in the diagnostic situation. This process of automatic script activation may be supplemented by some more controlled, additional search. For example, Weber et al. (1993) found evidence that physicians continue to generate diagnostic hypotheses until at least one serious alternative that can, be it with a low likelihood, account for the present symptoms and complaints, is found. If this diagnostic hypothesis is a standard alternative, given the Enabling Conditions and
the complaint, it may be automatically activated; if it is not, then it may have to be activated by a controlled search. Once the usually small, initial set of hypotheses has been formed, subjects are highly reluctant to generate new hypotheses further downstream the process of diagnostic problem solving (cf. Elstein et al., 1978; Wagenaar, 1987); new symptoms are used to test current hypotheses, not to generate new ones. People are often more inclined to neglect or "explain away" disconfirming symptoms than to reconsider their current hypothesis, at least if they can not think of a better one. Thus, it is not unusual for people to generate hypotheses they know do not account for all of the relevant facts. Up until recently, these kinds of behavior were interpreted as reasoning errors (e.g., a confirmation bias, cf. J. Evans, 1989); however, we think they can be more appropriately conceived of as efforts to reach a local maximum in a problem solving situation in which the optimum solution may be beyond reach. Subjects fail in a diagnostic problem solving situation because they are unable to activate an illness script that fits the present case; they stick to their current hypothesis as long as they can not think of a better one. With increasing experience, the situation gradually improves, because a richer base of illness scripts develops, and the individual scripts become more tuned toward the situations in which they are potentially applicable. After sufficient experience, problems that were initially hard may become routine, i.e., the appropriate illness script is activated almost immediately. Apart from an increased likelihood that the correct diagnosis will be found, a large repertory of well-tuned scripts may also have an additional advantage: it may free time and resources to deal with situations that are less familiar. A physician for whom a large proportion of cases is of a routine nature, may be able to spend more time on the non-routine patients, or on the non-routine aspects of otherwise routine cases. For example, an experienced physician may be able to recognize that the presenting complaint of a patient is only a pretext and use deliberate reasoning and problem solving to find out the "true" story behind it, while a less experienced physician has to invest all available resources to diagnose the "surface" problem. From this example it may also be inferred that the illness script theory certainly does not pretend that all knowledge used by experts in a medical diagnostic situation is embedded in full-fledged illness scripts, but only that physicians dispose of scripted disease knowledge for a large number of relatively frequently occurring diseases.

THE DEVELOPMENT OF MEDICAL KNOWLEDGE

In the previous chapters, as well as in the preceding sections, a great deal of information has already been provided about the development of medical knowledge. In general, we adhere to the model proposed by Schmidt and Boshuizen and Boshuizen and Schmidt in the early 1990s. The illness script is a central aspect of this model. Clinical knowledge is assumed to be represented as illness scripts; this applies to experts as well as to less experienced subjects. The major difference is that expert illness scripts are more finely tuned towards practical situations; this is, at least partly, a consequence of the integration of Enabling Conditions in the scripted structure, enabling the experts to quickly activate appropriate illness scripts in many situations. Experts' Consequences may also be better tuned than those of novices (cf. Hobus et al., 1988). As such, the development of illness scripts can be considered a form of weak restructuring of knowledge (cf. Vosniadou & Brewer, 1987). Experts represent more and/or different
relations between concepts than do novices, or, as seems to be the case for illness scripts, directionally more balanced relations: though advanced students are able to activate knowledge of Enabling Conditions given a diagnosis, the other way round is not yet passable by these subjects. A reason for this imbalance in the relations between concepts may be that Enabling Conditions are usually of a more general nature than Consequences. Thus, whereas a combination of a few Consequences often points unambiguously in a certain direction, for Enabling Conditions, this may much less often apply: They constrain the search space, but usually not to the point of excluding all but a few hypotheses. This is particularly the case for Enabling Conditions like age, sex, occupation (a few exceptions notwithstanding), environment, life style, and risk factors; these variables are probabilistically related to a large number of diseases. On the other hand, Enabling Conditions like previous diseases, use of medication and hereditary influences are often of a much more specific nature, and hence may play a role equivalent to that of Consequences. Neither the Hobus et al. studies, nor the present studies distinguished between different scopes of Enabling Conditions and Consequences; consequently, a suggestion for future research might be to replicate the Hobus et al. (1990) study to investigate the influence of specific versus general Enabling Conditions and Consequences.

Though no conclusive evidence has been gathered on the topic, there are some indications that the development of illness scripts and the concomitant integration of Enabling Conditions into these scripts follow a course relatively independent from the acquisition of biomedical knowledge, but contingent upon experience with actual patients. For example, we found (Chapter 3) that subjects' inclination to mention Enabling Conditions is associated to the number of patients they have seen with a specific disease, but not to medical experience in general. Weber et al. (1993) also report that medical experience has no independent effect on hypothesis generation beyond the "having seen of similar cases before." Furthermore, the contribution of biomedical knowledge in descriptions of patients or clinical pictures asymptotes to about zero long before the full integration of Enabling Conditions takes place (cf. Chapter 3); this also suggests a relatively independent development of illness scripts, at least as far as the final stages of this development are concerned. In fact, it may be theoretically possible that illness scripts develop without knowledge of the underlying biomedical model, just like subjects in the study of Ahn et al. (1992) were able to learn the details of a "pot-latch" ceremony, without knowing the reasons behind the actions that had to be performed. Similarly, physicians in non-western cultures, physicians in the middle ages, and physicians working according to principles of alternative medicine, may have illness scripts that are either not tied at all to an underlying biomedical model, or based on an alternative, maybe even incompatible, model. Indeed, even regular, late 20th century physicians in Western Europe or America may have idiosyncratic illness scripts (cf. Schmidt, Norman, & Boshuizen, 1990), that may not be properly embedded in regular biomedical knowledge. Nevertheless, it is likely that most students do employ biomedical knowledge in developing illness scripts, if only to be able to explain why particular Enabling Conditions and Consequences are (inter-)related, in addition to knowing that they are part of the same disease (i.e., an aspect of their clinical knowledge). If, in a diagnostic situation, an appropriate illness script can be quickly activated, no biomedical knowledge is necessary to establish a diagnosis; if this process fails, subjects will try to find a diagnosis by explaining, initially on basis of clinical knowledge, but if this is also lacking, on basis of biomedical knowledge, the possible common origin of the complaints, signs and symptoms.
In this thesis we have focused mainly on intra-individual cognitive aspects, and implicitly we have considered diseases as fixed, objective entities. However, we have to address the issue of the interaction between the "external world" and the human cognitive system here. A particular prominent aspect of this interaction concerns the nature of signs and symptoms. Throughout the previous chapters, it has been implicitly assumed that signs and symptoms are objectively perceptible and unchanging entities. However, there is evidence (cf. Bransford, Franks, Vye, & Sherwood, 1989; Lesgold, 1984) that experts may perceive quite different features than novices. Thus, the problem may not be that it is difficult for inexperienced subjects to learn or remember rules like: "diagnosis X is applicable if the patient is moderately depressed" – the problem is that perceiving a feature like "moderately depressed" is the result of a complex pattern recognition process in an expert, while the rule does not provide clues for a novice how to recognize this symptom. In addition, some features may only become perceptible if the illness script is already activated. Lesgold (1984) points to such a situation in which the triggering of the atelectasis (a pulmonary affliction) script on basis of some features of a chest X-ray, leads to the perception of the other features of this condition as well. Though the importance of this observation may be particularly high for complex visual domains like radiology and dermatology, they may sometimes also play a prominent role general medicine. A related phenomenon is the interaction between signs or symptoms as a consequence of the condition of the patient. For example, heart sounds of highly overweight patients may differ considerably from those of subjects with average weight. An important aspect of diagnostic expertise development may be a growing sensitivity for this interdependence between features. In summary, it should be emphasized that to consider signs and symptoms as objectively perceptible, unchangeable, and for subjects at all expertise levels alike, is in fact a simplification.

**IMPLICATIONS FOR MEDICAL EDUCATION**

Our experimental results do not provide direct suggestions for medical educational practice. However, especially in light of previous research, some tentative recommendations can be formulated.

First, how can education foster the development of illness scripts? In the study described in Chapter 3, we found that the development of illness scripts is associated with subjects' amount of experience with actual patients who suffer from the diseases in question. Thus, it is important to provide students with opportunities to see patients. In this respect, it is probably better for students to see many patients for a short time, than to delve deeply into a few cases. To foster illness script development, the cases presented to students should be as representative as possible for the specific illness scripts (cf. Bordage, 1984; Murphy & Wright, 1984; Van Rossum et al., 1990). Cases with much "noise" should be avoided, especially patients with an impressive, but irrelevant context (Van Rossum et al., 1990), because this may cause students to develop a wrong image of the Enabling Conditions. An additional reason why the use of actual patients should be recommended is that since patient problems are the future context in which medical knowledge is used, why not also teach it in this context? (cf. Godden & Baddeley, 1975; Norman, 1988). This notion is also implicitly present in the recommendation of Elstein et al. (1978), who advocate to provide students with repeated exercises in knowledge application for a similar range of problems as they will be experience in the future.
However, as for practical reasons the opportunities for students to see and examine patients probably will remain limited, it may make sense to encourage them to gather as much knowledge, especially clinical knowledge, about diseases and patients as possible, not only by studying textbooks, but also by watching videotapes, attending patient demonstrations, etceteras. Though an extensive theoretical clinical knowledge base is not sufficient for developing full-fledged illness scripts, it may be very helpful, because students may develop relatively well-defined expectations concerning what they will see when faced with concrete patients. Thus, it is likely that they can profit to a larger extent from their actual clinical experiences, if the theoretical background of their (future) illness scripts is already present.

On the other hand, it does not seem to make much sense to teach students explicitly the strategy to use Enabling Conditions in trying to diagnose a case; the integration of Enabling Conditions into illness scripts is something that takes much time and experience. The use of Enabling Conditions in the initial phase of a diagnostic situation is not under voluntary control. Furthermore, especially advanced students already possess this knowledge, and may know that it is important, even though they are unable to use it.

Similarly, it should also not be recommended to teach students hypothesis generation, as, for example, Benbassat & Shiffman (1976), advocate. Prolific generation of hypotheses is not characteristic of expert behavior; addition of a new hypothesis (or substitution of a rejected hypothesis by a new one) is only useful if this more recently activated hypothesis is qualitatively better than the best of the old ones. What we do recommend is that students be stimulated to generate reasons why their diagnostic hypotheses may be incorrect in a certain case (cf. Wageman, 1987). As already mentioned, people tend to stick to their current hypothesis, even if they know or may expect that it can not be correct, as it flies in the face of the empirical evidence. However, to ignore the evidence may be more comfortable than to end up with no hypothesis at all. In a medical diagnostic situation, such behavior may be manifested by the physician implicitly denying part of the information provided by the patient (e.g., "You really do not have a hoarse cough?") not necessarily because he suspects the patient to be insincere, but because the information is inconsistent with the, in his eyes, currently most likely diagnostic hypothesis. Students should be made aware of this kind of behavior, which may impair the quality of their diagnostic performances, apart from being unfavorable to the quality of the doctor-patient relationship.

What options are available if the student or physician gets stuck in such a situation? The illness script theory does not provide any clues as to how to find a way out of such a situation. Though it may intuitively be assumed that a search for additional information may lead to a solution, according to Elstein et al. (1978), insufficient information is usually not a critical diagnostic bottleneck. Stubbornly continuing to gather data, however remote they may be from the actual problem of the patient, is not a good strategy: The probability that the appropriate illness script will be activated is small, and in practical situations, it is apt to lead to costly and unnecessary diagnostic procedures. Rather, a physician should rely on good amnestic interview skills, because these are the most valuable and least expensive tools he has at his disposal; a fact already acknowledged by Wilkins (1970), who remarks that a good history should be given 80% of the weight in a diagnosis. If the correctly performed anamnestic interview does not provide clues as to what might be the most likely illness script, neither students nor experienced physicians should be reluctant to consult medical handbooks or colleagues in order to arrive at the appropriate diagnosis.
Summary and general discussion

A topic that has received considerable attention in the literature on medical expertise and medical education is the use of Bayesian statistics and medical decision theory in diagnosis and treatment (e.g., Biela, 1986; Camerer & Johnson, 1991; Christensen-Szalanski & Bushyhead, 1981; Hamm, 1988; Kuipers, Mascovich, & Kassirer, 1988; Schwartz & Griffin, 1986). From the present point of view, two things can be said about this. First, it surely is important that the episodic knowledge base of a physician — i.e., the pool of instances his illness scripts are based upon — is representative for the population frequencies (Weber et al., 1993). Experience with the "wrong" type of cases for a disease, as well as with the "wrong" diseases for a population, may result in a systematic bias in illness scripts. However, quite another question is whether it is effective to explicitly teach students and physicians to take Bayesian principles into consideration in arriving at a diagnosis. It should be clear from the previous discussion that the initial activation of illness scripts in a diagnostic setting is not a process under voluntary control. Thus, explicitly applying Bayesian principles in the early stages of diagnostic problem solving, is not possible. Nevertheless, there is evidence that "Bayesian-like" principles are implicitly incorporated into the activation of illness scripts. For example, Christensen-Szalanski and Bushyhead (1981) found that, though it is difficult for physicians to incorporate explicitly presented base-rate information into their diagnostic judgements, they do use apriori probabilities obtained from experience. This finding suggests that it indeed might be more sensible to provide students with cases in frequencies that are as representative as possible for the future patient population they will encounter, than to let them solve problems in which base-rate probabilities have to be explicitly used. On the other hand, it may be useful to teach some general principles in this respect, like the difference between representative features and diagnostic features of diseases (Klayman & Brown, 1993), or, as Hodgkin (1984) recommends, awareness of the fact that rare symptoms of common diseases occur more frequently than common symptoms of rare diseases. However, this kind of knowledge is probably more relevant in deciding between two competitive, simultaneously activated illness scripts, than in the process of illness script activation itself. In general, we think that the illness script theory and medical decision theory complement rather than contradict each other (cf. also Patel & Coughlin, 1985): The former theory is a descriptive theory about the way clinical diagnostic knowledge is represented and activated, the latter a normative theory about making decisions in situations that are characterized by a degree of uncertainty.

SUGGESTIONS FOR FURTHER RESEARCH

In the discussion sections of the Chapters 2 through 5, a number of recommendations and suggestions for future research have already been proposed. For example, the script recognition study in Chapter 5 might be extended to longer retention intervals and to recall measures, particularly in order to solve the discrepancy between the results of this study and previous research, which reported that subjects do infer unstated, but script or schema typical information. Furthermore, an experimental distinction between general and specific Enabling Conditions and Consequences might yield interesting results. For example, it may be found that general Enabling Conditions, or background information in the sense of Weber et al. (1993) (like age, sex, occupation, and environment), may play a different role compared to more specific information, like medical history, current medication, travelling to tropical countries, and hereditary factors.
In Chapter 5, it was hypothesized that the failure to find a false recognition effect of omitted signs and symptoms should be attributed to the absence, at least between separate symptoms, of a causal structure in illness scripts. Therefore, it would be interesting to compare illness scripts with classical scripts that also lack a causal structure, like a circus ring performance program (cf. Bower et al., 1979). In a circus performance, the appearance of lions, for example, does not imply the appearance of clowns, and vice versa; thus, in contrast to the restaurant situation, memories of a text about a circus program will contain no causal cues about the actual presence of the different acts. Another example of such scripts can be found in a study of Ahn et al. (1992), in which subjects had to learn a difficult script (the "potlatch ceremony") for which no causal or goal structure could be invoked. If these types of scripts would show the same memory phenomena as the illness scripts in the study in Chapter 5, particularly as far as the false recognition data of omissions are concerned, then this would provide additional support for the illness script theory. It would also indicate that illness scripts can in principle operate without knowledge of an underlying model, because the potlatch ceremony — and probably the circus performance, too — lack such a model. This would imply that illness scripts, at least in principle, can be learned almost without any underlying knowledge of the medical basic sciences.

Another question concerns the initial development of illness scripts. The results of the studies described in the present thesis show that sixth-year students already possess relatively well-developed illness scripts, while even fourth-year students may have script-like knowledge structures for some cases. It would be interesting to investigate whether preclinical students are inclined to construct illness scripts in an effortful manner, for example by inserting self-generated explanations into their beginning scripts. To a certain extent, this would be equivalent to the behavior of subjects in Bartlett’s (1932/1954) experiments with the mysterious, difficult to comprehend "War of the Ghosts" story. By generating explanations for the different actions in this story, subjects were able to remember it, though with a considerable distortion towards a more ordinary scripted story. In a similar vein, it is possible that preclinical students try to comprehend a difficult case, with many apparently disparate signs and symptoms, by generating explanations for the coherence of the features. Parts of these explanations may subsequently become part of the illness script structure, and hence intrude during recall, or lead to false recognitions in a recognition task. In general, we think that especially in the early stages of illness script development, biomedical explanations of the reasons why certain symptoms or patient characteristics are related to each other play an important role in interconnecting the elements of illness scripts; in later stages, firm associations between the constituent parts of these scripts will establish, and consequently the role of the underlying biomedical explanations decreases. If a physician is in doubt whether a certain illness script is applicable, he will tend to use clinical, rather than biomedical, reasoning to solve the problem.

Still another issue, already addressed in Chapter 3 but in need of more research, concerns the question of the inclusion, and possible role, of a fourth illness script component, i.e., Course & Treatment. Though this component may be of little interest for the initial stages of the diagnostic process, it may have important implications for later stages. The results of Chapter 3 suggest that knowledge of this component is acquired in a relatively late stage of expertise development. Similar studies like the ones described in Chapter 4 and 5 might be used to investigate whether knowledge of the expected course of a disease and the treatment eventually become part of the scripted structure.
However, though these studies could reveal interesting aspects of illness scripts, we think that only more fine-grained models, in which the basic elements of cognition are explored, can ultimately indicate why experienced physicians are so much more able to generate accurate hypotheses on basis of scarce information, than less expert subjects. Though this is of course not the only feature of medical expertise (e.g., confirming a proposed diagnosis, choosing and applying a treatment, advising a patient, and so on, are also important aspects), we think it is an extremely important one. Expert diagnostic knowledge is organized in a way that enables experienced physicians, faced with a patient, to quickly recognize in many cases a small number of possible diagnostic hypotheses, even to the point of already "knowing" at that moment which disease the patient suffers from. At the behavioral level, especially if qualitative measures are employed, it will be difficult to find cues to explain this capability. As Hofstadter (1985) argues, the most interesting aspects of cognition take place at the level of 100 msec or less - "the time it takes to recognize your mother." For most people, it will make simply no sense to ask how or why they recognize their mother — then why ask experienced physicians how or why they recognize that a patient has a disease the physician may have seen almost as often as his or her mother? Real progress in expertise research can only be expected from studies that go beyond — indeed, way beyond — the purely descriptive level. In this thesis, we have tried to make a small step towards a more explanatory model, to disclose something of what is generally called the "clinical look."
SAMENVATTING (SUMMARY IN DUTCH)

De ontwikkeling en functie van ziektescripts:
Een onderzoek naar de structuur van medisch-diagnostische kennis

In dit proefschrift wordt verslag gedaan van een viertal onderzoeken die uitgevoerd zijn met als doel een nader licht te werpen op de representatie en ontwikkeling van medisch-diagnostische kennis. Tezamen kunnen deze onderzoeken geplaatst worden in het kader van modellering en theorievorming met betrekking tot de ontwikkeling van expertise. In Hoofdstuk 1 wordt allereerst een historisch overzicht gegeven van het onderzoek in de psychologie dat aanzet heeft gegeven tot de huidige belangstelling voor de aard en ontwikkeling van expertise. Beargumenteerd wordt dat de wortels van het huidige expertise-onderzoek gezocht moeten worden in met name het onderzoek naar probleemoplossen en naar de wijze waarop kennis geregestoneerd kan worden. Een aantal theorieën met betrekking tot beide onderwerpen wordt kort besproken. Een belangrijke plaats wordt ingeruimd voor het onderscheid tussen modellen die ervan uitgaan dat kennis al redenerend en elaborerend wordt toegepast, en modellen die de nadruk leggen op kennis als kant-en-klare pakketjes, die min of meer automatisch als geheel toegankelijk zijn tijdens het probleemoplossen. Uit eerder onderzoek is gebleken dat het laatste het geval is met de diagnostische kennis van ervaren artsen, zeker in routinesituaties: deze kan geregestoneerd worden in de vorm van ziektescripts. Alhoewel Feltoovich en Barrows (1984) de eersten waren die het ziektescript als kennistructuur nader hebben uitgewerkt, en het begrip een relatief onafhankelijke status hebben gegeven, wordt in het huidige proefschrift gekozen voor een andere invalshoek, waarbij het ziektescript veel sterker gekoppeld wordt aan het "klassieke" script, zoals dat gedefinieerd en uitgewerkt wordt door Schank en Abelson (1977). In deze zienswijze worden ziektescripts in eerste instantie geclassificeerd door binnenkomende diagnostische informatie, en vervolgens geïnstantieerd met behulp van de verdere gegevens van de betreffende patiënt of casus.


Deze, en aanverwante, bevindingen hebben geleid tot een theorie over de ontwikkeling van klinisch-medische kennis (Schmidt, Norman, & Boshuizen, 1990).
Deze theorie stelt dat in de preklinische fase van de geneeskunde-studie studenten voornamelijk bezig zijn met het vergaren van ziektekennis en het opbouwen van een gigantisch netwerk van, met name biomedische, kennis. Gedurende de coschappen, wanneer ze in de kliniek rondlopen, wordt deze boekenkennis toegepast en afgestemd op gebruik in praktische situaties. Geleidelijk aan ontstaan volwaardige ziektescripts, zij het dat dit proces geruime tijd in beslag neemt en zeker nog niet voltooid is op het moment dat het basisarts-diploma wordt behaald. De bevinding dat net afgestudeerde basisartsen veel minder goed dan experts in staat zijn om gebruik te maken van gegevens uit de context van de patiënt (Enabling Conditions), wijst er op dat de ontwikkeling van deze component van het ziektescript pas in een relatief laat stadium wordt voltooid.

Van de vier experimenten waarvan in de hoofdstukken 2 tot en met 5 van dit proefschrift verslag wordt gedaan, zijn er twee rechtstreeks gericht op deze theorie van de ontwikkeling van medische kennis: er wordt geprobeerd aanvullende ondersteuning te vinden voor de ziektescript-theorie en de relatief late ontwikkeling van kennis met betrekking tot Enabling Conditions. De andere twee experimenten hebben als hoofdthema de vraag in hoeverre ziektescripts vergelijkbaar zijn met de Schank en Abelson (1977) scripts. In de jaren volgend op de introductie van het script-begrip zijn een groot aantal onderzoeken uitgevoerd waarin de bruikbaarheid en theoretische levensvatbaarheid van het begrip getoetst werd. Indien experimentele evidentie kan worden gevonden voor de veronderstelling dat ziektescripts tot dezelfde geheugenprestaties en verwerkingspatronen aanleiding geven als klassieke scripts, dan vormt dit eveneens ondersteuning voor het nut van het (theoretische) concept ziektescript.

In Hoofdstuk 2 worden de resultaten beschreven van een scrambled case reconstruction study. Achtergrond van dit onderzoek vormde de vraag in hoeverre ziektescripts vergelijkbaar zijn met de Schank en Abelson (1977) scripts. De bedoeling van deze taak was dat de proefpersonen zelf een casus samenstelden (reconstrueerden) op basis van een verzameling losse brokstukken casus-informatie. Deze informatie, die bestond uit korte zinnetjes afgedrukt op in totaal 48 kaartjes, had betrekking op zowel de achtergrond en context van de patiënt, als op diens huidige klachten en symptomen. Sommige kaartjes bevatten "echte" casusinformatie, andere slechts "ruis", en weer andere zelfs informatie die strijdig was met de bedoelde ziekte. De scrambled case was zodanig samengesteld dat het herkennen van de verborgen ziekte (i.e., de ziekte van Pfeiffer) op zichzelf niet moeilijk bleek, ook niet voor de vierdejaars-studenten; de vraag was dan ook niet of ervaren artsen beter in staat zijn een ziekte te herkennen temidden van een hoeveelheid ruis, maar of zij meer dan studenten geneigd zouden zijn Enabling Conditions op te nemen in hun casus-reconstructie. Als dit laatste het geval is, dan wijst dit erop dat het studenten kennelijk ontbreekt aan kennis van Enabling Conditions: ook wanneer informatie met betrekking tot deze ziektescript-component expliciet voorhanden is, wordt deze door hen niet herkend en geselecteerd.
De resultaten lieten echter geen enkel significant verschil zien tussen de drie expertiseniveaus. Niet alleen selecteerden proefpersonen op de verschillende expertiseniveaus gemiddeld een ongeveer gelijk aantal informatie-elementen voor de casus-reconstructie, ook op de aard van de geselecteerde informatie bleek geneeskundige ervaring geen enkele invloed te hebben. Ook uit enkele extra manipulaties, zoals het laten wegnemen van niet-cruciale informatie, bleek geen enkel effect van expertiseniveau. De algemene conclusie luidde dan ook dat nergens uit blijkt dat bij het reconstrueren van een casus experts meer geneigd zijn Enabling Conditions te selecteren dan niet-experts. Het lijkt er dus op dat deze kennis in principe aanwezig is, ook bij gevorderde studenten.

Deze conclusie wordt nader getoetst in het onderzoek waarover in Hoofdstuk 3 wordt gerapporteerd. Dit onderzoek is in feite een replicatie en uitbreiding van het experiment van Hobus et al. (1989). De replicatie hield in dat in het huidige onderzoek, net als in dat van Hobus et al., aan proefpersonen gevraagd werd om, voor twintig verschillende ziekten, een bijbehorende prototypische patiënt te beschrijven. De uitbreiding was tweeledig. Ten eerste werden twee extra expertiseniveaus toegevoegd, te weten vierdejaars-studenten geneeskunde en huisartsen in opleiding, naast zesdejaars-studenten en ervaren huisartsen. Ten tweede werd weliswaar aan de helft van de proefpersonen gevraagd om een prototypische patiënt te beschrijven met elk van de ziekten, maar de andere helft kreeg de opdracht om het ziektebeeld behorend bij de aandoeningen te beschrijven. Deze laatste conditie was toegevoegd omdat de mogelijkheid bestond dat het expliciete verzoek om een prototypische patiënt te beschrijven, proefpersonen ertoe kan hebben aangezet om Enabling Conditions, die eigenlijk (nog) geen deel uitmaaken van hun ziektescript-structuren, toch in hun beschrijvingen op te nemen. De achterliggende gedachte was dat de instructie om het ziektebeeld behorend bij een bepaalde ziekte te beschrijven, in plaats van een prototypische patiënt, proefpersonen in hun verhalen veel minder zou sturen in de richting van Enabling Conditions; immers, de term ziektebeeld werd geacht sterker te verwijzen naar klachten en symptomen, dan naar factoren uit de context van de patiënt.

Op grond van de vroegere resultaten werd verwacht dat ervaren artsen in het algemeen meer Enabling Conditions (zowel absoluut aantal als proportioneel) zouden noemen dan minder ervaren proefpersonen. Ook werd een interactie voorspeld tussen expertiseniveau en type instructie: ervaren artsen zouden in beide condities ongeveer dezelfde beschrijvingen geven, inclusief de relevante Enabling Conditions, terwijl met name proefpersonen op de tussenliggende expertiseniveaus (zesdejaars-studenten, huisartsen in opleiding) een zekere gevoeligheid voor het type instructie zouden vertonen, in die zin dat ze (relatief) meer Enabling Conditions zouden noemen indien hen expliciet gevraagd werd een prototypische patiënt te beschrijven, dan wanneer ze een ziektebeeld zouden moeten beschrijven. Voorts werd verwacht dat het aandeel van biomedische informatie in de beschrijvingen een monotoon negatief verband zou vertonen met expertiseniveau.

In grote lijnen werden deze hypotheses door de feiten gestaafd. Zeker tot het niveau van huisartsen in opleiding nam het aantal en de proportie Enabling Conditions in de beschrijvingen toe naarmate de proefpersonen meer ervaring hadden. Daarna is er geen toename meer; de data wijzen er zelfs op dat ervaren huisartsen, met name proportioneel, wat minder Enabling Conditions noemen dan huisartsen in opleiding. Daarnaast bleek er een groot verschil te bestaan tussen de twee condities: indien de proefpersonen een prototypische patiënt moesten beschrijven, werd een omgekeerd U-vormig verband gevonden tussen expertise-
niveau en aantal en proportie Enabling Conditions. Werd echter gevraagd naar het ziektebeeld, dan bleek er veeler sprake van een monotone toename van aantal en proportie Enabling Conditions in de beschrijvingen naarmate de proefpersonen meer ervaring hadden. Deze resultaten ondersteunen het idee dat met toenemend expertiseniveau, Enabling Conditions steeds meer integraal onderdeel gaan uitmaken van ziektescripts, in die zin dat zij ook toegankelijk zijn indien zij niet rechtstreeks, dat wil zeggen niet via de instructie om een prototypische patiënt te beschrijven, aangesproken worden.

Eveneens in overeenstemming met de voorspellingen bleek het aantal biomedische informatie-eenheden in de beschrijvingen zowel relatief als absoluut af te nemen naarmate het expertiseniveau toenam. Dit effect trad op bij beide instructies. Tenslotte werd ook gevonden dat niet zozeer geneeskundige ervaring in het algemeen, maar met name het daadwerkelijk gezien hebben van patiënten met de in het onderzoek gebruikte ziekten, positief gerelateerd was aan het aantal Enabling Conditions in de beschrijvingen: er bleken correlaties in de orde van .35 tot .50 te bestaan tussen het aantal patiënten dat de proefpersonen gemiddeld aangaven gezien te hebben met een bepaalde ziekte (gemeten met behulp van een rating-scale), en het gemiddelde aantal Enabling Conditions dat werd genoemd. Voor de overige ziektescript-componenten bleek er vrijwel geen verband te bestaan tussen aantal patiëntcontacten en hoeveelheid informatie in de beschrijvingen. Dus, het lijkt erop dat met name praktische ervaring met patiënten de ontwikkeling van ziektescripts, zeker wat betreft de integratie van Enabling Conditions in de kennisstructuren, bevordert.

De experimenten die in de hoofdstukken 4 en 5 worden beschreven hadden met name tot doel het onderzoeken van de "script"-aspecten van ziektescripts. Vergeleken met de onderzoeken uit hoofdstuk 2 en 3 waren deze studies relatief kwantitatief en fijnkorrelig van aard. In Hoofdstuk 4 wordt een experiment beschreven waarin de invloed van expertiseniveau, typiciteit van casuïstiek en ziektescript component (Enabling Conditions en Consequences) op de snelheid waarop casusinformatie gelezen, c.q. verwerkt wordt, nader werd onderzocht. Tevens werd nagegaan in hoeverre deze factoren van invloed zijn op de subjectieve waarschijnlijkheid (kansschatting) dat de in een casus beschreven patiënt daadwerkelijk lijdt aan de kwaal waarvan het ziektescript is geactiveerd. In dit onderzoek kregen zesdejaars-studenten geneeskunde en ervaren huisartsen een aantal korte, gecomputeriseerde casus te lezen, waarin steeds een patiënt werd beschreven met Enabling Conditions en Consequences die hetzij prototypisch, hetzij atypisch waren voor de ziekte waaraan deze patiënt verondersteld werd te lijden. De proefpersonen kregen de opdracht de casus zo snel en zo grondig mogelijk te lezen, om aan het einde te beslissen of de patiënt inderdaad aan deze ziekte, die was aangekondigd aan het begin van de casus, leed. Deze beslissing moest zowel door middel van een dichotome ja/nee keuze, als door middel van een waarschijnlijkheidsschatting, uitgedrukt als percentage, worden genomen.

De leestijden van afzonderlijke caseuselementen werden geregistreerd, en de gemiddelde leestijden van Enabling Conditions, Consequences, en gehele casus werden berekend. Voorspeld werd dat ervaren huisartsen de casusinformatie sneller zouden verwerken dan zesdejaars-studenten, dat prototypische informatie sneller verwerkt zou worden dan atypische, en mogelijk dat de meer ervaren proefpersonen wat gevoeliger zouden zijn voor typiciteit van Enabling Conditions dan de studenten. Ook werd verwacht dat typiciteit van de aangeboden informatie van invloed zou zijn op de kansschattingen, waarbij ervaren artsen met typiciteit van zowel Enabling Conditions als Consequences rekening zouden houden, ter-
wijl de zesdejaars-studenten hun waarschijnlijkheidsoordelen voornamelijk zouden laten bepalen door typiciteit van Consequences.

Inderdaad bleek dat ervaren proefpersonen casusinformatie sneller verwerken dan studenten, en dat prototypische informatie sneller wordt gelezen dan atypische. Echter, de invloed van ziektescript-component was additief van aard, en vertoonde geen interactie met expertiseniveau: de bijdragen van Enabling Conditions en Consequences aan de leestijden waren ongeveer gelijk voor beide expertiseniveaus. Voor de waarschijnlijkheidsschattingen werd wél een interactie gevonden tussen expertiseniveau en ziektescript-component: ervaren huisartsen blijken wat meer rekening te houden met typiciteit van Enabling Conditions bij het schatten van de kans dat een patiënt lijdt aan een bepaalde ziekte dan zesdejaars-studenten, alhoewel deze laatsten zeker niet helemaal ongevoelig zijn voor typiciteit van deze ziektescript-component. Deze resultaten ondersteunen de gedachte dat de kennisstructuren van experts (ziekte)script-achtige eigenschappen vertonen, en tevens dat bij de ontwikkeling van deze structuren beide ziektescript-componenten, t. w. Enabling Conditions en Consequences, zijn betrokken.

Het in Hoofdstuk 5 gepresenteerde onderzoek tenslotte is in wezen een replicatie, aanpassing, en uitbreiding van eerder script recognition onderzoek verricht door Yekovich en Walker (Walker & Yekovich, 1984, 1987; Yekovich & Walker, 1986) en Graesser en zijn collega's (Graesser, 1981; Graesser, Gordon, & Sawyer, 1979; Graesser, Woll, Kowalski, & Smith, 1980; Graesser & Nakamura, 1982; Smith & Graesser, 1981). In een herkenningstaak werd de invloed van expertiseniveau, het al of niet letterlijk gepresenteerd zijn van casusinformatie, typiciteit, en ziektescript-component op de nauwkeurigheid van de responsen en de bijbehorende reactietijden onderzocht. De experimentele taak bestond uit twee onderdelen. In een leerfase kregen vierdejaars- en zesdejaars-studenten geneeskunde en ervaren huisartsen gecompiteriseerde casus aangeboden. De diagnose verbonden aan de casus werd van tevoren bekend gemaakt; de enige taak die de proefpersonen moesten uitvoeren was het lezen en begrijpen van de casus. De presentatieduur van alle casuselementen, die de vorm hadden van korte netjes, was van tevoren vastgesteld en voor alle proefpersonen gelijk. Na een korte tussentaak werd de testfase gestart. Gedurende deze testfase kregen de proefpersonen casuselementen aangeboden, waarvan zij moesten vaststellen, zo snel en nauwkeurig mogelijk, of deze letterlijk op het scherm te zien waren geweest gedurende de presentatie van de overeenkomstige casus in de leerfase, of niet. Typiciteit, feitelijke presentatie, en ziektescript-component van de test-elementen werden experimenteel gemanipuleerd. De reactietijd voor elk element werd geregistreerd, evenals de kwaliteit van de response: hit, false alarm, correct rejection, of miss. In navolging van het voornoemde script-onderzoek werd voorspeld dat de proefpersonen atypische casus-elementen beter zouden kunnen herkennen dan prototypische, en dat dit verschil groter zou zijn voor de ervaren artsen dan voor de studenten. Eersten artsen zouden met name geneigd zijn prototypische, maar niet letterlijk gepresenteerde casusinformatie, ten onrechte te herkennen als wél letterlijk aangeboden. Wat betreft de reactietijden werd verwacht dat deze korter zouden worden met toenemend expertiseniveau, dat op atypische casuselementen sneller gereageerd zou worden dan op prototypische, en dat met name prototypische, niet letterlijk gepresenteerde casusinformatie aanleiding zou geven tot lange reactietijden. Daarnaast werd ook een interactie tussen expertiseniveau en typiciteit verwacht, in die zin dat ervaren artsen in het algemeen snellere responsen zouden geven dan gevorderde studenten en onervaren artsen, maar relatief meer moeite zouden hebben met prototypische, niet letterlijk aangeboden informatie. Tenslotte werden de huidige resultaten vergeleken met die van Yekovich.
en Walker (1986), om na te gaan of het door hen gepresenteerde stadia-model van het verwerken van script informatie, ook van toepassing zou zijn op ziektescripts.

Uit de resultaten bleek inderdaad dat prototypische, niet letterlijk aangeboden casuselementen vaak ten onrechte herkend werden, veel vaker dan atypische, niet letterlijk aangeboden elementen. Dit effect was sterker voor ervaren artsen dan voor studenten. Voor alle drie de expertiseniveaus gold tevens dat dit resultaat met name het gevolg was van de reacties op Enabling Conditions, en minder van die op Consequences. Daarnaast bleken proefpersonen op alle drie de expertiseniveaus relatief lange reactietijden te vertonen voor niet letterlijk gepresenteerde, prototypische informatie. Dit effect was ook weer voornamelijk toe te schrijven aan de Enabling Conditions. De vergelijking van de huidige resultaten met die van Yekovich en Walker (1986) liet een gemengd beeld zien. De reactietijden voor hits en correct rejections voor ziektescript-informatie vertoonden een patroon overeenkomstig met dat van dit type informatie in klassieke scripts. Echter, waar false alarms in het eerdere onderzoek voornamelijk een gevolg leken van snelle, onterechte, "herkenning", leken ze in het huidige onderzoek eerder het resultaat van een weloverwogen, doch verkeerd uitpakkinge beslissing.

Een opvallende bevinding van het onderzoek in Hoofdstuk 5 was dat, in tegenstelling tot eerder onderzoek (bv. Arkes & Harkness, 1980), de ervaren artsen in de huidige studie niet geneigd waren prototypische informatie die opzette- lijk uit de casus was weggelaten, maar in de testfase wel werd aangeboden, ten onrechte te herkennen. Dit in tegenstelling te geparafraseerde informatie, dat wil zeggen testelementen die qua inhoud overeenkwamen met oorspronkelijke casuselementen, maar niet letterlijk hetzelfde waren: deze laatste werden door ervaren artsen wél relatief vaak ten onrechte herkend. Kennelijk wordt de letterlijke bewoording waarin casusinformatie is gesteld snel vergeten, maar zijn artsen niet geneigd voor een ziekte prototypische informatie automatisch af te leiden, wanneer deze in de casus niet vermeld wordt.

Samenvattend kan gesteld worden dat ziektescripts een aantal belangrijke eigenschappen die tot uitdrukking komen in waarneembaar gedrag, zoals werkingssnelheid en geheugenprestaties op een herkenningstaak, gemeen hebben met de scripts van Schank en Abelson (1977). Bovendien kan ook geconcludeerd worden dat ziektescripts met toenemende praktijkervaring steeds meer geïntegreerd en geconsolideerd worden. De veronderstelde differentiële rol van Enabling Conditions en Consequences komt echter niet zo goed uit de verf: in een vrije produktie taak worden weliswaar aan expertiseniveau gerelateerde verschillen gevonden in toegankelijkheid als gevolg van type instructie, maar in een ziektescript-herkenningstaak blijken er nauwelijks verschillen te bestaan tussen ervaren huisartsen en studenten wat betreft geheugenprestaties voor beide ziektescript-componenten.

In Hoofdstuk 6 tenslotte worden de implicaties besproken van de huidige onderzoeksresultaten voor de representatie van medisch-diagnostische kennis in de vorm van ziektescripts, voor de ontwikkeling van deze scripts, en voor de geneeskundige onderwijspraktijk. Betoogd wordt dat ziektescripts een betere verklaring geven voor de onderzoeksresultaten dan andere vormen van kennisrepre- sentatie, zoals prototypie en mentale modellen. Voorts wordt beargumenteerd dat het ontwikkelingsaspect van de ziektescript-theorie een potentiële verklaring kan vormen voor de verschillen tussen beginners en experts in de geneeskunde, zoals die in het huidige onderzoek en de voorlopers ervan gevonden zijn. Enkele voorlo- pige suggesties worden gedaan met betrekking tot de implicaties van de huidige resultaten, en de ziektescript-theorie, voor de praktijk van het geneeskundig onderwijs. Tenslotte worden worden voorstel afgeleid voor vervolgorunderzoek.