The Imitation Game as a Method for Exploring Knowledge(s) of Chronic Illness

Robert Evans\textsuperscript{a} and Helen Crocker\textsuperscript{b}
\textsuperscript{a}Centre for the Study of Knowledge, Expertise and Science (KES), Cardiff School of Social Sciences, Cardiff University
\textsuperscript{b}Department of Public Health, University of Oxford

Abstract
Medical sociology has traditionally contrasted the bio-medical knowledge of healthcare professionals with the socially situated knowledge possessed by patients. These differences are particularly important in chronic conditions where patients can develop highly sophisticated understandings of their disease and its symptoms. In this paper we use a novel research method – The Imitation Game – to explore how far the experiential knowledge of patients can be appreciated and understood by medical professionals. We examine the extent to which professional dietitians were able to reproduce the discourse of people diagnosed with coeliac disease and show that dietitians were able to ‘pass’ as patients with coeliac disease but that a control group of lay people could not. We also briefly explore the domains of knowledge in which dietitians tended to be more or less successful and reflect on the utility of the research for medical sociology and the training of healthcare professionals.

Key Words: Imitation Game; Interactional Expertise; Lay Expertise; Coeliac Disease; Comparative Research

Introduction
One of medical sociology’s key achievements has been to show how bio-medical expertise can silence the socially situated knowledge of patients (e.g. Freund et al, 2003). Although social science has been a recognised part of the medical curriculum since the mid-1940s (Reid, 1976), concern about the tendency to treat ‘the chart’ and not ‘the patient’ remains (Epstein, 1999, cited in Freund et al 2003, p. 237-8). Here we approach this problem by exploring what medical practitioners are able to learn about the experience of living with a chronic condition through their interactions with patients. In doing so, we outline a new approach to practitioner-patient research and the method by which it can be conducted.

The paper is both methodological and substantive: it describes the Imitation Game method through its application to a genuine research question. The substantive issue is the extent to which medical professionals are able to break out of the bio-medical perspective and appreciate the lived experiences of the patients they treat. Dietitians are an interesting test case as their therapeutic effectiveness depends on patients implementing advice derived from bio-medical and other scientific research in everyday settings. By investigating how far dietitians are able to understand the practical reality of living with coeliac disease we also illuminate the broader question of what practitioners can learn about patient experience. Specifically, if dietitians are able to display high levels of understanding of what it means to live with coeliac disease, it suggests that bio-medical expertise can exist alongside an authentic understanding of the patient experience. If this is correct, then it
follows that a more patient-centred medical practice, in which medical professionals have access to both bio-medical knowledge and social understandings of illness, is an achievable goal.

The methodological interest lies in the development of a new research method – the Imitation Game – through which the distribution of these different bio-medical and patient expertises can be mapped. The idea that motivates the research is that dietitians’ interactions with patients with coeliac disease enable them to develop the specialist interactional expertise (Collins and Evans 2002, 2007) needed to understand the day-to-day problems of living with an illness they do not have and whose effects they do not, therefore, experience directly. Lay people will not have this experience and so will lack this understanding. The Imitation Game allows these differences to be quantified and hence shows the extent to which dietitians’ knowledge of coeliac disease is similar to that held by patients and different to that found in the lay population.

The paper is structured as follows. First we provide an overview of coeliac disease and briefly summarise some relevant literature on medical, experiential and lay knowledges. Next we introduce the two main concepts – interactional expertise and the Imitation Game – before describing our own research in some detail. Given the novelty of the method we devote considerable attention to the procedures used to collect and analyse the data. We conclude by reflecting on the implications of our work for future research.

**Background**

**Coeliac Disease**

Coeliac disease is a gastrointestinal autoimmune disorder affecting approximately one per cent of the UK population (NHS, 2010). The immune reaction is triggered by gluten, a protein found in wheat, barley, and rye, and causes damage to the villi that line the small intestine. If untreated coeliac disease can lead to malnutrition, but this is typically preceded by a number of other symptoms including chronic or intermittent diarrhoea, recurring abdominal pain, prolonged fatigue, unexplained anaemia, and sudden or unexpected weight loss (NICE 2009, p. 17).

The condition has no cure but symptoms can be controlled and damaged tissue repaired by avoiding all foods containing gluten. For this reason, dietitians play a significant role in the management of coeliac disease. Ideally, they provide dietary advice at diagnosis and ongoing support using:

the most up to date public health and scientific research on food, health and disease, which they translate into practical guidance to enable people to make appropriate lifestyle and food choices.
(BDA nd).

This translation role explains why understanding the patient perspective is crucial. Not only do medical advice and dietary rules have to be applied in social settings, the need for dietary vigilance imposes social, emotional and financial costs on patients and their families (Hallert et al, 2002; Sverker et al, 2005, 2007; Whitaker et al, 2009). Understanding the multiple social contexts in which patients enact medical advice is, therefore, central to the work of dietitians.

**Medical, Experiential and Lay Knowledges**

The recognition that patient experience can be a vital component in the successful management of chronic illness has many echoes with work in Science and Technology Studies (STS). For example, Wynne’s studies of the nuclear industry (Wynne, 1992), Irwin’s analysis of environmental issues (Irwin, 1995), and the more
general calls for more ‘constructive’ or ‘interactive’ forms of technology assessment (Rip et al, 1995; Grin et al, 1997) all highlight the relevance of local knowledge in the application (and contestation) of scientific advice. Studies of health and illness that reinforce these findings include Popay and Williams’ work on public health (Popay and Williams, 1996), Arksey’s study of Repetitive Strain Injury (Arksey, 1998), Layton et al’s study of parents of children with Down’s Syndrome (Layton et al, 1993) and Epstein’s study of AIDS treatment activists (Epstein, 1996).

As a result of this work, the increased status of lay and other forms of knowledge is now being reflected in the almost routine calls for public engagement in science and technology policy (e.g. House of Lords, 2000; Wilsdon and Willis, 2004). In the case of medical practice, the recognition of the patient’s perspective is changing the provision and evaluation of medical care in a number of ways. Examples include participatory mechanisms that seek to enrol lay citizens and patients into the design and evaluation of healthcare policies (Martin, 2008), the evaluation of routine medical care through the use of Patient Reported Outcome Measures (PROMs), and the increasing use of the internet by patients to research their condition and treatment options (Ziebland, 2004). Taken together, this decline in the authority of medical authority and associated shift towards more shared or patient-centred decision-making has led some to suggest that a ‘new medical conversation’ is developing (Mazur, 2002).

Whilst many of these changes are undoubtedly positive there is a risk that the epistemological levelling may erase legitimate distinctions between lay, patient and medical knowledge. This point is made particularly clearly by Lindsay Prior (2003) and, like Prior, we are concerned that failing to distinguish between different kinds and levels of expertise may mask important differences. That said, we are certainly not claiming that patient knowledge is irrelevant. Rather, our concern is with what follows from the recognition that different social groups have different experiences. In particular, how do the different kinds of expertise that these different experiences give rise to get articulated and shared with others?

If we accept that the STS and other research cited above has shown that patient groups are able to understand the relevant bio-medical science, then the critical question for medical practice is how far the experiences that inform the patient perspective can be shared by the scientifically trained professionals who treat them. The implication of patient-centred care is that there should be a substantial overlap between the knowledge that patients have about living with their condition and the understanding that medical professionals have of the same issue. As clinicians generally do not suffer from the conditions they treat, their knowledge of the patient perspective must be developed through the linguistic interactions with patients. More formally, we can say that, if patient-centred practice is to be realised, then medical staff will need to develop interactional expertise (Collins and Evans 2002, 2007) in the domain of patient experience.

**Interactional Expertise and the Imitation Game**

The idea of interactional expertise was first introduced by analogy to sociological fieldwork in science studies (Collins and Evans 2002, 2007). Competent practitioners (e.g. scientists) have the embodied, practical expertise needed to contribute to a domain of science (contributory expertise) and this includes expertise in the language used to describe practices within that domain (interactional expertise). In contrast, the sociological researcher will typically acquire fluency in the language of the domain but not the practices (i.e. sociologists of science do not usually contribute to the science they research, just as criminologists do not commit crimes or make arrests). This means that the outcome of sociological fieldwork is best seen as the acquisition of interactional expertise without the corresponding contributory expertise. Nevertheless, to the extent that the researcher has genuinely acquired interactional expertise in their field of study then he or she
can say everything a practicing contributory expert would say and for linguistic tasks is indistinguishable from such a person (Collins, 2004, 2011; Giles, 2006).

The Imitation Game method that we describe in this paper was developed by Collins and Evans to test this claim (Collins et al, 2006; Evans and Collins, 2010; Collins and Evans, 2013). The method allows researchers to examine how far one social group is able to reproduce the discourse of another and hence the extent to which that individual or social group can be said to possess interactional expertise in the relevant domain of practice. The format of the Imitation Game is similar to the Turing Test (Turing, 1950) used in Artificial Intelligence research in which a Judge has to distinguish between a computer and a human. In the Imitation Game, the Judge, who has the contributory expertise being investigated, has to distinguish between the discourse of a contributory expert (the ‘Non-Pretender’) and another person (the ‘Pretender’) who may, or may not, have interactional expertise. The basic Imitation Game design is summarised in Figure 1, which shows how the participants communicate using computers.³

**Figure 1: Basic Imitation Game Set-Up**

If the idea of interactional expertise is correct then Pretenders who have been immersed in the target community for a significant period will be able to reproduce the discourse of that community even if they have no practical expertise. Initial tests, conducted using the colour blind and those with perfect pitch have confirmed the existence of interactional expertise and more recent studies suggest the method can also be used to research topics such as sexuality, gender and religion (Collins et al, 2006; Collins and Evans, 2013).⁴

**Imitation Games with Dietitians**

Here we use the Imitation Game to explore how expertise about ‘living with coeliac disease’ is shared and distributed amongst patients, dietitians and lay people. The core argument is that if dietitians have succeeded in developing interactional expertise about coeliac disease – that is, if they have developed an understanding of the patients’ perspective – this will be revealed by the failure of judges with coeliac disease to reliably distinguish between genuine patients (Non-Pretenders) and dietitians (Pretenders) answering as if they were patients. In contrast, if dietitians lack interactional expertise they will be bad at pretending and judges should be both confident and accurate in their identifications.

To ensure that any success on the part of dietitians is not due to a more ubiquitous expertise in experiences related to food, we also ran Imitation Games in which the dietitians were replaced by lay people. If the expertise needed for Pretenders to succeed is not related to interactions with patients with coeliac disease then there should be no difference in the performance of lay people. If, however, the expertise is specific to the experience of coeliac disease, then we would expect lay participants to be worse at pretending.
Research Questions and Hypotheses

In principle, Imitation Games can give rise to one of two ‘ideal-type’ outcomes. If the participants charged with pretending have had a significant amount of interaction with the target group then they will have had the opportunity to develop interactional expertise in that domain of practice. In this scenario, the Judge should find it very difficult to tell who is who as both sets of answers (i.e. Pretender and Non-Pretender) will be equally plausible. This means that judges are forced to guess when asked to decide who is the Pretender and who is the Non-Pretender, giving rise to an equal number of right and wrong identifications. As this is effectively a chance outcome, Imitation Games that start with this expectation are called a chance condition. In this research, the Imitation Games involving dietitians should be a chance condition.

The other possibility is that there is little or no interaction between the groups. In this case the hypothesis is that the person charged with pretending will lack the interactional expertise needed to produce plausible answers. As judges should be able to recognise these mistakes and correctly identify the participants this is called an identify condition. In this research, the Imitation Games involving lay people should be an identify condition as we do not expect them to have the specialist knowledge needed to fool a judge with coeliac disease.

In practice, the extent to which the two conditions produce different outcomes is revealed by comparing the distribution of final guesses and calculating a summary statistic called the ‘Identification Ratio’ (IR). Although the underlying statistical principles are quite common, their application in this context is new. We now explain the procedure in order to set out its rationale and lay the groundwork for the interpretation of the Imitation Games involving dietitians.

First, the raw data for each identify or chance condition is re-coded as follows. Final guesses with a confidence level of 3 or 4 are classed as ‘Right Guesses’ when the judge was correct and ‘Wrong Guesses’ when they are not. All final guesses with a confidence level 1 or 2 are classed as ‘Don’t Knows’ regardless of whether the judge was right or wrong. The aggregate performance of Judges can then be summarised as shown in Table 1 (data adapted from Collins et al, 2006).

| Condition | Identify | Chance |
|-----------|----------|--------|
| Wrong Guess | 3 | 10 |
| Don’t Know | 7 | 51 |
| Right Guess | 16 | 9 |
| **Total** | **26** | **70** |

Next, the ‘excess’ or ‘net’ number of Right Guesses is calculated by subtracting the number of ‘Wrong Guesses’ from the number of ‘Right Guesses’ in each condition. This is done because the initial Right Guess category includes correct identifications based on a genuine difference between the two responses (e.g. one participant makes a mistake) and lucky guesses based on some spurious difference between the two sets of answers (e.g. the answer was shorter). If we assume that the second sort of guess is as likely to be right as it is to be wrong, then subtracting the Wrong Guesses from the Right Guesses eliminates this ‘noise’ and reveals the number of Right Guesses that can be attributed to the information contained in the dialogs. The difference between the Excess Right Guesses and the total number of Imitation Games in that condition makes up the balance of Remaining Guesses. This gives the simplified 2 x 2 table of results shown in Table 2.
Table 2: Sample 2x2 Table of Imitation Game Results

| Condition          | Identify | Chance |
|--------------------|----------|--------|
| Remaining Guesses  | 13       | 69     |
| Excess Right Guesses | 13     | -1     |
| Total              | 26       | 70     |

Finally, the Identification Ratio (IR) is calculated by dividing the number of Excess Right Guesses by the total number of Imitation Games in that condition. For the data given above, the IRs, to one decimal place, are 0.5 for the Identify condition and 0.0 for the Chance Condition.\(^5\)

To establish whether the difference between the Identify and Chance conditions is statistically significant, we use a Monte Carlo-style simulation method to estimate the distribution of IRs consistent with the frequencies of Right, Wrong and Don’t Know answers shown in Table 1 and hence the probability of the observed difference being due to chance. In this case the probability is less than 0.0001 so the conclusion that the two conditions are different is more than justified.\(^6\)

Using this terminology, we can now re-state the main research questions as follows:

- Imitation Games involving dietitians as Pretenders should conform to a chance condition outcome, meaning that the IR should be close to zero.
- Imitation Games involving lay people as Pretenders should conform to an identify condition outcome, meaning that the IR should be positive and based on previous research, more than about 0.3 (Collins et al, 2006; Collins and Evans 2013)
- The difference between the two conditions should be statistically significant.

Before discussing the results in detail, it can be noted that any bias in the recruitment works against the hypotheses being tested. For example, in the chance condition, members of coeliac support groups are more likely to be knowledgeable and pro-active in the management of their condition than non-members. As more knowledgeable judges should be better discriminators, this should make it harder for dietitians to ‘pass’ as people living with coeliac disease. This should have the effect of raising the IR and make the Chance condition look more like an Identify condition.

Similarly, for the identify condition, recruiting lay participants via a University notice board may have created a sample in which internet-savvy and articulate respondents are over-represented. As one would expect these people to be at least as good at ‘faking’ coeliac disease as participants who were, for example, less well educated or less familiar with the internet, any bias is unlikely to favour the hypothesis. Specifically, if the lay participants are especially good at pretending, judges should find it harder to correctly identify the participant with coeliac disease. This will lower the IR and make the expected Identify condition look more like a Chance condition.
Samples and Recruitment

For the Imitation Games involving dietitians, participants with coeliac disease were primarily recruited through a ‘Facebook’ group set up for the discussion of coeliac disease. Messages were sent to a random sample of 113 of the 2,222 members listed in July 2009. This approach resulted in thirteen research participants. A similar message posted to the Facebook group’s discussion board was less effective, recruiting just one extra participant. An additional five participants were recruited using a snowball sample based on people with coeliac disease that were known to the researchers. All participants met the following inclusion criteria:

- aged 16 or over;
- diagnosed with coeliac disease by endoscopy at least one year ago;
- access to an email account and the Internet.

Social networking sites were also used to recruit the participants needed to act as judges. The administrators of one Yahoo group and one Facebook group agreed to distribute participant request messages to all members of their groups. The Yahoo group had 1,340 members at the time of distribution, of which twelve members responded and seven actually participated. The Facebook group had 1,447 members at the time of distribution, which resulted in 113 replies of which 37 actually participated. As some individuals belonged to both groups, the total number of people receiving the participant request is less than the sum of the groups’ memberships.

Dietitians were recruited in three ways. Firstly, a database of freelance Dietitians – Dietitians Unlimited (www.dietitiansunlimited.co.uk) – was used to send information to 31 dietitians listing coeliac disease as one of their specialist areas. This resulted in five dietitians agreeing to take part in the research. Secondly, NHS hospitals with dietetics departments were identified using an Internet search engine. Ten departments were contacted by post or email, which led to the recruitment of a further five dietitians. Finally, dietitians were recruited using a Facebook group for dietitians. Sixteen of the group’s 395 members were contacted after being selected at random, and, of these, three agreed to take part in the research. As anybody could join this Facebook group, the professional status of members was checked against the Health Professions Council’s register (www.hpcheck.org).

The inclusion criteria for dietitians were similar to those for the participants with coeliac disease:

- registered dietitian;
- experience of working with patients with coeliac disease;
- access to an email account and the Internet;
- must not have coeliac disease themselves.

For Imitation Games involving lay people recruitment took place in July 2010. A new set of participants with coeliac disease were recruited using adverts on five different Facebook groups concerned with coeliac disease and the gluten-free diet. In addition, 84 members of these groups were contacted directly, and adverts were placed on a health journalist’s online blog, two other online gluten-free message boards, and on Cardiff University’s electronic noticeboard. Of those who responded, a total of 35 took part in the research.

Lay participants were also recruited using Cardiff University’s electronic noticeboard. This resulted in 19 respondents, of which 10 participated. Unlike the previous groups of participants, the inclusion criteria for lay people were designed to exclude participants with experience of dietary restrictions that might be similar to coeliac disease. Lay participants were thus asked to confirm that they:
did NOT have coeliac disease or any other dietary condition;
• did NOT have any close friends or family members with coeliac disease or a similar dietary condition;
• did NOT work with people with coeliac disease, or others with a dietary condition, as part of their job;
• did NOT follow, or have close friends or family members who follow, an unusually restrictive diet for other reasons (e.g. ethical, religious etc.).

Procedure
The Imitation Games began with each judge inventing a set of six questions, which were then sent to the researcher. Judges were free to ask questions on whatever topics they thought would distinguish between genuine patients and dietitians. The instructions were kept as open ended as possible, with advice limited to the types of questions that might be most useful. For example, Judges were told to avoid factual and other questions that could be answered by looking things up on the internet and to focus instead on the experiential aspects of coeliac disease. In this way, the questions should relate not only to the current concerns of people living with coeliac disease but also to the shared experiences that make them a social group and not simply a collection of individuals who happen to share a diagnostic label. As an example, the questions produced by two participants with coeliac disease are re-produced in Table 3.

Table 3: Sample Questions Asked by Coeliac Judge

| Question Number | IG8                                                                 | IG9                                                                                                                                 |
|-----------------|---------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------|
| 1               | You pop into a restaurant to check they can serve Gluten Free food before you book. Nobody seems very sure. What do you do? | When ordering in a restaurant, would you say that you had coeliac disease or would you ask for gluten/wheat free food?         |
| 2               | You’ve made some new friends who want you to come round for dinner. How do you deal with this? | Eating with friends and family – Just one of these won’t hurt will it? How would you reply?                                    |
| 3               | You are invited to a wedding and there is going to be a finger buffet. What do you do? | Invited to a meal with people who do not know you, to what extent would you quiz them on the food to be served?                |
| 4               | When explaining Coeliac disease, somebody comments about ‘faddy eaters’. How do you deal with this? | Shopping, do you read the labels to see if the contents had been changed since previously bought?                             |
| 5               | In a self service restaurant you notice the waitress using the same tongs to serve breaded food. What do you do? | Are you aware of any difficulties you might encounter on a trip to another country?                                           |
| 6               | Chatting to a friend they suggest that ‘a little bit of Gluten won’t hurt’. What do you say? | When having a medical prescription, or over the counter medication would you check for gluten/wheat?                         |

Next, the researcher removed any identifying information and forwarded the complete set of questions to one dietitian and one person with coeliac disease. When both participants had returned their answers to the researcher, these answers were anonymised, consolidated into a single document and sent to the judge who had created the questions. Finally, the judge informed the researcher which set of answers they thought came
from the person with coeliac disease. The judge also provided a reason for this judgement and an indication of how confident they were using the four point scale:

1. I have little or no idea who is who
2. I am more unsure than sure
3. I am more sure than unsure
4. I am pretty sure I know who is who

In total 12 Imitation Games were completed in this way, using 11 different judges. This is referred to as ‘Stage 1’.

In Stage 2, transcripts of these 12 Imitation Games were distributed to new judges. Each judge received three transcripts, which they were asked to read before returning their verdict, confidence level and reason, to the researcher. In total, this generated 107 new judgements from 42 judges.

For the Imitation Games with lay participants, the same procedure was used, but with the dietitians being replaced by the lay participants. In this case, 10 Stage 1 transcripts were created and, when distributed at Stage 2, these gave rise to 57 judgements from 23 judges.

Results
The substantive aim of the paper is to investigate the extent to which dietitians have interactional expertise in the embodied experience of living with coeliac disease. In order to establish this, we now analyse the performance of dietitians and compare this with that of lay people.

Chance Condition: Dietitians as Pretenders
The outcome of the 12 Stage 1 Imitation Games is summarised in Table 4, which also shows how many times each transcript was used in Stage 2.

Table 4: Number of Imitation Games Conducted

| Dietitian | Imitation Game Transcript ID | Original Judge’s Verdict | Number of Stage 2 Games |
|-----------|-----------------------------|--------------------------|------------------------|
| A         | IG1                         | Wrong                    | 6                      |
| B         | IG2                         | Right                    | 12                     |
| C         | IG3                         | Right                    | 2                      |
| D         | IG4                         | Wrong                    | 12                     |
| E         | IG5                         | Don’t Know               | 9                      |
| F         | IG6                         | Right                    | 11                     |
| G         | IG7                         | Right                    | 11                     |
| H         | IG8                         | Right                    | 11                     |
| D         | IG9                         | Wrong                    | 7                      |
| I         | IG10                        | Don’t Know               | 10                     |
| J         | IG11                        | Right                    | 11                     |
| K         | IG12                        | Right                    | 5                      |
| **Total** |                            |                          | **107**                |
The distribution of judgements about these transcripts is shown in Table 5. The first two columns show the verdicts of the Judges at Stage One. As the sample size is small, these are not analysed further but simply used as a means to generate data at Stage 2.12

Table 5: Proportion of Right, Wrong and Don’t Know Transcripts

| Verdict of Stage One Judges | Distribution of Transcripts in Stage 2 | Verdicts of New Judges at Stage 2 |
|-----------------------------|----------------------------------------|-----------------------------------|
|                            | N | %  | N | %  | N | %  |
| Wrong Guess (i.e. Dietitian) | 3 | 25 | 25 | 23 | 37 | 35 |
| Don’t Know                  | 2 | 17 | 19 | 18 | 24 | 22 |
| Right Guess (i.e. Coeliac)  | 7 | 58 | 63 | 59 | 46 | 43 |
| Total                       | 12| 100| 107| 100| 107| 100|

Identification Ratio (IR) | 0.08

The next two columns in Table 5 show the distribution of transcripts that were returned by judges at Stage 2. As can be seen the proportions of Right, Wrong and Don’t Know transcripts at Stage 2 are very similar to those at Stage 1. This suggests that the distribution of verdicts at Stage 2, which is shown in the final two columns, has not been caused by, for example, ‘Right Guess’ transcripts being under-represented in the Stage 2 data.

If we now examine the data displayed in the final two columns more closely, we see that the proportion of right guesses and wrong guesses is quite similar (43% and 35%). This, in turn, is reflected in the IR for the 107 judgements, which is 0.08. This is clearly consistent with Chance condition runs in other topics (Collins et al, 2006; Collins and Evans, 2013) and supports the hypothesis that the dietitians will have the interactional expertise needed to reproduce the patient perspective.

Identify Condition: Lay People as Pretenders

Although the data show that dietitians can pass as people living with coeliac disease, it is possible that this success reflects more general social experiences such as cooking and preparing food or eating out in restaurants. To test this hypothesis we ran a second set of Imitation Games in which lay people attempted to pass as someone living with coeliac disease. This involved creating 10 new transcripts that were then used to generate 57 judgements, distributed as shown in Table 6.

Table 6: Proportion of Right, Wrong and Don’t Know Transcripts

| Verdict of Stage One Judges | Distribution of Transcripts in Stage 2 | Verdicts of New Judges at Stage 2 |
|-----------------------------|----------------------------------------|-----------------------------------|
|                            | N | %  | N | %  | N | %  |
| Wrong Guess (i.e. Dietitian) | 3 | 30 | 22 | 39 | 4 | 7  |
| Don’t Know                  | 0 | 0  | 0 | 0  | 13| 23 |
| Right Guess (i.e. Coeliac)  | 7 | 70 | 35 | 61 | 40| 70 |
| Total                       | 10| 100| 57| 100| 57| 100|

Identification Ratio (IR) | 0.63
In contrast to the dietitians, where the IR was 0.08, the IR for lay people is 0.63. This is a clear Identify Condition result with the judges having much more success in working out who was who. Again, this is consistent with the hypothesis that any success by the dietitians would be due to specialist interactional expertise gained through encounters with their patients and not as a result of everyday experience.

As with the chance condition, we also examined the extent to which this result might have been affected by the distribution of transcripts. In this case, the distribution of the transcripts returned at Stage 2 appeared to deviate from the distribution of results in the original 10 Imitation Games. As the difference takes the form of an increase in the proportion of wrong guesses it seems unlikely to affect the overall conclusion. To make certain, however, we ran a series of five exercises in which the difference between the Stage 1 and 2 distributions was eliminated by removing a random sample of 7 of the Stage 2 verdicts that were based on Stage 1 wrong guesses. The data were then re-analysed for each of these five sub-samples. As shown in Table 7, the resulting Identification Ratios were all very similar, both to each other and to the whole sample, suggesting that the distribution of returned dialogues had no significant effect on the outcome.

### Table 7 Checks for Distribution Effects on IR

|                  | Sample One | Sample Two | Sample Three | Sample Four | Sample Five |
|------------------|------------|------------|--------------|-------------|-------------|
| Wrong Guess (i.e. Dietitian) | 4          | 3          | 3            | 4           | 2           |
| Don’t Know       | 10         | 11         | 13           | 10          | 13          |
| Right Guess (i.e. Coeliac) | 36         | 36         | 34           | 36          | 35          |
| **Total**        | **50**     | **50**     | **50**       | **50**      | **50**      |
| Identification Ratio (IR) | **0.64**   | **0.66**   | **0.62**     | **0.64**    | **0.66**    |

### Differences Between Dietitians and Lay People

Having established the IR for each set of Imitation Games, we now turn to the question of their difference. The main results are summarised in Table 8.

### Table 8: Identify and Chance Condition Results

|                  | Identify (Lay) | Chance (Dietitian) |
|------------------|----------------|--------------------|
| Wrong Guess (i.e. Dietitian) | 4              | 37                 |
| Don’t Know       | 13             | 24                 |
| Right Guess (i.e. Coeliac) | 40             | 46                 |
| **Total**        | **57**         | **107**            |
| Identification Ratio (IR) | **0.63**       | **0.08**           |

Using the Monte Carlo simulation reveals that the probability of this difference emerging by chance is less than 0.0001\(^{13}\). In other words, there is a clear difference between the two conditions. This provides further
support for the hypothesis that, as a result of their professional experience, dietitians will develop the interactional expertise needed to pass as ‘patients’ in an Imitation Game.

**Variations in Dietitians’ Expertise**

In addition to the quantitative data needed to calculate the IR, the Imitation Game also generates more qualitative data in the form of the questions, answers and reasons provided by participants. We now turn to a more exploratory analysis of this data for the Imitation Games with dietitians. To do this, we used the Stage 2 judgements to calculate the Identification Ratios for each of the 12 original transcripts. We then grouped the dietitians into three categories, defined as follows:

- **No Interactional Expertise**: IR between 0.60 and 1.00, so easily identified by judges
- **Some Interactional Expertise**: IR between 0.20 and 0.59, so sometimes identified by judges
- **High Interactional Expertise**: IR less than 0.19, performance comparable to that of contributory experts

The Identification Ratio for each transcript, and corresponding level of Interactional Expertise for each dietitian, is shown in Table 9.

**Table 9: Identification Ratios and Levels of Interactional Expertise**

| Imitation Game ID Number | Number of Imitation Games | Identification Ratio | Level of Interactional Expertise |
|--------------------------|---------------------------|----------------------|----------------------------------|
| IG9                      | 7                         | -0.63                | High                             |
| IG4                      | 12                        | -0.54                | High                             |
| IG5                      | 9                         | -0.20                | High                             |
| IG10                     | 10                        | -0.18                | High                             |
| IG2                      | 12                        | -0.15                | High                             |
| IG1                      | 6                         | 0.00                 | High                             |
| IG11                     | 11                        | 0.08                 | High                             |
| IG12                     | 5                         | 0.33                 | Some                             |
| IG6                      | 11                        | 0.58                 | Some                             |
| IG7                      | 11                        | 0.58                 | Some                             |
| IG8                      | 11                        | 0.92                 | None                             |
| IG3                      | 2                         | 1.00                 | None                             |

As can be seen, performance varied considerably but, although some demographic data was collected, there was no clear relationship between length of career or number of clients seen and the level of Interactional Expertise. This suggests that how experience is used matters more than its absolute quantity, though further research is needed to discover what factors help develop interactional expertise most effectively.

**Dietitians’ Successes**

We then examined whether dietitians with higher levels of expertise were more likely to provide plausible or convincing answers. This was done by coding the content of the reasons given by judges into one of nine categories (e.g. ‘eating out’, ‘explaining/disclosing illness’ etc) and selecting those cases where the judge explicitly acknowledged that the dietitian had provided a ‘good’ answer. These answers were then grouped
by the dietitian’s level of interactional expertise and the number of correct answers per level standardised by dividing the total number of ‘successes’ by the number of dietitians in each level. These results are summarised in Figure 2.

Figure 2: Number of Acknowledged Successes per Dietitian per Category

Perhaps unsurprisingly given the nature of the game, which encourages judges to look for mistakes, dietitians achieved acknowledged successes in only four of the nine topics covered by the questions asked. This does not mean that dietitians only had interactional expertise in these areas as it is possible that some categories (e.g. ‘Language’ or ‘Technical Facts’) were only mentioned by judges if they spotted an error. In addition, the distribution of successes (and errors) is influenced by the questions that were asked. Judges asked a large number of questions about different aspects of eating out, so it is not surprising that a large proportion of successes (and failures) are found in this category.

Nevertheless, there are some tentative conclusions that can be drawn from Figure 2. For example, it appears that the relationship between the dietitian’s level of interactional expertise and their success in different topics is transitive or cumulative in nature. Dietitians who were ultimately ranked as having no interactional expertise failed to score any explicit successes; dietitians with at least some interactional expertise scored ‘successes’ in response to questions about ‘Explaining/Disclosing Illness’ and ‘Eating Out’; whilst dietitians with the highest levels of interactional expertise scored successes in these topics plus two additional categories relating to food selection, and emotions. These latter topics, which included issues such as diagnosis, coming to terms with no longer being able to eat certain foods, and the reactions of other people, suggest that interactional expertise can include ‘emotional’ as well as ‘technical’ or ‘theoretical’ knowledge.
Dietitians’ Mistakes

The transcripts were also examined for instances of acknowledged mistakes. These are summarised in Figure 3, which was produced using the same procedures as for Figure 2.

**Figure 3: Number of Errors Made per Dietitian per Category**

Unlike successes, errors were dispersed fairly evenly across the categories, with the ‘Eating Out’ category once more appearing somewhat disproportionate due to the number of questions in this category. It is also clear that the transitivity which characterised the successes of dietitians is not present, though it does appear that dietitians with the highest levels of interactional expertise made the fewest mistakes. On the other hand, dietitians who demonstrate some interactional expertise appear in all but one category, whereas dietitians with no interactional expertise appear in only five out of nine categories. Similarly, there is one category – ‘Food Selection’ – that includes only dietitians who demonstrated the highest levels of interactional expertise.

Perhaps a more useful way of analysing the mistakes is to examine the kinds of errors that were made. In general, the mistakes made by judges fell into one of three categories: an overly risky decision about what to eat; making too much fuss when confronted with something they could not eat; or providing an overly medical or ‘textbook’ answer. A typical example of the first kind came in IG5, where participants were asked how they would respond if offered gluten-free bread at a café, and were asked whether they would like it toasted. The dietitian responded:

‘Yeah, that will be great. I love toast and it’s good to know they have gluten free bread’.
Many judges (IG14, IG25, IG41, IG52, IG55, IG111) pointed out that this answer ignores the risk of cross-contamination (i.e. gluten-free bread could become contaminated by using the same toaster as gluten-containing bread). Similar mistakes were seen in questions relating to other kinds of food selection, such as being given a standard meal on an aeroplane or being served ice-cream with a wafer in a restaurant.

An example of the second kind occurred in IG8, where the dietitian said they would speak to the manager if their food was served with the same tongs as gluten-containing food. Many judges commented that they would not want to ‘make such a big deal’ (IG20, IG34, IG38, IG81). Similar concerns were also expressed about the demands dietitians would make of potential hosts in order to ensure gluten-free food, with many judges suggesting that they would simply eat before attending buffets or similar meals.

It is worth noting that there was some variation between judges, particularly in relation to gluten consumption. Whilst many judges said they would send contaminated restaurant food back, some did acknowledge that the sensitivity to gluten varies so that, for at least some people, the apparently cavalier response of the dietitian could be plausible even if it was not what the judge would do themselves:

‘I would never risk the fact that I could eat something that wasn’t safe so only if I was sure would I do this. (not everyone thinks the same way as that though)’ (IG95).

That said, however, these views were comparatively rare and the majority of judges interpreted the dietitians as making a social misjudgement that underestimated the seriousness or difficulty of the situation they were being asked to imagine.

The third kind of mistake occurred when the dietitian’s answer was seen as overly concerned with medical aspects of the condition or as omitting the emotional and experiential aspects. For example, in IG7, when asked about what lifestyle changes they had made since diagnosis, the dietitian mentioned, as part of their answer, that they try to eat more dairy products to ensure they have enough calcium in their diet. Many judges felt that this answer lacked any practical understanding, with one judge commenting:

‘The answers from [the dietitian] are very text book and just sound like someone that has not lived with the condition but has read up on how to advise people with it’ (IG19).

It appears, therefore, that in some instances, dietitians answered the question from the ‘wrong’ perspective (i.e. the clinician’s perspective of disease rather than the patient’s perspective of illness). In other words, although there is some evidence to support the claim that interactional expertise in the embodied experiences of chronic illness is possible, its acquisition is not something that can be taken for granted. This, in turn, is something that has consequences for healthcare professionals as it may discourage patients from following treatment advice and/or continuing with dietetic appointments.

**Conclusions**

This paper has used a new research method to explore the extent to which medical practitioners are able to appreciate the lived experience of their patients. Using the Imitation Game we have shown that dietitians can develop the interactional expertise needed to ‘pass’ as patients living with coeliac disease. We have also shown that this ability cannot be attributed to the more ubiquitous experiences of everyday life. In relation to current concerns with patient-centred practice, the research shows that it is possible for medical professionals to complement their professional training with an understanding of their patients’ life-worlds.
Clearly, however, there is still much to do. The research showed that the performance of dietitians varied widely. By analysing the successes and failures, and replicating the study with a larger sample, it would be possible to develop a more robust understanding of how well dietitians understand different aspects of living with coeliac disease. Over time, it is even possible that these methods could be incorporated into professional training, allowing dietitians to explore their own knowledge and identify those areas where their expertise could be developed. For example, the Imitation Game could be used to support the growing number of medical education initiatives that utilise interaction between students and the chronically ill to promote patient-centred practice. If Imitation Games were used at the beginning and end of the programme, the Identification Ratios would then provide a measure of the student’s individual development and the program’s success.

In addition, there is no reason why the research has to be limited to dietitians. Not only can the educational use described above be generalised to other settings, the technique can be used to research the relationship between patients and any of the other medical professionals who treat them. By comparing the success of different professional groups, such as doctors, nurses, physiotherapists and so on, it would be possible to identify those domains in which patients were indeed enjoying the ‘new medical conversation’ and those in which more traditional roles remained entrenched. Over the longer term, the data gathered in this way could even be used to explore the extent to which increased understanding of the patient perspective by healthcare professionals is reflected in the patient experience and their health outcomes.

Finally, the method can also be used in non-medical settings. As noted above, the Imitation Game is currently being used to research social understandings of religion, sexuality, and ethnicity across Europe. Other potential topics include political issues where expert and lay knowledges are in conflict and where it is important that those in authority are seen to understand the issue from the perspective of those who will be affected. Using the Imitation Game would provide a new way of exploring the extent to which citizen concerns are fully appreciated by policy-makers and, perhaps, of making policy-makers more aware of the gaps in their own knowledge. There is, therefore, a significant programme of comparative research to be undertaken. In this paper, we have set out the method by which it can be undertaken and hope that others will attempt to replicate and extend its results.

Bibliography

Arksey, H. (1998) *RSI and the Experts*, London: UCL Press.

BDA – British Dietetic Association. (nd) The British Dietetic Association [Online]. Available at: http://www.bda.uk.com [Accessed: 26 June 2012].

Collins, H.M. (2004) ‘Interactional Expertise as a Third Kind of Knowledge’, *Phenomenology and the Cognitive Sciences*, 3(2) 125-143.

Collins, H.M. (2011) ‘Language and Practice’, *Social Studies of Science*, 41(2): 271-300.

Collins, H.M. and Evans, R. (2002) ‘The Third Wave of Science Studies: Studies of Expertise and Experience’, *Social Studies of Science*, 32(2): 235-296.

Collins, H.M. and Evans, R. (2007) *Rethinking Expertise*. Chicago: University of Chicago Press.
Collins, H.M. and Evans, R. (2013) ‘Quantifying the Tacit: The Imitation Game and Social Fluency.’ Sociology (published as Online First article on 25 February 2013, doi: 10.1177/0038038512455735)

Collins, H.M. Evans, R., Ribeiro, R. and Hall, M. (2006) ‘Experiments with Interactional Expertise,’ Studies In History and Philosophy of Science, 37(4): 656-674.

Epstein, S. (1996) Impure Science: AIDS, Activism, and the Politics of Knowledge, Berkeley: University of California Press.

Epstein, Ronald M. (1999) ‘Mindful Practice’, Journal of the American Medical Association, 282(9): 833-839.

Evans, R. and Collins, H.M. (2010) ‘Interactional Expertise and the Imitation Game’, in Gorman, M. (ed.) Trading Zones and Interactional Expertise, Chicago: MIT Press. pp. 53-70.

Freund, Peter ES, McGuire, Meredith B and Podhurst, Linda S (2003) Health, Illness and the Social Body: A Critical Sociology (4th Edition), Upper Saddle River, NJ: Prentice Hall.

Giles, Jim, 2006 ‘Sociologist Fools Physics Judges’, Nature 442:8

Grin, J., van de Graaf, H. and Hoppe, R. (1997) Technology Assessment Through Interaction: A Guide, The Hague: Rathenau Institute.

Hallert C, Granno C, Hulten S, Midhagen G, Strom M, Svensson H, Valdimarsson T. (2002) ‘Living with Coeliac Disease: Controlled Study of the Burden of Illness’, Scandinavian Journal of Gastroenterology, 37: 39–42.

House of Lords (2000) Science and Society: Science and Technology Select Committee, Third Report, London: HMSO.

Irwin, A. (1995) Citizen Science: A Study of People, Expertise and Sustainable Development, London and New York: Routledge.

Layton, D., Jenkins, E., Macgill, S. and Davey, A. (1993), East Yorkshire: Studies in Education Ltd.

Martin, Graham P. (2008) “Ordinary People Only’: Knowledge, Representativeness, and the Publics of Public Participation in Healthcare’, Sociology of Health & Illness, 30(1): 33-54.

Mazur, D. (2002) The New Medical Conversation, Boulder; CT: Rowman and Littlefield

NHS – National Health Service. (2010) Coeliac Disease [Online]. Available at: http://www.nhs.uk/conditions/Coeliac-disease/Pages/Introduction.aspx [Accessed: 12 January 2011].
NICE – National Institute for Health and Medical Excellence. (2009) Coeliac Disease: Recognition and Assessment of Coeliac Disease (medical guideline 86) [Online]. Available at: http://www.nice.org.uk/nicemedia/pdf/CG86FullGuideline.pdf [Accessed: 06 August 2009].

Popay, J. and Williams, G. (1996) ‘Public Health Research and Lay Knowledge’, Social Science & Medicine, 42(5): 759-768

Prior, L. (2003) ‘Belief, Knowledge and Expertise: The Emergence of the Lay Expert in Medical Sociology’, Sociology of Health and Illness, 25(3): 41-57.

Reid, Margaret (1976) The Development of Medical Sociology in Britain, Discussion Papers in Social Research No 13, University of Glasgow. Available online at: http://www.britsoc.co.uk/medsoc/Medsoc+Origin.htm [Accessed 7 Jan 2011]

Rip, Arie, Misa, Thomas J. and Schot, John (Eds) (1995) Managing Technology in Society: The Approach of Constructive Technology Assessment, London ; New York: Pinter Publishers.

Sverker, A., Hensing, G. and Hallert, C. (2005) ‘Controlled by Food’ – Lived Experiences of Coeliac Disease’, Journal of Human Nutrition and Dietetics, 18(3): 171-180.

Sverker, A., Östlund, G., Hallert, C. and Hensing, G. (2007) ‘Sharing Life with a Gluten-Intolerant Person: The Perspective of Close Relatives’, Journal of Human Nutrition and Dietetics, 20: 412-22.

Turing, A.M. (1950) ‘Computing Machinery and Intelligence’, Mind, 49: 433-460.

Whitaker, J.K.H., West, J., Holmes, G.K.T. and Logan, R.F.A. (2009) ‘Patient Perceptions of the Burden on Coeliac Disease and its Treatment in the UK’, Alimentary Pharmacology & Therapeutics, 29(10): 1131-1136.

Wilsdon, J. and Willis, R. (2004) See-through Science: Why Public Engagement Needs to Move Upstream, London: Demos [in association with the Green Alliance, Environment Agency and RSA]. Available on line at http://www.demos.co.uk/catalogue/paddlingupstream/ [accessed 12 June 2005].

Wynne, B. (1992) ‘Misunderstood Misunderstanding: Social Identities and Public Uptake of Science’, Public Understanding of Science, 1(3): 281-304.

Ziebland, S. (2004) ‘The Importance of Being Expert’, Social Science and Medicine, 59: 1783-1793.

Biography

Helen Crocker is a DPhil student in Public Health within the Health Services Research Unit at the University of Oxford. Helen joined the unit in 2009 as a Research Officer, where she was involved in a study piloting the use of patient-reported outcome measures (PROMs) for long-term conditions in primary care. Her DPhil research focuses on identifying and assessing patients’ experiences of healthcare services accessed to diagnose and manage the gastrointestinal condition coeliac disease. In addition, she plans to identify how the condition
impacts on health-related quality of life (HRQOL), and to explore relationships between patients’ experiences and HRQOL.

Robert Evans is a Reader in sociology at the Cardiff School of Social Sciences, where his research focuses on research methods, the sociology of science and technology and the nature of expertise. Previous projects have examined economic forecasting, GIS models for sustainable development and genetics. His current research is devoted to developing the ideas set out in the ‘Third Wave of Science Studies’ paper co-authored with Harry Collins (Social Studies of Science, 2002) and, in particular, the use of imitation game method for comparative research.

1 We would like to thank to Carole Sutton and the two anonymous referees for their constructive comments on an earlier draft of this manuscript. We are also grateful to Harry Collins, the members of the KES group at Cardiff School of Social Sciences and the participants at the SEESHOP workshops who have all listened to us present this work and provided a wealth of useful ideas and feedback.

2 It is possible for social scientists to develop contributory expertise in their field of study. The point here is simply that it is not necessary.

3 A more detailed description is available in Evans and Collins, 2010.

4 The further development of the method is now supported by a €2.25M Advanced Research Grant from European Research Council (ERC Advanced Grant 269463). The grant provides funding for comparative and cross-national Imitation Game research on a number of different topics. The aim is to explore the robustness of the method and to further develop the research protocols needed to enable the method to be used by others.

5 Arithmetically this is the same as coding wrong guesses as ‘-1’, don’t knows as zero, right guesses as ‘+1’ and then calculating the mean of the resulting distribution. Although the result is the same, we think presenting the data categorically has some advantages (e.g. it makes it easier to see the proportion of don’t know responses).

6 The Monte Carlo method involves simulating a sampling distribution by creating 10,000 pairs of samples based on weights derived from the two original samples. Counting the number of sample pairs in which the difference exceeds that observed in the data provides an estimate of the probability that two samples with the observed difference could be created purely by chance. Other statistical tests that could be used include the Chi Square test for data displayed in Table 1 or the t-test for the difference between the means if the data is coded as described in note 7. Both give a similarly clear result. We are grateful to Professor Bernard Silverman and the late Professor Tony Coxon for affirming our method of statistical analysis but the responsibility rests with us.

7 This is a variation on the method used in earlier research (e.g. Collins et al, 2006) where bespoke software relayed questions and answers between participants in real time.

8 The instructions provided to judges explained this in more detail and also included a number of sample questions to illustrate the ideas. After the first few Imitation Games it became apparent that judges were relying too heavily on these sample questions and the information sheet was modified to include fewer examples. Examples of recent instruction sheets and other information can be found at http://www.cardiff.ac.uk/sosci/expertise.

9 For more on this ‘proxy stranger’ method, see Collins and Evans, forthcoming.

10 The two sets of questions are chosen as they represent transcripts in which the Pretenders (i.e. the Dietitians) were least successful (IG8) and most successful (IG9). See Table 9 for more detail.

11 The ‘stages’ are similar to the interactive ‘Phase 1’ and non-interactive ‘Phase 2’ described in Collins et al 2006.

12 A similar procedure is now becoming standard practice in the much larger research project referred to in note 4.

13 As with the earlier example, using alternative statistical tests gives equally clear results.

14 The categories were developed inductively from the transcripts.