Medical decision making is heavily influenced both by medically relevant and extraneous (nonmedical) factors (McKinlay et al., 1997). Nonmedical factors include characteristics of the patient, physician, health care system, and the decision-making environment (Flynn, van Schaik, van Wersch, Douglass, & Cann, 2003). A specific type of medical decision—the referral of patients by one physician (typically in primary care) to another physician (typically in secondary care), either temporarily or permanently for specialist advice, investigations, or treatment—has attracted a substantial amount of international research attention because of the high costs involved and the potential negative impact on patients. As with other types of medical decision making, a plethora of nonmedical factors is reported to exert an influence on general practitioners’ (primary care physicians’) referral behavior for a wide variety of medical and surgical procedures (Flynn et al., 2003).

The finding that referral behavior varies as a function of medical and nonmedical factors and possible interaction effects between these factors (Flynn et al., 2003) raises questions as to whether patients are receiving appropriate management in cases of overreferral (high referral rates) or are being denied access to medical and surgical procedures that could be beneficial to management in cases of underreferral (O’Donnell, 2000). These questions become ever more pressing with an increasing awareness of errors in medicine. For example, it is recognized that there is a relatively small, but important and alarming, number of poorly performing or incompetent physicians, with estimates of 15% in primary care (Boseley, 2004). Nevertheless, there is still a lack of evidence to suggest that the variation in referral rates of primary care physicians has detrimental clinical consequences for patients in terms of increased morbidity and mortality (Hippisley-Cox et al., 1997). Moreover, there is a dearth of research on referral behavior that considers the knowledge structures of primary care physicians.

Illness Script Theory

The concept of illness scripts has been proposed as the most pertinent type of knowledge structure operating in medical contexts (Charlin, Tardiff, & Boshuizen, 2000; Custers, Regehr, & Norman, 1996; Feltovich & Barrows, 1984). According to illness script theory (Charlin et al., 2000), illness scripts are narrative structures with three main components (see Figure 1): (a) enabling conditions—both medical and nonmedical contextual patient-related factors that influence the probability of disease (e.g., age, gender, occupation); (b) faults—the major real malfunctions that typify a particular illness, described in biomedical terms (e.g., the invasion of tissue by pathogenic organisms); and (c) consequences—various complaints, signs, and symptoms that might result from a specific fault. Illness scripts are believed to be activated as integrated units, instantiated by data available for a particular patient. Scripts are instantiated by substituting default slot values and characteristics with the actual features and values of that patient. In clinical contexts, physicians are frequently reasoning under constraints of time pressure, and therefore lengthy deliberations based on “first principles” of biomedical knowledge are often precluded (even more so in primary care), but scripts provide an efficient representation for storing and using medical knowledge (Charlin et al., 2000; Patel, Arocha, & Kaufman, 1999). Because of this, illness scripts aid classification in medical contexts.
Enabling conditions are typically available to a physician in the early stage of a medical consultation. These conditions can include medical and specific other information that influences the probability of a specific disease. Enabling conditions are probabilistically, rather than deterministically and causally, related to faults. For example, excessive smoking may increase the probability of lung disease, but in a particular patient this factor may be unrelated. Some consequences may also be available early during a consultation from a patient’s complaints or from a conspicuous sign. However, a physician predominantly learns about consequences as a result of questioning and examining the patient or requesting further investigations such as full blood count tests.

As a result of the content and nature of medical training, in terms of illness scripts, less experienced physicians are expected to be mainly guided by consequences. However, experienced physicians are assumed to have developed richer illness scripts and use both enabling conditions and consequences in diagnostic situations. These predictions were confirmed in previous research (Custers, Boshuizen, & Schmidt, 1996). Further research (Custers, Boshuizen, & Schmidt, 1998) found that reporting of enabling conditions increased and reporting of biomedical information (faults) decreased with increasing expertise in medicine. More specifically, reporting of enabling conditions for a particular disease was also positively related to the number of patients seen with that disease (or fault).

Given that illness scripts are recalled from long-term memory, and memory is based on experiences of the social world (i.e. socially constructed as a result of experience practicing as a physician), they hold the potential for confounding the assessment of the objective evidence that is presented (i.e. consequences). In particular the enabling components may be susceptible to confounding. (van Schaik & Flynn, 2003, p. 250)

This component is potentially prone to biases such as the representativeness and availability heuristics, which may arise when physicians are exposed to unrepresentative samples of patients with particular diseases. Reliance on representativeness may trigger the base-rate fallacy (Kahneman & Tversky, 1973), with physicians ignoring the importance of base-rate information available from book learning, clinical trials, and epidemiological studies. Reliance on the availability heuristic may lead to inaccurate estimates of the prevalence of common diseases (“salience bias”), where decision making is influenced by the characteristics and outcomes of patients from recent consultations (Wilkin, 1992).

The reliance on these two heuristics may account for the influence of nonmedical factors on medical decision making.

Reporting of illness script components still continues to change after basic medical training (Custers et al., 1998). However, from previous research it is not clear whether, and to what extent, changes in illness scripts still occur in experienced physicians. More specifically, given the increasing awareness of errors in medical practice, the question arises as to what extent knowledge structures differ between experienced physicians as a function of measurable differences in other variables. (In no way whatsoever do we suggest that the participants in the current research study were performing poorly.) Furthermore, previous research has not investigated systematically whether, and to what extent, specific values of individual illness script components (“slots”)—rather than a set of components—and their interactions influence the referral behavior of primary care physicians. Charlin et al. (2000) highlighted “a need for more research to validate this theoretical construct [illness script] and to determine its effective contribution to the diagnostic process” (p. 187). Therefore, the current research investigates illness script components in the context of primary care physicians’ referral behavior for suspected gastrointestinal disorders.

Referral Behavior and Illness Scripts

Given the extensive training that primary care physicians receive and their subsequent experience in practicing medicine, it is likely that referral behaviors—just as other types of medical decision—are influenced by their knowledge structures (Charlin et al., 2000). Indeed, it has been found that differences in clinical expertise between physicians are related to the knowledge that they possess rather than their problem-solving ability (Feltovich, Coulson, Spiro, & Dawson-Saunders, 1992; Patel & Groen, 1986). Previous research on referral behavior has identified various (medical and nonmedical) factors that affect referral behavior (Flynn et al., 2003), but no attempt has been made to explain variability in this behavior using a theoretical framework of knowledge representation such as illness script theory. Therefore, the current study uses the framework of illness scripts to investigate referral behavior. Because of the probabilistic nature of the activation of illness scripts primarily and because of the imperfect state of knowledge in medicine more generally, primary care physicians cannot always arrive at a definitive diagnosis during a consultation. Consequently,
there frequently is a need for primary care physicians to refer patients to secondary care. Major motivations for referrals to secondary care include learning more about the consequences, suspected faults, or both to exclude, confirm, or identify pathology to facilitate patient management. Typically, decisions to refer are based on an incomplete provisional diagnosis that is informed by the activation of illness scripts that involve particular combinations of consequences and enabling conditions. According to Patel, Groen, and Patel (1997), test results are mainly used to confirm a diagnosis or to rule out alternatives. However, in the framework of illness script theory, the results of tests gained from referrals can also provide new values for consequences that lead to new provisional diagnoses.

On the basis of previous work on the nature and development of illness scripts and their potential role in referral behavior, we proposed the following specific hypotheses:

Hypothesis 1: Consequences act as referral drivers (factors that increase the likelihood of a decision to refer being made).

Hypothesis 2: Enabling conditions act as referral drivers.

Hypothesis 3: There are interaction effects of enabling conditions and consequences on referral behavior.

Hypothesis 4: Experience in practicing as a physician influences referral behavior.

Hypothesis 5: Because knowledge of enabling conditions develops as a function of experience, this experience moderates the effect of enabling conditions on referral behavior. (Knowledge about consequences has been learned at medical school and is assumed to remain stable subsequently.)

These hypotheses are illustrated in Figure 2.

Method

Research Methodology

To date, research investigating referral behavior has been dominated by nonexperimental methods such as observational (database) and postal questionnaire surveys (Flynn et al., 2003; McKinlay et al., 1997). On the one hand, observational studies typically enable only a narrow focus of investigation in terms of referral behavior, as medical databases contain more medical factors than nonmedical factors and are generally not designed to answer the research questions being investigated, although they can contain many thousands of cases that are representative of the population (van Wersch, van Schaik, & Flynn, 2003). On the other hand, questionnaire surveys suffer from poor response rates and can produce nonrepresentative samples (Feldman et al., 1997).

The use of full-factorial designs is restricted by the large number of medical and nonmedical factors that warrant attention and the prohibitively lengthy process of recruiting adequate numbers of physicians to meet conventional levels of statistical power (McKinlay et al., 1997). Half-factorial (as well as other fractional-factorial) designs are an economical alternative to full-factorial designs, as they enable researchers to cost-effectively achieve perfect balance between multiple factors and maintain adequate statistical power (Kirk, 1995). Half-factorial designs reduce the number of study conditions by 50%, so fewer numbers of participants need to be recruited, but at the same time limiting the number of interaction effects that can be tested. Research has used half-factorial designs previously to investigate medical decision making, for example examining factors influencing prescribing behavior for respiratory tract infections and the management of breast cancer (McKinlay et al., 1997; Stephenson, Henry, & Norman, 1988). However, an extensive search of the literature revealed that half-factorial design methodology has not been applied to investigating the referral behavior of primary care physicians.

Referral drivers and referral inhibitors (factors that decrease the likelihood of decision to refer) associated with decisions to refer to a specialist domain such as gastroenterology has received little attention in the literature (Cornford & Cann, 2002). The current study conceptualizes referral drivers and inhibitors in the framework of illness scripts. Therefore, the current research aimed to examine the influence of enabling conditions and
consequences on general practitioners’ referral behavior for common upper
and lower gastrointestinal symptoms using half-factorial design methodology.

**Participants**

The participants were one group of 32 full-time general practitioners in the Tees Health District, Middlesbrough, United Kingdom, who referred patients with suspected gastrointestinal disorders to the Department of Gastroenterology at James Cook University Hospital (JCUH). The hospital is part of South Tees Hospitals National Health Service Trust and serves a population of over 1.5 million in the surrounding areas. Using a list of general practitioners in the district who referred patients to JCUH (all general practitioners in the United Kingdom must refer patients to secondary care for endoscopy procedures), we selected a representative mix of general practitioners in terms of size and location of practice and time since qualification as a general practitioner.

The majority of the general practitioners were male (75%), practiced in suburban/urban areas (78%), and operated under a system of general medical services (69%). The means (Midn; SD) for physician’s age, years in general practice, years in current practice, practice list size, and booking rate (consultations per hour) were 44.2 (43; 8.1) years, 15.4 (16; 8.8) years, 13.3 (15; 8.9) years, 8.715 (8,250; 2,831), and 7.2 (7; 1.4), respectively. All the general practitioners worked in group practice.

**Materials and Design**

The findings of a previous qualitative study that identified potential drivers and inhibitors of referral decisions for patients with gastrointestinal symptoms (Cornford & Cann, 2002) were used as a basis for constructing written vignettes. Five dichotomous patient factors were selected for inclusion in the vignettes: weight loss (present or absent), full blood count (abnormal or normal), age (38 or 52), gender (male or female), and socioeconomic status (SES; working class or professional). In this article age always refers to patient’s age, and physician’s age refers to physician’s age.) Weight loss was represented as a reduction in body mass of 4 lbs (1.81 kg) in the previous 2 months. In terms of illness script theory, the first two factors were consequences. The other three were enabling conditions and have been reported as significant predictors of referrals in previous research (gender: Delnoij & Spreeuwenberg, 1997; Franks & Clancy, 1997; age: Lee, Lum, Hillman, & Bauman, 1994; Wilkin & Smith, 1987; SES: Fleming, Crombie, & Cross, 1991).

Using a half-factorial design methodology \((\frac{1}{2} \times 2^{k-1} = \text{number of combinations, where } k \text{ is the number of predictor variables})\), 16 combinations of the five dichotomous patient factors (i.e., one level of each factor was included in exactly 50% of the vignettes) were embedded in a \(3 \times 2\) full-factorial design to create 96 \(6 \times 16\) different vignettes (see Figure 3). The factors in the full-factorial design were gastrointestinal symptom complex (dyspepsia, rectal bleeding [fresh with the stool], and change in bowel habit [two or three loose motions per day]) and symptom duration (long and short represented by 4 and 2 weeks, respectively). In terms of illness script, both symptom complex and symptom duration in the full-factorial design were consequences. To uncover physicians’ use of and potential interactions between referral drivers, patient characteristics such as 2-week duration of symptoms were included in the design, which made the appropriateness of referral decisions ambiguous in some cases and more definite in others. The fault in each case was suspected gastrointestinal disorder, but this was not stated because the participants’ task was to decide whether to refer or not refer, and stating the specific fault would have invalidated this task.

Vignettes with abnormal full blood counts included a simulated laboratory report with fixed values as follows: hemoglobin, 10.9; mean corpuscular volume, 70; and mean corpuscular hemoglobin concentration, 22, with ratios of white blood cells and platelets indicated as normal. A series of caveats were included with each vignette (see the Appendix). In case of a normal test result for full blood count no simulated lab report was included.

**Procedure**

General practitioners were contacted by telephone and invited to participate by either Paul Cann or Andrew Douglass. At the study session, conducted face-to-face by Darren Flynn in a general practitioner’s office, informed written consent was obtained in each case. Each general practitioner was presented with six vignettes in random order (one from each combination of values of symptom complex and duration); each vignette was selected as follows: a combination of values of the five factors from the half-factorial design was randomly selected without replacement. The physician was then requested to make a decision to refer the patient or not to refer (see the Appendix for an example of a vignette). In each case the general practitioner was informed that the patient was presenting for the first time and that subsequent questions that were put to him or her were dependent on the initial referral decision.

In cases of an initial decision to refer, general practitioners were asked to state what they expected to gain from the referral in terms of confirming, excluding, or identifying pathology. This was followed by asking general practitioners to estimate the probability of pathology being confirmed, excluded, or identified after a further assessment and/or investigation (e.g., colonoscopy) in secondary care had been undertaken. After the study session, general practitioners were debriefed and thanked for their cooperation. The study materials were initially piloted with 6 general practitioners, and improvements were made to the materials and procedure on the basis of their comments.

**Data Analysis**

We used logistic regression models to test the hypotheses, with initial referral decision as the outcome measure. Similar to multiple regression, logistic regression relates to a set of predictor variables with an outcome variable, but the outcome is measured at the nominal level instead of the interval level (Field, 2005; Menard, 2002); \(R^2\) in logistic regression is analogous to \(R^2\) in multiple regression, representing the improvement in prediction of the outcome by applying a logistic regression model with a given set of predictors compared to a model without predictors. The odds ratio (OR) of an individual predictor in logistic regression is similar to a regression coefficient \((b)\) in multiple regression, representing the change in odds of the outcome that result from a unit change in the predictor. First, models with both consequences and enabling conditions were investigated (Models 1 and 1a) to test Hypothesis 3. Next, models with only consequences were examined (Models 2 and 2a) to test Hypothesis 1. Models with and without enabling conditions were compared to determine the significance of enabling conditions over and above the effect of consequences to test Hypothesis 2. Furthermore, the effect of physician factors such as experience in practicing medicine was investigated with both consequences and enabling conditions (Models 3, 3a, and 3a′) and with consequences only (Models 4 and 4a) to test Hypotheses 4 and 5 (as well as to allow further tests of Hypotheses 1 and 2). Next, the hypotheses were tested with reasons given by general practitioners for their referral decisions (using logistic regression analysis) and estimated probability of disease (using multiple regression analysis) as outcome measures.

**Results**

**Analysis of Referral Rates**

The outcome measure was initial referral decision (“refer” or “do not refer”). An initial inspection of Table 1 appears to show that referral rates are higher with presence of weight loss, abnor-
Figure 3. Full-factorial design of Symptom Complex × Duration interaction with half-factorial design of five patient factors embedded. 52 = age 52; 38 = age 38; PWL = presence of weight loss; AWL = absence of weight loss; M = male, F = female.
mal full blood count test result, male gender, older age, and rectal bleeding than when these characteristics are not present in a patient. The following results of inferential statistics back up this first impression and, most important, are used to test the hypotheses set out previously.

Consequences and enabling conditions as predictors: Models 1 and 1a. The effect of physician as a covariate (the 32 physicians coded as 31 dummy variables) on the outcome referral decision was significant. Controlling for the effect of physician, we found that Model 1 with the following significant enabling conditions and consequences explained a significant amount of variability in the outcome referral decision: age, gender, weight loss, full blood count test result, symptom complex (rectal bleeding, dyspepsia, and change in bowel habit), and duration (see Tables 1 and 2). Referrals were statistically significantly more likely for patients with the following characteristics: age 52, male gender, presence of weight loss, abnormal full blood count test results, rectal bleeding (as opposed to change in bowel habit and dyspepsia), and symptom duration of 4 weeks.

Explained variability increased significantly when the effects of the following significant interactions were added to Model 1 (Model 1a; supporting Hypothesis 3): Full Blood Count/H11003 Gender, Symptom Complex/H11003 Weight Loss/H11003 Full Blood Count, Duration/H11003 Weight Loss, and Symptom Complex × Weight Loss × Age (see Table 2). A Hosmer-Lemeshow test of overall model fit showed that actual frequencies of referrals increased with increased predicted probability of referrals, $\chi^2(8) = 1.89, p = .98$, and sensitivity (95%) and specificity (97%) were extremely high. Moreover, a test of absolute (rather than relative) fit of the model demonstrated that the model predicted referral decisions without significant error, $\chi^2(45) = 44.35, p > .25$.

The first interaction effect shows that the relative increase in referral rates from female to male patients was higher with abnormal than with normal full blood count test results. The second interaction effect demonstrates that with a normal full blood count test result, there was a relative increase in referral rate from no weight loss to weight loss for all three symptom complexes; however, with an abnormal full blood count test result, there was an increase for two of the symptom complexes but not for dyspepsia. The third interaction effect shows that for patients aged 38, the relative increase in referral rate from no weight loss to weight loss was higher with a 4-week duration; however, for patients aged 52, the relative increase in referral rate from no weight loss to weight loss was higher with a 2-week duration. The fourth interaction effect demonstrates that the relative increase in referral rate from no weight loss to weight loss was higher for male patients than for female patients. The fifth interaction effect shows that for patients aged 38, the relative increase in referral rate from no weight loss to weight loss was highest with rectal bleeding; however, for patients aged 52 the relative increase in referral rate from no weight loss to weight loss was highest for change in bowel habit.

Consequences as predictors: Models 2 and 2a. Controlling for the effect of physician, we conducted further logistic regression analysis, which revealed that Model 2 with the following significant consequences explained a significant amount of variability in the outcome referral decision (supporting Hypothesis 1): weight loss, full blood count, symptom complex, and duration (see Tables 1 and 2). Explained variability increased significantly when we
added the effect of the following significant interaction to Model 2 (Model 2a): Weight Loss/H11003 Full Blood Count Test Result/H11003 Symptom Complex (see Table 2). A Hosmer-Lemeshow test showed that actual frequencies of referrals increased with increased predicted probability of referrals, χ²(8) = 7.79, p = .45, and sensitivity and specificity were 86% and 83%, respectively.

A test of Model 1a against the nested Model 2a revealed a significant difference in explained variability. Thereby it was established that—after controlling for the effect of physician—enabling conditions had a significant effect over and above consequences (supporting Hypothesis 2). A conceptual summary of the test results of Models 1, 1a, 2, and 2a is presented in Figure 4.

Physician variables, consequences, and enabling conditions as predictors: Models 3 and 3a. The effect of physician factors (individual-practitioner and practice characteristics) on referral decision was investigated next. It was not possible to expand the previous models with physician factors as additional predictors because these factors were nested within and linearly related with physicians. The effect of the following physician factors as covariates on the outcome referral decision was significant (supporting Hypothesis 2).

### Table 2

<table>
<thead>
<tr>
<th>Model Source</th>
<th>df</th>
<th>χ²</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Physician</td>
<td>31</td>
<td>71.02**</td>
<td>.27</td>
</tr>
<tr>
<td>1 Consequences and enabling conditions</td>
<td>7</td>
<td>90.97**</td>
<td>.61</td>
</tr>
<tr>
<td>1a Interactions involving consequences and enabling conditions</td>
<td>7</td>
<td>59.64**</td>
<td>.83</td>
</tr>
<tr>
<td>2 Physician</td>
<td>31</td>
<td>71.02**</td>
<td>.27</td>
</tr>
<tr>
<td>2 Consequences</td>
<td>5</td>
<td>61.92**</td>
<td>.50</td>
</tr>
<tr>
<td>2a Interactions involving consequences</td>
<td>2</td>
<td>6.28*</td>
<td>.52</td>
</tr>
<tr>
<td>1–2 Enabling conditions</td>
<td>2</td>
<td>29.05**</td>
<td>.09</td>
</tr>
<tr>
<td>1a–2a Enabling conditions and interactions involving enabling conditions</td>
<td>7</td>
<td>82.41**</td>
<td>.31</td>
</tr>
<tr>
<td>3 Physician variables, including interactions</td>
<td>4</td>
<td>17.49**</td>
<td>.07</td>
</tr>
<tr>
<td>3 Consequences and enabling conditions</td>
<td>7</td>
<td>80.07**</td>
<td>.38</td>
</tr>
<tr>
<td>3a Interactions of physician variables with consequences and enabling conditions</td>
<td>6</td>
<td>35.88**</td>
<td>.52</td>
</tr>
<tr>
<td>3a* Interactions of physician variables with consequences and enabling conditions + interactions from Model 4a</td>
<td>7</td>
<td>36.41**</td>
<td>.52</td>
</tr>
<tr>
<td>4 Physician variables, including interactions</td>
<td>4</td>
<td>17.49**</td>
<td>.07</td>
</tr>
<tr>
<td>4 Consequences</td>
<td>5</td>
<td>56.23**</td>
<td>.29</td>
</tr>
<tr>
<td>4a Interactions of physician variables with consequences</td>
<td>1</td>
<td>6.40*</td>
<td>.31</td>
</tr>
<tr>
<td>3–4 Enabling conditions</td>
<td>2</td>
<td>23.75**</td>
<td>.09</td>
</tr>
<tr>
<td>3a–4a Enabling conditions and interactions involving enabling conditions</td>
<td>8</td>
<td>53.76**</td>
<td>.21</td>
</tr>
</tbody>
</table>

**Note.** Model 3a* was constructed to include the interactions from Model 4a to make sure Model 4a was nested within Model 3a*. This was necessary because Model 4a was not nested within Model 3a, which would invalidate a test of Model 3a against Model 4a.

* p < .05. ** p < .01.

![Figure 4](image-url)

**Figure 4.** Conceptual representation of test results of Models 1, 1a, 2, and 2a. Arrows indicate significant main and interaction effects. Consequences include main effects of consequences and the effect of the Symptom Complex × Weight Loss × Full Blood Count interaction.
ing Hypothesis 4): physician’s age, experience, and booking rate as well as the effect of the Physician’s Age × Experience interaction (see Table 2). With increasing booking rate, the odds for referrals increased. With increasing physician’s age and decreasing experience, referrals increased, but the (negative) effect of experience was only significant for physician’s age above the median age (43 years; OR = 0.89, 95% confidence interval [CI] = 0.81, 0.98, p < .05) but not for the remaining physicians.

Controlling for the effect of the significant physician factors, logistic regression analysis demonstrated that Model 3 with the following enabling conditions and consequences explained a significant amount of variability in the outcome referral decision (refer or do not refer): weight loss, full blood count, symptom, duration, age, and gender (see Table 2). Explained variability increased significantly when we added the effects of the following significant interactions to Model 3 (Model 3a; supporting Hypothesis 5; physician variables in italics): Experience × Booking Rate, Experience × Experience, Booking Rate × Weight Loss, Booking Rate × Weight Loss × Age, and Experience × Full Blood Count Test Result × Symptom Complex (see Table 2). A Hosmer-Lemeshow test showed that actual frequencies of referrals increased with increased predicted probability of referrals, $\chi^2(8) = 11.00, p = .20$, and sensitivity and specificity were 85% and 84%, respectively.

The first interaction effect shows that at lower booking rates the increase in relative referral rate with age was greater with increasing experience; however, at higher booking rates this increase was smaller with increasing experience. The second interaction effect demonstrates a higher relative referral rate for male patients in less experienced physicians but no difference in relative referral rate in more experienced physicians. The third interaction effect shows that the increase in relative referral rate from no weight loss to weight loss was higher with increasing booking rate. The fourth interaction effect demonstrates that for patients aged 38 relative referral rates decreased from no weight loss to weight loss with lower booking rates, but there was increase at higher booking rates; however, for patients aged 52 relative referral rates increased with presence of weight loss and lower booking rates but decreased with higher booking rates. The fifth interaction effect demonstrates that in less experienced physicians the increase in referral rate from normal to abnormal full blood count test results was highest for change in bowel habit; however, in more experienced physicians the increase was largest for dyspepsia.

Physician variables and consequences as predictors: Models 4 and 4a. Controlling for the effect of the same significant physician factors, further logistic regression analysis revealed that Model 4 with the following consequences explained a significant amount of variability in the outcome referral decision (providing support for Hypothesis 1): weight loss, full blood count, symptom, and duration (see Table 2). Explained variability increased significantly when we added the effect of the following significant interaction to Model 4 (Model 4a): Experience × Full Blood Count (see Table 2). A Hosmer-Lemeshow test showed that actual frequencies of referrals increased with increased predicted probability of referrals, $\chi^2(8) = 8.56, p = .38$, and sensitivity and specificity were 75% and 76%, respectively. The interaction effect shows that the increase in relative referral rate from normal to abnormal full blood count test results was higher in more experienced physicians.

Model 3a+. The interaction effect from Model 4a and all the interaction effects of Model 3a were added to Model 3 to obtain Model 3a+, in which Model 4a was nested. A test of the difference between Model 3a+ and the nested Model 4a revealed a significant difference in explained variability (see Table 2), thereby establishing—that enabling conditions had a significant effect over and above consequences when taking into account physician factors that were predictive of referral decisions. This provided further support for Hypothesis 2. A conceptual summary of test results of Models 3, 3a, 3a+, 4, and 4a is presented in Figure 5.

Analysis of Reason for Referral and Probability of Disease

We analyzed reason for referral—confirm, exclude, or identify disease—as an outcome with a separate logistic regression analysis for each reason. (Physicians could choose more than one reason for a particular vignette.) The effect of physician on all three reasons

![Figure 5](image_url)
for referral was significant: confirming disease \( (R^2 = .28), \chi^2(31) = 71.28, p < .01; \) excluding disease \( (R^2 = .33), \chi^2(31) = 48.21, p < .05; \) and identifying disease \( (R^2 = .43), \chi^2(31) = 110.45, p < .01. \) Controlling for the effect of physician, we found that none of the three enabling conditions and four consequences was a significant predictor of confirming, excluding, or identifying disease.

In the analyses of effects of physician factors on reason for referral, only one factor—number of clinical sessions per week—was a significant predictor of confirming disease \( (OR = 0.82, 95\% CI = 0.68, 0.99, R^2 = .02), \chi^2(1) = 4.21, p < .05, \) and with more sessions per week, the odds of referring in order to confirm disease decreased. Only one physician factor—type of practice—was a significant predictor of excluding disease \( (OR = 0.80, 95\% CI = 0.66, 0.97, R^2 = .02), \chi^2(1) = 5.56, p < .05, \) and with more sessions per week, the odds of referring in order to identify disease decreased. Controlling for the effect of these predictors, we found that none of the three enabling conditions and four consequences was a significant predictor of confirming, excluding, or identifying disease.

Estimated probability of disease—to confirm, exclude, and identify pathology—was analyzed as an outcome, with a separate multiple linear regression analysis for each probability. Prior to this, probabilities were transformed to logarithmic odds, resulting in distributions closer to a normal distribution. The effect of physician on probability for all three reasons for referral was significant: confirming disease \( (R^2 = .46), F(31, 157) = 4.29, p < .01; \) excluding disease \( (R^2 = .62), F(27, 49) = 2.99, p < .01; \) and identifying disease \( (R^2 = .46), F(31, 157) = 4.33, p < .01. \) When we controlled for the effect of physician, none of the three enabling conditions and four consequences was a significant predictor of probability to confirm, exclude, or identify disease.

In the analyses of effects of physician factors on estimated probability of disease, only one factor—booking rate—was a significant predictor of probability to confirm disease \( (b = -0.27, 95\% CI = -0.48, -0.07, R^2 = .08), F(1, 75) = 6.87, p < .05, \) and with more sessions per week, probabilities to confirm disease decreased. None of the physician factors was a significant predictor of probability to exclude disease. Only one physician factor—number of clinical sessions per week—was a significant predictor of probability to identify disease \( (b = 0.27, 95\% CI = 0.04, 0.51, R^2 = .03), F(1, 187) = 5.11, p < .05, \) and with increasing booking rate, probabilities to identify increased. When we controlled for the effect of booking rate (in the cases of confirming and identifying disease), none of the three enabling conditions and four consequences was a significant predictor of probability to confirm, exclude, or identify disease.

**Discussion**

The results for the outcome initial referral decision support the five hypotheses: (1) consequences were significant predictors (test results: Models 2, 2a, 4, and 4a); (2) enabling conditions were significant predictors (test results: Model 1a compared with Model 2a, and Model 3a* compared with Model 4a); (3) there were significant interaction effects between consequences and enabling conditions (test results: Models 1a, 3a); (4) experience in practicing as a physician was a predictor (test results: Model 3); and (5) there were interaction effects between physician factors and enabling conditions (test results: Model 3a). However, physician factors also interacted with consequences (not predicted; test results of Models 3a and 4a), and there was a three-way interaction of a physician factor with an enabling condition and a consequence (not predicted; test results of Model 3a). The latter two findings suggest that knowledge and the use of consequences still develops in practicing physicians, although it was assumed previously that this knowledge does not develop significantly further after training at medical school (e.g., Custers, Boshuizen, & Schmidt, 1996).

From the test results of Models 1, 1a, 2, and 2a, the effect of the combined predictors (consequences, enabling conditions, and their interactions) was considerable, explaining more than 50% of the variability in the outcome measure referral decision, after taking into account the variability due differences between physicians. In the absence of established guidelines for \( R^2, \) Cohen’s (1988) guidelines for effect size of \( R^2 \) in multiple regression are applied here to \( R^2, \) the counterpart to \( R^2 \) in logistic regression. Following these guidelines, consequences and their interactions explained a quarter of the variability (large effect size), enabling conditions and their interactions explained more than 30% (large effect size), and the interactions of consequences and enabling conditions in combination explained more than 20% (large to moderate effect size) in referral rates. The variability between physicians was also relatively large, explaining more than 25% of variability. The latter finding confirmed previous research that found large variations in referral rates between groups of physicians (e.g., Franks, Williams, Zwanziger, Mooney, & Sorbero, 2000). Nevertheless, the very substantial effects of illness script components and their interactions demonstrate that our experimental manipulations explained a majority of the variability in referral decisions.

The results of Models 3, 3a, 3a*, 4, and 4a demonstrate that physician variables explained slightly more than 5% of the variability (small to moderate effect size) in referral rates. The interaction effects of physician variables with consequences and enabling conditions explained just under 15% of the variability (moderate effect size). The effects of consequences and enabling conditions and their interactions remained large, explaining more than 20% and more than 30% of the variability, respectively.

On the other hand, the main effects of consequences explained approximately twice as much variability in referral decisions (>20%) as the main effects of enabling conditions (9%). On the other hand, the interaction effects involving only consequences were relatively small, explaining 2% of variability, but the interaction effects involving enabling conditions explained 10 times as much variability (>20%). These results can be explained in terms of illness script theory (i.e., the characteristic consequences and enabling conditions). Consequences (or combinations thereof) can be sufficient to conclude that a particular fault exists. Enabling conditions, however, are only probabilistically related to faults, and enabling conditions without consequences being present (i.e., there are no symptoms) do not allow a conclusion about any fault. As a result of this, enabling conditions can be expected to affect decision making by moderating the effect of consequences, but consequences may affect this more independently; this assertion was confirmed by the different effect sizes for the main and
interaction effects between consequences and enabling conditions. Indeed, the majority of significant interaction effects in Model 1a were between an enabling condition and one or more consequences. However, one of the interactions did not involve any enabling conditions—the moderating effect of weight loss on the effect of full blood count test results was in turn moderated by symptom complex. This demonstrates that the power of particular consequences as referral drivers (e.g., weight loss) can depend on the presence of additional consequences (e.g., blood test results and symptom complex), and this is compatible with illness script theory. Although only experienced physicians were included in the current study, on the basis of the theory it would be predicted that in inexperienced physicians, enabling conditions would be less influential and there would only be interactions between consequences.

The main effect of specific physician variables (experience, physician’s age, and booking rate) was relatively small compared (small to moderate) with interaction effects of these variables with consequences and enabling conditions (moderate). Again, this result is not surprising, because particular characteristics of physicians that make them more likely to diagnose a particular fault without consequences being present (i.e., there are no symptoms) do not allow a conclusion about any fault. There were significant interaction effects in Model 3a between one or more physician factors and one or more consequences and/or enabling conditions. These findings demonstrate the moderating effect of physician factors on consequences and enabling conditions and, by implication, individual variations in illness scripts. For instance, there was a higher relative referral rate for male patients than for female patients in less experienced physicians but no difference in relative referral rate in more experienced physicians. Although it has been assumed that experienced physicians possess fully developed illness scripts, these findings demonstrate that who they are (in terms of age and experience) and their work pattern (booking rate) influences their referral behavior. The interaction effect between physician’s age and experience indicates that (a) in older physicians—but not in younger physicians—referral rate decreased with experience and (b) in physicians with less experience—but not in those with more experience—there was a trend toward an increase in referral rate with a physician’s age.

The lack of predictive power of consequences and enabling conditions in outcome measures other than initial referral decision may be due to the following. First, reasons for referral may have been varied and included factors such as fear of malpractice (a variable not included in the current study but a significant predictor in the results of Franks et al., 2000). Second, probability of confirming, excluding, or identifying an unspecified disease after results of further investigations prompted by a referral (as investigated in the current study) are different from probability of a specific disease (as in Custers, Boshuizen, & Schmidt, 1996). Therefore, consequences and enabling conditions may have failed as predictors because of this lack of specificity. Furthermore, the prediction of probabilities may have suffered from ceiling and floor effects (median estimated probabilities to confirm, exclude, and identify disease were .10, .90, and .10, respectively).

Prior to the current study, no research on referral behavior has been conducted (a) in the framework of illness scripts and (b) in the medical domain of gastroenterology. Furthermore, previous studies on illness scripts have not used factorial or half-factorial (or other fractional-factorial) designs that systematically combine values of individual illness script components. Despite this, studies on referral behavior have been conducted without reference to illness scripts and knowledge representation more generally. Franks and Clancy (1997) and Delnoij and Spreeuwenberg (1997) established that patient gender was a significant predictor of referrals. Previous research reported that patient age was also a predictor of referral decisions (Lee et al., 1994; Wilkin & Smith, 1987). Both these enabling conditions were confirmed as predictors in the current study. Although SES was the only nonsignificant enabling condition in the current study, previous research (Fleming et al., 1991) reported this variable to be a referral driver. The failure to confirm this finding in the current research may be due to a lack of differentiation in status (working class or professional, without further distinctions), the presence of more powerful referral drivers (consequences and other enabling conditions), and the use of vignettes rather than real patients. The consequences included in the current study (symptom complex, duration, full blood count test, weight loss) are classic alarm symptoms for gastrointestinal disorders and were confirmed as referral drivers.

Other related research on medical diagnosis, rather than referral behavior, has concluded that illness script components (both consequences and enabling conditions) are present in experienced physicians’ illness scripts (Custers, Boshuizen, & Schmidt, 1996), and the test results of our first three hypotheses are compatible with this assertion. Custers, Boshuizen, and Schmidt (1996) also found that with increasing experience from medical students to experienced physicians, there was a shift from using consequences toward the use of both consequences and enabling conditions. The results of testing our fifth hypothesis indicate that even within experienced physicians the use of enabling conditions still changed with increasing experience. Moreover, the use of consequences changed, although according to illness script theory (Custers et al., 1998), the use of consequences (because they are taught at medical school and their use should remain stable) should not change in experienced physicians.

The current study has demonstrated interaction effects between consequences and enabling conditions and between different consequences in experienced physicians. These results pertain to the first research issue identified by Charlin et al. (2000, “the nature of links within clinical knowledge networks,” p. 187). They also confirm illness script theory by providing support for a network of illness script components that act in concert to influence referral behavior for gastrointestinal disease (and by identifying particular dominant links between illness script components). The finding that physician factors interacted with consequences and enabling conditions when making referral decisions provides evidence for a differentiation in the use of illness script components in experienced physicians. This finding pertains to the third research issue identified by Charlin et al. (2000, “the dynamic or stable nature of scripts,” p. 187). This differentiation is not predicted by illness script theory, but it can be interpreted as a confirmation of previous research that has shown a wide variability in referral rates between physicians (Cummins, Jarman, & White, 1981; Delnoij & Spreeuwenberg, 1997; Franks et al., 2000; Wilkin & Smith, 1987). Furthermore, the interactions of booking rate with consequences and enabling conditions demonstrate that work pattern influences the use of illness script components, a finding not predicted by illness script theory. This indicates a need to incorporate factors of
the (work) environment into a theory of knowledge representation in medical contexts, as suggested by research on “cognition in context” (Hutchins, 1995). Our findings suggest that—at least in the context of referral behavior—illness scripts are influenced by even more factors than previously believed, including both individual and environmental physician factors.

Within the framework of illness script theory, previous research has demonstrated the effects of consequences and enabling conditions as well as physicians’ experience on their diagnoses (Boshuizen & Schmidt, 1995; Custers, Boshuizen, & Schmidt, 1996; Schmidt, Norman, & Boshuizen, 1990), and the current study has shown that these effects extend to referral behavior. These results indicate that consequences are not sufficient to account for medical decision-making behavior. In principle, the concept of illness scripts applies more generally to any medical or nonmedical domain that includes diagnosis and referral, for example clinical psychology, counseling psychology, and educational psychology, and research should be extended to these domains. In these domains, consequences are various psychological test results and records of clients’ behavior that are used to inform diagnoses, for example those based on criteria described in the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 1994). However, experienced psychologists may use information about enabling conditions even more so than consequences to inform their diagnoses and referrals to other medical, psychological, or other professionals. Potential reasons for this are that the mapping of consequences onto faults is more ambiguous in psychological domains than in medical domains and practitioners are poor at using the information of two or more variables simultaneously in making a judgment (Hastie & Dawes, 2001). Further research is required on practitioners’ use (or their lack of use) of illness script components to establish the relative weight that is assigned to different illness script components in various domains. This is important in particular because existing research, as summarized by Hastie and Dawes (2001), appears to have neglected the distinction between consequences and enabling conditions at best and to have focused on consequences at the expense of enabling conditions at worst. Results of this research can form a basis for the development of training programs (Charlin et al., 2000) and improving decision making.

A limitation of the current study is the use of written vignettes. More realistic representations are desirable to increase external validity of research results while retaining experimental control. For this purpose videotaped consultations can be produced (McKinlay, Potter, & Feldman, 1996). More realistic settings can be created by the use of interactive video productions, where physicians can “interrogate” patients; and simulated patients can provide even more realism. Interactive video can show a wider range of enabling conditions and more specific information, for instance about SES (e.g., workers in the asbestos industry and painters using solvents), that is relevant for clinical reasoning more generally and referral behavior in particular (in relation to the specific consequences of a particular patient). Another limitation of the current study is that the sample was drawn from a single health district. Therefore, the effect of regional differences on referral rate (“geography is destiny”; Gigerenzer, 2002, p. 101)—in addition to other referral drivers, including illness script components—could not be investigated.

A very important issue for future research is computer-based decision support (van Schaik, Flynn, van Wersch, Douglass, & Cann, 2004). In this context, van Schaik et al. (2004) concluded that use of a direct booking system in primary care has the potential to reduce the discomfort associated with clinical uncertainty by reducing the cognitive demands on general practitioners. Clinical uncertainty is indeed ubiquitous within medical contexts (more so in primary care), and it is widely regarded to be the strongest factor influencing the observed variability in the use of medical resources (Bachman & Freeborn, 1999). Computer-based decision support has the potential to reduce inappropriate decisions by those with poorly developed or biased illness scripts. Finally, a promising application of illness scripts is their use in teaching (Charlin et al., 2000), and further research in this area should include diagnosis, treatment, and management of disease.

Conclusion

Illness scripts are efficient knowledge representations that aid classification in diagnostic situations. In the current research, illness script components, both consequences and enabling conditions and their interactions, were found to influence referral behavior. Evidence was obtained for a linked structure of illness script components, and it appears that this structure is influenced by individual and environmental physician factors even after medical training, not only in terms of enabling conditions, but also in terms of consequences.

References


Appendix

Example of a Case Vignette (Change in Bowel Habit—4 Weeks’ Duration of Symptoms)

Please read the following case scenario of a patient consulting you for the first time:

The patient describes having two or three loose bowel motions per day for the last 4 weeks. The patient has lost 4 lbs of weight in the previous 2 months, and the results of a full blood count test indicated iron deficient anemia. The patient is 52 years old, male, and from a professional background.

Please assume the following:

1. There are no other signs of infection, pathology, or comorbid disorders.
2. There is no evidence of other causes of weight loss.
3. A routine examination (rectum and/or abdomen) revealed no abnormal findings.
4. There is no evidence of rectal bleeding.
5. The patient’s dietary habits and appetite are both healthy.
6. The patient’s general health and well-being is good.
7. The patient has no past or family history of GI disorder.
8. The patient drinks below the recommended safe level of alcohol.
9. The patient is a nonsmoker.
10. The patient has not been prescribed has been taking aspirin or NSAIDs.
11. The patient has not been away on an overseas holiday in the last 12 months.
12. Plasma viscosity, culture sensitivity, and fecal occult blood tests are all negative.

FULL BLOOD COUNT TEST—Summary of Results

HB = 10.9 MCV = 70 MCHC = 22 White Blood Cells/Platelets—Normal

What is your decision?

1. Refer the patient?
2. Do not refer at this time?

Note. GI = gastrointestinal; NSAIDs = nonsteroidal anti-inflammatory drugs; HB = hemoglobin; MCV = mean corpuscular volume; MCHC = mean corpuscular hemoglobin concentration.

Received February 3, 2005
Revision received May 20, 2005
Accepted May 20, 2005