The Schum Challenge: 7th June 2005

January 13, 2006
Contents

1 Introduction ........................................ 3
   1.1 Ways To Read This Document ................ 3
   1.2 Modifications to the Email Exchange ........ 3

2 Setting Up the Challenge ........................ 5
   2.1 Negotiating the Structure .................... 5
   2.2 Establishing the Programme .................. 11

3 After the Challenge: Feedback ................... 18
   3.1 First Thoughts: Positive .................... 18
   3.2 Second Thoughts: A Note of Caution ......... 22
   3.3 Attempts at Resolution ....................... 24
   3.4 Problems Reiterated ........................ 26
   3.5 Problem Compounded ......................... 30
   3.6 Problem Rejected ............................ 32
   3.7 Framing the Problem More Widely .......... 39
   3.8 The Final Word: An Outsider’s Synthesis ... 46
   3.9 Later Discussion .............................. 50

4 Schum’s response ................................ 64
   4.1 Introductory Comments ....................... 64
   4.2 Substance-Blindness .......................... 65
   4.3 When Substance is Never Blind ............... 67
   4.4 Epilepsy Services in Ferndale ............... 69
   4.5 Academics Teach Better if They are Also Research Active ... 71
   4.6 A Historian’s Perspectives on Evidence ... 73
   4.7 Ethnicity Issues in the Use of Evidence .... 74
   4.8 In Conclusion ................................ 76
   4.9 wrote with title Re: [Evidence] The Schum Challenge: an inter-disciplinary study .... 77
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>Appendix: Email Conversations</td>
<td>79</td>
</tr>
<tr>
<td>5.1</td>
<td>Formatting Email in Print</td>
<td>79</td>
</tr>
<tr>
<td>6</td>
<td>Appendix: Evidence Personnel</td>
<td>82</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Ways To Read This Document

This document is a hyperlinked pdf, which means that it will always print with identical pagination on any machine: it can therefore be referenced reliably during discussion by page and paragraph number. It contains hyperlinks which enable the reader to navigate the cross references within the text and in the index. It is also designed to be usable as a printed document. Cross-references refer to the email in question by number and page, as do the page numbers in the index entries. In practice, it may be easiest to use both at once. Depending on your pdf viewer, the links may not be visible: recent versions of Acrobat Reader should show them, albeit in a rather ugly way. It is, however, intended that the printed version should be sufficient: the hyperlinks will not appear as any different from normal text in the printed version.

1.2 Modifications to the Email Exchange

Some of the material quoted in the email exchange has been excised (where clearly superfluous by dint of repetition) and the emails have been reordered to show the current response first: you will see (indented) what the person was replying to. This inevitably leads to some repetition of material, despite the editing. Quotation levels are indicated by indentation, thus:

original message

Response 1 to original message
1.2. MODIFICATIONS TO THE EMAIL EXCHANGE

Response 2 to response 1 (pertinent message)

Please note that some conversations moved away from the group (either deliberately or accidentally) and at other times, people quoted material that had apparently been sent to them privately back to the wider list. In such cases, the quoted material is included in full. For those interested in the messy details of how to manage such a collection of disparate messages and other special pleading, see page 79.

Misspellings and syntax have not been corrected, though a fewemailer-induced errors have been checked and altered to reflect the intended outcome. Some long URLs have been changed to ‘tinyurl’ entries which avoid excessively long lines (but still point to the same place on the web).
Chapter 2

Setting Up the Challenge

2.1 Negotiating the Structure

2.1.1 Philip Dawid wrote with title Schum challenge

addressing Evidence Personne\footnote{1For a list of this sub-group, see page 82}(Thursday, April 7, 2005 1:15:36 pm)

Dear all,

Arising out of a recent “Narrative Study Circle”, which discussed David Schum’s book “The Evidential Foundations of Probabilistic Reasoning”, it has been suggested that, during his next visit here, planned for the week beginning 6 June, we hold a (say) half-day open meeting at which he can be invited to show how his conception of a unified “Science of Evidence” can indeed be made to work across a variety of fields. He has (like the Queen) graciously agreed to this suggestion.

The idea would be to line up say 4 speakers, from the more sceptical wing of the Evidence Programme, and have each present a problem relating to using evidence in their field that they feel might pose a suitable challenge to this thesis. Ideally this would be of a fairly specific and concrete nature, although in some way representative of the generality of such problems.

So I am hereby soliciting such contributions. Please reply quickly, if possible, with your vaguely formed ideas, but be prepared to flesh these out fairly soon. If we can get written material to Schum in advance that would obviously be a good thing. He will be given an opportunity to respond at the meeting, but I would expect (hope) that the problems will be meaty enough to require further sustained effort over a period of time.

I am sending this message only to those significantly involved with the Evidence Programme, but please feel free to draw it to the attention of any colleagues.
2.1. NEGOTIATING THE STRUCTURE

you feel might be willing to get actively involved in such an enterprise. Initial soundings suggest a suitable meeting date could be Tuesday 7 June, p.m. Please let both me and Victoria know if you want to be involved but would prefer some other timing.

Philip

2.1.2 Tony Gardner-Medwin wrote with title Re: Schum challenge

addressing Philip Dawid and the Evidence Personnel (Thursday, April 14, 2005 10:50:16 am)

At 17:31 12/04/2005, Philip wrote:
Dear all,
We have had a good response to this, with offers from:
1. Trisha Greenhalgh/Janet McDonnell/Emma Byrne
2. Jason Davies
3. Tony Gardner-Medwin
4. Stephen Rowland
Trisha et al. even offered a double session, but that seems a bit OTT to me (and I think to poor Dave Schum)!
So far as I can assess, the only day that you could all manage is Tuesday 7 June – and Stephen is not free between 12 and 1:30. But we can take a lunch break then...
Can all involved please quickly confirm your availability then? Also indicate how long is the *minimum* time you feel it would take to present the issues clearly.
We now want, at the very minimum, titles and brief abstracts for the presentations. But the more detail can be provided in advance the better. And the earlier the better...
Dave: Please circulate the above list with any suggestions, requests or demands of your own.
Best wishes,
Philip

Dear Philip,
Is this OK? 7th is fine. I can go on at any length!
Tony
Race: taboo or not taboo? Its different status as evidence in medical and legal decisions.
2.1. NEGOTIATING THE STRUCTURE

It can be right, when deciding on a patient’s diagnosis and treatment, to take account of statistics correlating race and disease. It is generally wrong, in a criminal trial, to use statistical correlations between race and crime as evidence. I will explore the difference, considering whether it is a matter of ethics, prejudice, utility, relevance, or a difference in the nature of medical and legal decisions. Background is available at http://tinyurl.com/9gedo.

Tony GM

2.1.3 Stephen Rowland wrote with title Re: Schum challenge

addressing Philip Dawid and the Evidence Personnel (Friday, April 15, 2005 9:31:09 am)

At 17:31 12/04/2005 +0100, Philip Dawid wrote:
Dear all,
We have had a good response to this, with offers from:
1. Trisha Greenhalgh/Janet McDonnell/Emma Byrne
2. Jason Davies
3. Tony Gardner-Medwin
4. Stephen Rowland
Trisha et al. even offered a double session, but that seems a bit OTT to me (and I think to poor Dave Schum)!
So far as I can assess, the only day that you could all manage is Tuesday 7 June – and Stephen is not free between 12 and 1:30. But we can take a lunch break then...

Dear Philip and All
I just want to be clear about the purpose of this meeting.
I see it as a valuable opportunity to invite Schum to apply his framework for tackling questions of evidence in relation to our problems. I don’t see it as an opportunity for us to present papers on how we deal with our evidential questions. We have the opportunity to do that at other times.
I think it might be useful to have a structure which ensures that this happens and that we don’t end up focussing on our presenters’ ideas rather than Schum’s.
No doubt each of us could take an hour to outline our problem. But if we have a reasonable understanding of what we are talking about we should be able to do it in ten minutes.
I’m not sure what time frame you had in mind, but if we had four problems and an hour for each, then it might be reasonable to use the time as follows:
2.1. NEGOTIATING THE STRUCTURE

10 minute presentation of the problem
5 min opportunity for Schum to respond ‘off the cuff’ if he wishes
25 min discussion of the problem by all (including Schum)
10 min response from Schum in the light of the discussion
5 min response by the presenter on the value of Schum’s contribution
leaving 5 min to get ready for next session.

I would recommend that timing is kept to closely.
I wonder what others think about this.
Best wishes
Stephen

2.1.4 Tony Gardner-Medwin wrote with title Re: Schum challenge

addressing Stephen Rowland and the Evidence Personnel (Friday, April 15, 2005 9:52:55 am)

In response to Stephen, I think an hour much too long to devote to each of
4 issues like this. Maybe 2 hours in all for 4 topics including discussion, then a
chance after lunch for David Schum to bring together thoughts in a synthesis and
response.
Tony

2.1.5 Philip Dawid wrote with title Re: Schum challenge

addressing Stephen Rowland and Evidence Personnel (Friday, April 15, 2005 11:08:35 am)

On Fri, 15 Apr 2005, Stephen Rowland wrote:
Dear Philip and All
I just want to be clear about the purpose of this meet-
ing.
I see it as a valuable opportunity to invite Schum to
apply his framework for tackling questions of evidence in
relation to our problems. I don’t see it as an opportunity
for us to present papers on how we deal with our eviden-
tial questions. We have the opportunity to do that at other
times.
2.1. NEGOTIATING THE STRUCTURE

On Fri, 15 Apr 2005, Tony Gardner-Medwin wrote:
In response to Stephen, I think an hour much too long to devote to each of 4 issues like this. Maybe 2 hours in all for 4 topics including discussion, then a chance after lunch for David Schum to bring together thoughts in a synthesis and response. Tony

Dear All,
I concur with Stephen’s description of the aim of the exercise, and with both his and Tony’s concerns over length. But it might not be appropriate to impose a one-size-fits-all timing constraint. If, as I propose, presenters focus on one specific case of evidential significance arising in their discipline, rather than generalities, it may take a little longer that 10 minutes to make the background and the evidential issues intelligible. But this should be kept to the absolute minimum – the shorter and punchier the better!

In the light of this, can presenters please advise me (again!) of the time they think they would require?

For those of you who have not yet done so, I very much recommend looking at the book “Evidence and Inference in History and Law” (eds William Twining and Iain Hampsher-Monk, Northwestern University Press, 2003) – in particular, as some sort of model for what I have in mind, the exchange between Mark Geller and Terry Anderson. Anderson applied Wigmorean analysis to a very specific problem posed by Geller: when did the ability to read cuneiform script die out? This analysis was developed over a period of time and involved very substantial work – it is not something that could be replicated off-the-cuff in a meeting like ours. But I would encourage a similar tight focus for our own “challenges”, and hope that at least some of these will lead to such longer term investigation.

Philip

2.1.6 Janet McDonnell wrote with title Re: Schum challenge

addressing Philip Dawid and the Evidence Personnel (Friday, April 15, 2005 2:33:38 pm)

Hello All, I think the one-hour model from Stephen would be right for us – we may need a little more than 10 minutes to set out the case but we can try to satisfy the spirit of Stephen’s suggestion and avoid giving a seminar. I think we might need 15 mins and then just be ready to supply more detail/clarification as the need arises during the session.

Janet
2.1. NEGOTIATING THE STRUCTURE

2.1.7  Tony Gardner-Medwin wrote with title Re: Schum challenge

addressing Philip Dawid and the Evidence Personnel (Friday, April 15, 2005 5:13:09 pm)

Dear Philip, What I proposed for discussion is certainly not comparable to the sorts of case studies in the Dialogues of Twining and Hampsher-Monk. So I think I should pull out. My proposal does I think challenge a science of evidence in a really fundamental way by contrasting different fields: the importance of defining, clarifying and avoiding confusion between the different hypotheses one might address in a problem, and to which evidence should be relevant. But this is a different sort of problem from what you want, and it sounds as if this occasion would not be the right one on which to try to present it.

Yrs, Tony GM

2.1.8  Philip Dawid wrote with title Re: Schum challenge - draft programme

in response to Jason Davies (private exchange) (Tuesday, April 26, 2005 4:02:57 pm)

On Tue, 26 Apr 2005, Jason Davies wrote:
I haven’t had a chance to look at it (Schum’s book on Evidence) again though we’ve ordered a copy for CALT which I hope to pick up very soon. How long before you want a title? it will get priority as soon as I can look at the book properly.

you’re not obliged to look at anything... just to come up with a specific problem that you feel may raise some interesting and challenging evidential points.

schum’s main thesis is that there is (or could be) a context/substance-blind approach to evidence analysis, so a good challenge would be one that says “in my discipline the problems are really rather special”

phil
2.2 Establishing the Programme

2.2.1 Philip Dawid wrote with title URGENT: Schum challenge

addressing Terry Anderson and Evidence Personnel (Friday, May 6, 2005 1:42:18 pm)

Dear all,

I am attaching a revised programme for the “Schum challenge” event on 7 June. Every one I have heard from agrees that we should go ahead with this, even if Dave Schum himself can not be there (though I remain hopeful that he can be).

I will leave it to speakers to decide on the balance between presentation and discussion within the overall 45 minute slot.

Please will ALL INVOLVED let me know BY NEXT MONDAY 9 MAY whether or not you are happy with this, or have any changes to make.

(JASON: Please provide a title!)

Philip

Evidence Workshop: Challenges to an Integrated Theory of Evidence
Tuesday 7 June 2005, Galton Lecture Theatre (Room 115, 1-19 Torrington Place)

Draft Programme (6 May 2005)
9:30 Arrival, coffee
9:45 Introductory remarks Philip Dawid
10:00 Challenge 1. The relationship between research and teaching Stephen Rowland
10:45 Coffee
11:15 Challenge 2. Race: Taboo or not taboo? Tony Gardner-Medwin
12:00 Challenge 3. TBA Jason Davies
12:45 Lunch
2:00 Challenge 4. “Are you coming from where I’m coming from?” Competing constructions of evidence in health care policy making. Trisha Greenhalgh, Janet McDonnell, Emma Byrne.
2:45 Case Study: Wigmore meets the last wedge: Mark Geller, Terry Anderson
3:45 Tea
4:15 General discussion

2.2.2 Philip Dawid wrote with title [Evidence] Evidence Workshop: 7 June

addressing Evidence list (Monday, May 9, 2005 3:12:35 pm)

Please see below for details of an important “Evidence” meeting.
2.2. ESTABLISHING THE PROGRAMME

Philip Dawid

Evidence Workshop: Challenges to an Integrated Theory of Evidence
Tuesday 7 June 2005 Galton Lecture Theatre (Room 115, 1-19 Torrington Place)

BACKGROUND

One of the principal motivations behind the “Evidence” programme at University College London is the possibility of identifying and developing an “integrated theory of evidence”, comprising fundamental features of the nature and behaviour of evidential reasoning that apply across the very broad range of disciplines and types of evidence that are represented in the overall programme. Arguments in favour of such an integrated theory, and some account of its possible nature and content, have been forcefully presented by David Schum in many articles over a long period, and brought together in his book “The Evidential Foundations of Probabilistic Reasoning” (Wiley, 1994).

Unsurprisingly, the “Schum thesis” has met with a degree of scepticism. This one-day workshop has been organised to give some of the sceptics an opportunity to present challenges to the Schum thesis, in the form of specific evidential problems that present themselves from within their own fields of study. There will also be a presentation by Mark Geller and Terry Anderson of their collaborative attempt to apply methods of organising legal evidence, as developed by Wigmore, to a specific historical question. This is described in the collection “Evidence and Inference in History and Law: Interdisciplinary Dialogues” (W L Twining and I Hampsher-Monk, eds), Northwestern University Press (2003).

It is very much hoped that David Schum will be able to participate, in spite of recent illness. In any case, others will be there to represent his point of view (as best they can), and there will be plenty of time for discussion of both specific and general issues.

This is an important event for the whole Evidence programme, and it is hoped that many participants and interested outsiders will be able to attend.

Refreshments and lunch will be provided for those who give advance notice of their attendance. Please inform Victoria Kingham before 31 May if you plan to be there.

Evidence Workshop: Challenges to an Integrated Theory of Evidence
Tuesday 7 June 2005, Galton Lecture Theatre (Room 115, 1-19 Torrington Place)

PROGRAMME

9:30 Arrival, coffee
9:45 Introductory remarks Philip Dawid Statistical Science, UCL
2.2. ESTABLISHING THE PROGRAMME

10:00 Challenge 1. The relationship between research and teaching: Stephen Rowland; Centre for the Advancement of Learning and Teaching, UCL
10:45 Coffee
11:15 Challenge 2. Race: Taboo or not taboo? Differences between legal and medical decisions. Tony Gardner-Medwin Physiology, UCL
12:00 Challenge 3. The failure of knowledge: the place of ignorance (an anthropological perspective). Jason Davies Centre for the Advancement of Learning and Teaching, UCL
12:45 Lunch
2:00 Challenge 4. “Are you coming from where I’m coming from?” Competing constructions of evidence in health care policy making. Trisha Greenhalgh, Primary Care and Population Science, UCL; Janet McDonnell; Emma Byrne. Greenwich University Business School
2:45 Case Study: Wigmore meets the last wedge: Mark Geller, Hebrew and Jewish Studies, UCL; Terry Anderson, Law, University of Miami
3:45 Tea
4:15 General discussion

2.2.3 Philip Dawid wrote with title Schum challenge

addressing Terry Anderson and Evidence Personnel (Friday, May 20, 2005 6:25:30 pm)

Dear Challengers,

It occurs to me that it might be a good idea for you to refresh your memory of the talks Dave Schum gave last summer. If you go to the website at https://www.evidencescience.org and click on Seminars in the events sections, you can then access (near bottom of list) the announcements of those talk. Attached to those for 27 May and 2 June 2004 you will find pdf files of the slides he used.

Hope this is helpful...

Philip

2.2.4 Philip Dawid wrote with title Schum challenge

addressing Tony Gardner-Medwin and Evidence Personnel (Sunday, May 29, 2005 11:13:30 am)

Dear Challengers,

You each have a 45 minute slot for your challenge, to include the presentation and some discussion. There is also a session set aside for further discussion at the end of the day.
2.2. ESTABLISHING THE PROGRAMME

I originally suggested the following indicative timings:
(a) Presentation of the problem: 15 minutes
(b) Initial response by Schum: 10 minutes
(c) General discussion: 15 minutes
(d) Response by presenter(s): 5 minutes

If you feel this is inappropriate for you, by all means make an alternative suggestion. I strongly suggest that you come armed with visual aids, handouts or whatever, to enable you to make your presentation in as compact yet informative a style as possible. I also strongly suggest that, so far as possible, you focus on particular examples we can get our teeth into, rather than generalities.

I am not sure at the moment of the extent to which Dave Schum will wish to contribute, but William Twining, Terry Anderson and I will be there to try and make some form of response. We all 4 need to get as much advance notice as possible of what you plan to present *and any materials you plan to use*.

I am looking forward to the event (though not without some trepidation). Very many thanks for agreeing to participate.

Philip

2.2.5 Philip Dawid wrote with title Challenge workshop 7 June

addressing Evidence challenge participants and Gianluca Baio (Sunday, June 5, 2005 3:07:19 pm)

Dear “Evidence Challenge” participants,

I am attaching as a Word file, and also appending below as plain text, the programme for Tuesday’s workshop.

Please see the website [www.evidencescience.com](http://www.evidencescience.com) for more information, maps, etc.

I look forward to seeing you there.

Philip Dawid

_Evidence Workshop Challenges to an Integrated Theory of Evidence_

Tuesday 7 June 2005, Galton Lecture Theatre (Room 115, 1-19 Torrington Place)

PROGRAMME
9:30 Arrival, coffee
2.2. ESTABLISHING THE PROGRAMME

10:15 Challenge 1 The relationship between research and teaching Stephen Rowland, Centre for the Advancement of Learning and Teaching, UCL
11:00 Coffee 11:15 Challenge 2 Race: Taboo or not taboo? Differences between legal and medical decisions. Tony Gardner-Medwin, Physiology, UCL
12:00 Challenge 3 The failure of knowledge: the place of ignorance (an anthropological perspective). Jason Davies, Centre for the Advancement of Learning and Teaching, UCL
12:45 Lunch
2:00 Challenge 4 “Are you coming from where I’m coming from?” Competing constructions of evidence in health care policy making. Trisha Greenhalgh, Primary Care and Population Science, UCL Janet McDonnell & Emma Byrne, Greenwich University Business School
2:45 Case Study: Wigmore meets the last wedge Mark Geller, Hebrew and Jewish Studies, UCL Terry Anderson, Law, University of Miami
3:45 Tea
4:15 General discussion
5:30 Close

The purpose of these challenges is to promote discussion, in the context of specific problems, of the possibility and scope of an “integrated theory of evidence” comprising fundamental features of the nature and behaviour of evidential reasoning that might apply across a very broad range of disciplines and types of evidence. Lead discussants will include David Schum, Philip Dawid, Terry Anderson and William Twining (Law, UCL).

2.2.6 Philip Dawid wrote with title “Challenge” session chairs and timings

addressing Nigel Harvey and others (see message for recipients) (Monday, June 6, 2005 2:46:52 pm)

Dear Challenge Chairmen (cc. Presenters/Respondents):

Many thanks for agreeing to chair a session in tomorrow’s “Evidence Challenge” workshop.

The designated respondents to the 4 challenges are Schum (DS), Twining (WT), Anderson (TA) and Dawid (PD). Any or all of these should be given opportunities to contribute, before inviting floor discussion.

The 2:45 session is not a challenge, but a presentation of a case study.

Appended below are some tentative suggestions for the division of the time on the various sessions. Following his stroke, Schum’s contributions are particularly
tentative, and nobody should feel offended if he has to skip out of some sessions.

I stress the “tentative” nature of all these timings: chairmen should be prepared for some flexibility – to some extent you will have to play it by ear. But it will be important to prevent the whole programme from running late. I appreciate it will be very hard to keep to such tight overall timings, but we simply can’t afford too much overshoot, so chairmen are requested to be (pleasantly) vicious.

I am looking forward to seeing you all tomorrow.

Best wishes,

Philip

SUGGESTED TIMINGS
SESSION 1. Chair: Nigel Harvey
9:45- PD 5 mins
10:15 DS 15 mins
WT 5 mins
TA 5 mins
10:15- Rowland 15 mins
11:00 PD 10 mins
WT 5 mins
TA 5 mins
Floor discussion 5 mins
Rowland response 5 mins

SESSION 2. Chair: Hasok Chang
11:15- Gardner-Medwin 15 mins
12:00 TA 10 mins
PD/DS 10 mins
Floor discussion 5 mins
G-M response 5 mins
12:00- Davies 15 mins
12:45 WT 10 mins
TA 5 mins
DS/PD 5 mins
Floor discussion 5 mins
Davies response 5 mins

SESSION 3. Chair: Mike Joffe
2:00- Greenhalgh et al. 15 mins
2.2. **ESTABLISHING THE PROGRAMME**

2:45 WT 10 mins  
TA 5 mins  
DA/PD 5 mins  
Floor discussion 5 mins  
Greenhalgh... response 5 mins

2:45- Geller/TA 30 mins  
3:45 DS/PD/WT 10 mins  
Floor discussion 10 mins  
Geller/TA response 10 mins

SESSION 4. Chair: Jim Smith  
4:15- General discussion 1 hour  
5:30 Rounding off  
APD/WT/TA/DS 15 mins
Chapter 3

After the Challenge: Feedback

3.1 First Thoughts: Positive

3.1.1 Philip Dawid wrote with title Challenge feedback

addressing Evidence challenge personnel (Wednesday, June 8, 2005 10:56:39 am)

Dear all,

My heartfelt thanks for your stimulating contributions to the “challenge” workshop yesterday.

I think now that it would be very helpful if each of you (including those in the “cc:” who were openly in the audience) were to make, and circulate to this list, some brief jottings of what (if anything) you yourself got out of the day, and your overall view of it. Can I urge you to do this while the memory is still fresh?

My own general view is that we made some small progress towards clarifying some ambiguities, resolving some misunderstandings, and developing the beginnings of a shared understanding of the scope and limitations of an “integrated science of evidence”. If I am right this is no mean achievement, and the day will have served its purpose very well.

Starting pretty much from rock bottom, I think we have now advanced to base camp. I should like to believe that we can move on from there, pulling together to make further progress onwards and upwards (though this may well be slow, and punctuated by diversions and detours to evade a range of obstacles and impediments.)

Perhaps now is the time to take up an earlier suggestion of mine that we should try and establish cross-disciplinary partnerships to explore the evidential problems of one area from the perspective of another. Stephen and I have agreed to do this in relation to his “teaching and research” challenge. Let me encourage you all to explore such possibilities seriously, and let us know what develops.
3.1. **FIRST THOUGHTS: POSITIVE**

Once again, many thanks for your part in what I think was a very worthwhile event.

Philip

3.1.2 **Stephen Rowland wrote with title Re: Challenge feedback**

addressing Philip Dawid and Evidence Personnel (Wednesday, June 8, 2005 2:05:31 pm)

Dear All

Thank you to Phil for arranging this day - and the evening - which I very much enjoyed (with the exception of the noisy air conditioner).

I agree with many of Phil’s comments below. I think the start we have made consists largely in beginning to sketch out some of the difficulties. While the day was appropriately structured around David Schum’s ideas, I think it is important that these are seen as only one of a number of important resources for the programme to draw upon.

I think some felt that Schum’s ideas had not really been challenged. I don’t share that view. I think the challenges had the effect of making it clear that the kind of evidential reasoning involved in his approach is of value, but does not help us with some of the problems in social enquiry. Specifically, it does not (nor is it, I believe, intended to) help us with the difficult issue of identifying and clarifying questions, concepts and theoretical resources. These are so often the difficulties - and most of the work - with social and policy related enquiry.

It may be that disciplinary differences hinge not so much around the acceptability of the Schum approach, but rather around the identification of questions, concepts, etc. which need to be framed before this approach can get a foot-hold. Certainly in my own research approach, which tends to be towards illuminative enquiry rather than hypothesis testing, the kind of step by step processes of evidential reasoning have only a limited place. In some forms of illuminative enquiry many would argue that it is unwise to formulate questions too precisely at an early stage and that the major purpose of such research is to frame, rather than answer, questions.

These difficulties, or limitations, in evidential reasoning raise for me the question of the scope of the programme. In my field of work I think it is important that the programme can raise a thoughtful critical response to the rhetoric of evidence-based practice/policy/etc that usually means no more than the most shallow forms of empiricist research in which there is virtually no thought given to the relationships between correlation and causation, for example. If it is to do this, it will need to cast its interests much wider than simply forms of evidential reasoning, to consider wider questions. This would inevitably lead us into methodological
3.1. FIRST THOUGHTS: POSITIVE

debates around empiricism, positivism, relativism, etc. (I know such words are often over used and vague.) I think we have shied away

I don’t know whether others might share any of these ideas, but they do suggest to me that we are now at an appropriate stage to think about the next steps in the programme as a whole, or rather, the next steps in how the various projects might ‘speak’ to each other more effectively in order to generate a genuinely interdisciplinary ‘science’ (in the widest sense) of evidence. This would support Phil’s idea of cross-disciplinary partnerships, although, of course, such things have to arise out of the projects and cannot be ‘forced’.

So I think that’s all quite an achievement and good use of a day.
Best wishes
Stephen

3.1.3 Jason Davies wrote with title Re: Challenge feedback
addressing Stephen Rowland and Evidence Personnel (Wednesday, June 8, 2005 2:13:35 pm)

A historian’s perspective.

Overall, the Schum Challenge day went a great deal better than might have normally been expected of such a mixture under such time constraints. The smooth running seemed to facilitate what is fairly described as a breakthrough (no less) for the Project, to my mind.

Essentially, the day provided a chance to discover that participants had been far more unclear than we realised about what different interests and claims were. In some ways, this seemed absurdly easy when it actually happened: in brief, it seemed to me that various concerned parties discovered to some relief that we were almost entirely ‘on the same page’. This is a bold claim to make for such a diverse group, and it may be that concerns were left unvoiced. But for many (to whom I spoke) there was a realisation that key terms and interests had been badly misunderstood, and the Schum Challenge day allowed many of these to be clarified. I list some below:

the claims made for a ‘substance-blind’ approach to evidence were clarified as being much more limited than many had feared;
certain terms were being heard as highly charged (eg ‘science’) when they were not intended to be particularly strong terms; the role of what can loosely be described as ‘considered judgement’ was specifically excluded from the process suggested.
3.1. **FIRST THOUGHTS: POSITIVE**

To an outsider, it might seem tardy to have discovered these ‘basics’ after such a long time. But all the hard work to date has, in retrospect, been a process of becoming self-conscious about such matters: in a sense, we have been like the PhD student deciding on an exact title (which can take at least a year of a three-year process in the humanities.) In a literal sense, then, we have defined the issue, in that we have demarcated what it is not.

I will say more in an interdisciplinary report that Stephen and I will draw up more fully in due course.

3.1.4 **Nigel Harvey wrote with title Re: Challenge feedback**

addressing Philip Dawid and Evidence Personnel (Wednesday, June 8, 2005 3:02:36 pm)

At 10:56 08/06/05 +0100, Philip Dawid wrote:

Dear all,

My heartfelt thanks for your stimulating contributions to the “challenge” workshop yesterday.

I think now that it would be very helpful if each of you (including those in the “cc:” who were only in the audience) were to make, and circulate to this list, some brief jottings of what (if anything) you yourself got out of the day, and your overall view of it. Can I urge you to do this while the memory is still fresh?

Dear All

I thought the day went well. Planning and organization were good and the meeting succeeded in clarifying some ideas and dispelling some misconceptions. I felt people came away having understood more about the aims that other researchers had and how those aims differed from their own. By the end of the discussion, the boundaries of the notion that evidence can be analysed in a substance-blind way seemed to be clearer to everyone. The challenges then seemed to pertain to how the results of such analyses, once completed, can be used in decision making in particular domains (rather than to the claim that such analyses are possible in those domains.) Reasons for misunderstanding and misconception of ideas in other disciplines are surprisingly difficult to identify sometimes. I think that the day may have helped to identify some of the ones that may have been afflicting our programme. If I’m right, it attained a degree of success, even if it did not always have the character that people expected it to have beforehand.

I enjoyed the day and would like to thank everyone, especially Philip and Victoria, who put so much effort into it.

Best wishes

Nigel
3.2 Second Thoughts: A Note of Caution

3.2.1 Trisha Greenhalgh wrote with title *RE: Challenge feedback*

addressing ‘Jason Davies’ and Evidence Personnel (Wednesday, June 8, 2005 3:15:31 pm)

A historian’s perspective.

Overall, the Schum Challenge day went a great deal better than might have normally been expected of such a mixture under such time constraints. The smooth running seemed to facilitate what is fairly described as a breakthrough (no less) for the Project, to my mind. . . .

I’m not so sure.

I feel that there is a temptation to veer towards ‘group think’, having had a very enjoyable day and decided what fine chaps we all are.

I felt the day was very useful NOT because we came any closer to reaching “agreement” (I certainly didn’t agree with much of what was being said), but because the areas of disagreement (dare I use the word incommensurability) were clarified.

I geninely admire the Wigmorian approach and feel it has a place in the scholarly and systematic analysis of evidence in relation to clear, focussed, hypothesis-driven questions of a certain type. I also feel that after seeing some worked examples, particularly Mark and Terry’s fascinating presentation at the end, I understand much more clearly how this approach is actually applied in practice, and how it might add value. So it was all worth it, just for that.

Here’s where I part company with many of yesterday’s discussants. I feel that the circumstances in which this “integrated science of evidence” is helpful and appropriate are relatively limited, whereas some people yesterday seemed to think it was virtually universally applicable. I felt at some stages in yesterday’s discussion, we were back discussing some pretty elementary philosophical questions such as ‘does all intellectual enquiry have to be highly focussed, hypothesis driven, and summarisable in terms of algorithmic alternatives?’ If the answer is yes, my response is ‘to a man [sic] with a hammer, everything is a nail’.

In terms of where this leaves the Leverhulme project, I think we should look critically at whether we have really “reached agreement”, and indeed at whether this was an appropriate goal to aim for. I feel just a tad like I’m being asked to give my life to Jesus....

Trish
3.2. SECOND THOUGHTS: A NOTE OF CAUTION

3.2.2 David Lagnado wrote with title *Re: Challenge feedback*

addressing Philip Dawid and Evidence Personnel (Wednesday, June 8, 2005 3:40:03 pm)

At 10:56 08/06/2005, Philip Dawid wrote:

Dear all,
My heartfelt thanks for your stimulating contributions to the “challenge” workshop yesterday.
I think now that it would be very helpful if each of you (including those in the “cc:” who were only in the audience) were to make, and circulate to this list, some brief jottings of what (if anything) you yourself got out of the day, and your overall view of it. Can I urge you to do this while the memory is still fresh?

Hi
I too enjoyed the meeting. It clarified several of the ‘opposing’ positions, and helped dissolve some issues of contention. For me it also raised several questions:

What would an integrated science of evidence look like? Would it consist in a set of core principles that could be used to generate a variety of formalisms and methodologies for acquiring, processing & using evidence? (e.g., could Wigmore charts, Bayesian networks and argumentation logics all be derived from these core principles?) Would it be primarily normative, or would it have prescriptive and descriptive dimensions (i.e., how much do we need to worry about the accessibility of this science to non-expert reasoners)? What criteria could we use to assess any potential set of principles? (Would it be circular to talk about evidence here?)

I also think that it important to sharpen and refine some of the challenges that were voiced, and to continue the dialogues in a more focused manner. There are fundamental disagreements that remain unresolved, and it would be a shame if these fell by the wayside.

cheers Dave
3.3. ATTEMPTS AT RESOLUTION

3.3 Attempts at Resolution

3.3.1 Jason Davies wrote with title RE: Challenge feedback
addressing Trish Greenhalgh and Evidence Personnel (Wednesday, June 8, 2005 3:55:15 pm)

I’m not so sure.

I can’t see anything in your account that disagrees with anything in mine...we got clearer about where evidence is relevant and where it is not. The only thing I can see between what we (you and I) said is that I thought they were being very clear about where evidence was relevant - and at every occasion, they stressed that they were talking about specific types of instance, and refused to try to appropriate (eg) judgement or (eg) policy-making. I was less convinced that this was others’ perception though...

3.3.2 Jason Davies wrote with title Re: Challenge feedback
addressing David Lagnado (accidentally omitting the rest of the list!) (Wednesday, June 8, 2005 4:13:00 pm)

What would an integrated science of evidence look like

not sure if we are supposed to be discussing this in this group, so please tell me if not...

Firstly, I think one thing that the term science should be largely dropped (bear with me). It is used to mean so many things that it can *only* be misleading. What do we gain over calling it ‘a science of evidence’ that we don’t have in talking about (eg) ‘an interdisciplinary approach to evidence’ or ‘a subject-blind approach to evidence’? Those in science departments are accustomed to using the word in a rather innocuous way (often because they are ignorant of its history....); those in the humanities, on the other hand, are used to hearing it as a *contrast* with what they do (papers on ‘a science of history’ always kick up a storm, often - but not always - based on misunderstanding). For instance, if ‘science’ is a mild claim, then there should be no problem dropping it for an equally general term. But if it claims things like ‘universal application’, ‘repeatability’ etc – such are the usual claims for science, then we are in trouble because the humanities will simply not be able to use it, for all the reasons they don’t speak about being scientific now. History is no less meticulous, rational, comprehensible etc for not being scientific.
I stress this is not a storm in a teacup - unless you want to privilege non-humanities perspectives, which means you have dropped the interdisciplinary ball already.

Secondly, my interest (allow me to flag it now) in the near future is precisely the opposite: I think (and cautiously invoke some of Trish’s argument, without her permission) we need to clarify further in what context this approach to evidence operates, and not just to allay fears. It will be far easier to discuss it when we know what it is *not*, because its scope will be clearer to us.

3.3.3 **David Lagnado wrote with title Re: Challenge feedback**

addressing Jason Davies (still off-list, perhaps not realising) (Wednesday, June 8, 2005 4:31:42 pm)

1. Point taken about the use of ‘science’.
2. I think we’re just emphasizing two sides of the same coin here. Discussing in more detail what it would look like clarifies what it is *not*, and it’s intended scope etc. It also has the additional benefit of helping us develop something positive (rather than just knowing what we’re not planning to do).

Trisha’s comments aside, I think some common ground is being developed...
3.4 Problems Reiterated

3.4.1 Janet McDonnell wrote with title *Re: Challenge feedback*

directing Philip Dawid and Evidence Personnel (Thursday, June 9, 2005 1:13:39 pm)

Dear Philip and All,

Assuming the phatic bits (thanks to organisers, lot of work, nice reception in
evening, etc), I too am in the less-optimistic-about-what-was-achieved-camp. I
think the original idea of the ‘challenge’ was to try to get down to some specifics
about the problems different groups are attempting to address and in particular to
see what part if any the schumian et al armoury could play in addressing this
by brings some real problems to the ‘table’. One outcome would have been
a better appreciation of what each ‘side’ is dealing with, another would be to
scope/identify the limitations of the armoury.

There wasn’t much engagement at all with any specifics, and I think this weak-
ened what was possible to achieve in the way of richer understanding all round.
Perhaps this was to do with the structure of the event, and/or perhaps, in retro-
spect the rhetoric of ‘challenge’ was not particularly helpful. Maybe we should
try collaboratively working on some problem(s) and thus get to understanding via
non-confrontational constructs. Others have also suggested this.

More fundamentally though I feel that we do have some issues which will have
to be addressed by going down the route Stephen has outlined more eloquently
than I can in his response.

“If it is to do this, it will need to cast its interests much wider than simply
forms of evidential reasoning, to consider wider questions. This would inevitably
lead us into methodological debates around empiricism, positivism, relativism,
etc. (I know such words are often over used and vage.) I think we have shied
away from such debates so far, perhaps because of entrenched positions, or tacit
assumptions.”

I am still not sure how clear it is, or I should say whether it is accepted, that the
situations we are interested in are ones where people are trying their best to make
informed decisions based as far as possible on evidence but where the reasoning
they (must) use has, of necessity, to go beyond what can be proved categorically
or probabilistically to what is plausible. The reasons why people do this is an
ontological issue not a question of poor reasoning with evidence. Basically we
are saying that people aren’t (just acting) stupid(ly), implying that if they used
wigmore charts, or whatever they would get to the truth or as close as they can?
let’s say probable truth.

I think where we have agreement, then, is in the not-very-interesting-region -
namely if deductive or probabilistic reasoning is possible it should be done, and
3.4. PROBLEMS REITERATED

to go a bit further, even when it might not be always possible, a good training on cases where it is applicable would heighten one’s awareness about one’s reasoning in general.

What we are doing is working in the swampy ground. On the high ground? if we were there, we should use the weapons of logic and probability, and somewhere there are the badlands where evidence plays little part (machiavelli, acceptance of bad rhetoric like ad hominem, post hoc ergo propter hoc, ad misericordium, petitio principii) where none of us want or need to go. However, I am not at all sure that there is agreement that the nature of the situations we are interested cannot be reduced to one where logic and probabilistic reasoning will give us ‘truth’. Like those who advocate using Wigmore charts, etc we are concerned to help people improve the way they reason with evidence. However we cannot help them unless we accept the characteristics of the situations in which they wish to bring evidence to bear. We itemised some of these yesterday? they are the well-rehearsed distinctions between ‘rational problem solving’ and practical decision making? all discussed ubiquitously in the literature from at least the last 30 years on policy making, decision making, and so on. So, whilst we are admitting as ‘rational’ such rhetorical devices as use of analogy and metaphor, and acknowledging that different audiences find arguments (and evidence) more or less plausible depending on what value hierarchies they adhere to and what argumentative moves they consider legitimate we still have an interest in evidenced-based action.

Janet

3.4.2 Emma Byrne wrote with title Re: Challenge feedback

addressing Janet McDonnell and Evidence Personnel (Thursday, June 9, 2005 1:40:16 pm)

Dear Evidencers,

I’d like to add my voice to that of Janet’s and to stress again (sorry to belabour the point) that we wouldn’t for a moment propose “abolishing” recourse to probabilistic or formal logical reasoning. Our position, simply put, is that there are other, still rational, approaches to problem solving, policy making, decision making and so forth. These methods can be brought to bear when logic and probability are not appropriate, nor even desirable, tools.

1This may well be a reference to the framework and language of Schon, D. (1983) The Reflective Practitioner: How Professionals Think In Action; see http://tinyurl.com/dmvhf for further details. We cannot tell, of course, who would have recognised the reference. Thus, it may have seemed to be a metaphor relying on ‘generic’ understanding, or a precise reference to an academic frame of reference.) Here, we have a further response to the metaphor of establishing a base camp – we are not there yet, nor are we in a position to establish any base camp.
3.4. PROBLEMS REITERATED

I believe that the event helped to further understanding of the respective positions, though there may still be some way to go. However, it’s not the case (nor should it be) that understanding = agreement. I’d like to lend my weight to calls for a collaborative framing of the discourse in order that such understanding that we do have be increased and the resulting areas of disagreement be profitably explored - who knows what exciting findings lurk therein!

Thanks to Janet for her rather more eloquent treatment of our position, and thanks to the organisers and participants of the challenge day

3.4.3 Philip Dawid wrote with title Re: Challenge feedback
addressing Janet McDonnell and Evidence Personnel (Thursday, June 9, 2005 1:53:20 pm)

VERY briefly, in response to Trisha, Janet and Emma:
An “integrated” theory of evidence need not be monolithic, any more than a “uni” versity should seek to shoehorn all academic enquiry inside a single approach. What (I hope) binds us together is a shared interest in deepening our understandings, of both our own and of others’ positions. Out of this common ground, let a thousand flowers bloom!

Philip

3.4.4 Trisha Greenhalgh wrote with title RE: Challenge feedback
addressing ‘Jason Davies’ (once again, privately, possibly in error) (Thursday, June 9, 2005 4:25:25 pm)

I’m not so sure.
I can’t see anything in your account that disagrees with anything in mine...we got clearer about where evidence is relevant and where it is not. The only thing I can see between what we (you and I) said is that I thought they were being very clear about where evidence was relevant – and at every occasion, they stressed that they were talking about specific types of instance, and refused to try to appropriate (eg) judgement or (eg) policy-making. I was less convinced that this was others’ perception though...

At one stage Janet asked them a specific question - what sort of problem is this approach relevant to? And the reply came loud and clear: “all problems”.
Trisha Greenhalgh
3.4.5 Philip Dawid wrote with title Challenge feedback: email post

addressing Evidence challenge personnel – Terry Anderson (Thursday, June 9, 2005 6:06:29 pm)

Dear all,

Many thanks to those who have provided some feedback on the “Evidence challenge” day. That has been an interesting interchange in itself.

I am thinking of putting these messages together and circulating them round the full evidence list. Does any one object to that? Is there any one who wants to add further to this interchange before I do so? (this could alternatively, and perhaps preferably, be done by posting further messages to the list).

Philip
3.5 Problem Compounded

3.5.1 Jason Davies wrote with title Re: Challenge feedback: email post

addressing Philip Dawid and Evidence Personnel (Thursday, June 9, 2005 11:14:53 pm)

Does any one object to that?

nope.
I would like to add that Janet has clarified what I didn’t understand about her position nicely: but I understood the ‘pro-Evidence’ statements differently from her. To be more precise, when the question that Trish mentioned was asked, I think it was misunderstood...(clear now?:-)

3.5.2 Stephen Rowland wrote with title Re: Challenge feedback: email post

addressing Jason Davies privately (Friday, June 10, 2005 9:48:09 am)

At 23:14 09/06/2005 +0100, Philip wrote:
Does any one object to that?

nope.
I would like to add that Janet has clarified what I didn’t understand about her position nicely: but I understood the ‘pro-Evidence’ statements differently from her. To be more precise, when the question that Trish mentioned was asked, I think it was misunderstood...(clear now?:-)

NO
Stephen Rowland

3.5.3 Janet McDonnell wrote with title Re: Challenge feedback: email post

addressing Jason Davies privately (Friday, June 10, 2005 9:59:55 am)
Does any one object to that?

nope.

I would like to add that Janet has clarified what I didn’t understand about her position nicely: but I understood the ‘pro-Evidence’ statements differently from her. To be more precise, when the question that Trish mentioned was asked, I think it was misunderstood...(clear now?:-)

not really but would be interested in elaboration via back channel? BTW very much enjoyed your piece - am always amazed by a historian’s perspective - i.e. looking over millennia, j
3.6. PROBLEM REJECTED

3.6 Problem Rejected

3.6.1 Philip Dawid wrote with title RE: Challenge feedback: email post

addressing Trisha Greenhalgh and Evidence Personnel, apparently bringing a private discussion back to the list (this message has multiple quotation levels – they should make sense in the sequence presented: (Friday, June 10, 2005 11:07:35 am)

On Thu, 9 Jun 2005, Trisha Greenhalgh wrote:

Good idea - funny how the emails seem
to cut deeper than some of the discus-
sion on the day - or perhaps it’s just that
I missed some of it.

did you get to any of the geller-anderson session?
that was quite educational.

i am particularly sorry that you missed anderson’s off-the-
cuff wigmore chart for your problem. i think that had quite
an impact on some folks at least...

philip

On Fri, 10 Jun 2005, Trisha Greenhalgh wrote:

Geller-Anderson presentation was excellent, but the wigmore chart
on our evidence (reported to me) didn’t help at all - I could have drawn
it up myself and indeed my own scholarly analysis looked pretty sim-
ilar. The point is, it didn’t answer the question we posed - which is
what should be DONE. Read Janet’s email carefully – I really don’t
think you’ve grasped where we’re at yet, and that’s the disappoint-
ment :-).

Dear Trish,

Almost certainly I’m NOT coming from where you’re coming from... and con-
versely! It seems to me that you are quite unreasonably expecting something that
nobody is claiming to be able to provide, and are, again unreasonably, frustrated
not to be getting it.

The current state of a putative “integrated theory of evidence” (to give it a
grand and certainly undeserved name) is extremely primitive. We do not now
3.6. **PROBLEM REJECTED**

have, and can not reasonably expect ever to have, some sort of super-hero theory that will whizz out of a clear sky and solve every one’s difficult problems (and I don’t believe that Schum or any one else involved should be read as making any claim to such magical powers). Until now the main concerns have been with understanding the basic purely inferential aspects of reasoning with evidence. How to use evidence to inform decision and policy is certainly an important question – and there is of course a formal theory of one – and many-person decision problems, heavily studied by psychologists, economists, statisticians etc., though the scope of this is pretty limited (though perhaps greater than some are willing to admit) — and we have not yet studied such policy issues in a broader context. I do think that this is an important area for further study, but it won’t be a doddle. I also think that it should be treated as a distinct problem, that can only really be addressed AFTER we have some understanding of the inferential impact of the evidence in play. In your own problem, this would involve, inter alia, addressing the issue of the perceived quality hierarchy of different types of evidence. This is I think the first problem (itself of course non-trivial) to attack, before we can begin to move on to policy issues.

The overall task as I see it is to subject such problems to various existing and still-to-be-developed general techniques (including, but certainly not confined to, Wigmore charting), to see what light these can shed. Then by identifying the ways in which these fall short, we can try and improve or replace them. I don’t think that we should give up on this enterprise just because it is difficult and still very rudimentary, and certainly not out of frustration that the tools we currently have aren’t up to the job. I very much hope that every one who thinks they have some sort of place on our “Evidence” programme would be happy to sign up to this very minimal “mission statement”.

The term “Science” has been a cause of some confusion in our programme, but I would like to emphasise one aspect of the “scientific method” that seems to me crucial: that is that we should take an experimental and incremental approach in our studies of evidence. We come with tentative suggestions and standpoints, not ready made answers. These suggestions need to be tried out on real-world problems, and their strengths and weaknesses identified. From this point of view I think your challenge was very appropriate, and (though you may not agree...) a great success!

I can understand, but not really sympathise with, your disappointment that the problems you can’t solve we can’t solve either. We are at the very bottom of the learning curve and have a long climb ahead. Let’s not lose heart too quickly.

All good wishes,

Philip
3.6. PROBLEM REJECTED

3.6.2 Trisha Greenhalgh wrote with title *RE: Challenge feedback: email post*

addressing ‘Philip Dawid’ and Evidence Personnel (Friday, June 10, 2005 11:37:22 am)

Dear Trish,

Almost certainly I’m NOT coming from where you’re coming from... and conversely! It seems to me that you are quite unreasonably expecting something that nobody is claiming to be able to provide, and are, again unreasonably, frustrated not to be getting it.

Yes, maybe you’re right! We came in on Tuesday trying to see whether we could mesh some incommensurables, and we weren’t able to, and I’ve no right to be disappointed. I do feel that exposing these “disappointments” and talking them through is critical for the academic ‘reframing’ that is the key byproduct of interdisciplinary work. So sorry if anyone took that personally. I wasn’t disappointed in the people, honest.

I’ve just had a fascinating conversation with William T who was keen to point out that Terry A is generally very keen to emphasise that the Wigmorean approach is NOT a universal solution for every problem but a specific approach designed for (and best applied to) the task of inferential reasoning. As someone’s already said, it may have been a distortion to bring our guests over and present them with a raft of problems and say ‘go on, solve THAT’. In some ways my team are skiing off piste (miles away from inferential reasoning), and maybe subconsciously expecting others to follow. But I disagree with you Phil that the way to research policymaking is to get better at inferential reasoning and THEN apply this approach to the difficult problems faced by policymakers. We’re approaching it very differently - from a different conceptual and theoretical framework, and we think we’re onto something.

But the most important thing to note is surely that for once at UCL, we’re spending some time using our brains to address difficult intellectual problems instead of filling out bits of paper for the bureaucrats. That for me is a real plus of what we’ve called the ‘Schum challenge’. I don’t really care whether we “reach agreement”!

DARE I suggest that we should all pick up in a year’s time and report back with MORE DATA?? And that aim for a better understanding of our different perspectives rather than an “integrated theory of everything”?

Trisha Greenhalgh
Professor of Primary Health Care
3.6.3 Philip Dawid wrote with title RE: Challenge feedback: email post

addressing Trisha Greenhalgh and Evidence Personnel (Friday, June 10, 2005 11:52:12 am)

On Fri, 10 Jun 2005, Trisha Greenhalgh wrote:

... But I disagree with you Phil that the way to research policymaking is to get better at inferential reasoning and THEN apply this approach to the difficult problems faced by policymakers. We’re approaching it very differently - from a different conceptual and theoretical framework, and we think we’re onto something.

I confess that, right after sending that, I had some second thoughts! I in fact believe that, when the specific goal of an evidential analysis is to inform some policy choice, that aim needs to be kept clearly in mind to guide the framing and development of an appropriate “purely inferential” analysis.

DARE I suggest that we should all pick up in a year’s time and report back with MORE DATA??? And that aim for a better understanding of our different perspectives rather than an “integrated theory of everything”?

Wonderful idea. Suggest a date!
    Philip

3.6.4 Jason Davies wrote with title Re: Challenge feedback: email post

addressing Jason Davies and the Evidence Personnel (Friday, June 10, 2005 12:17:14 pm)

I would like to add that Janet has clarified what I didn’t understand about her position nicely: but I understood the ‘pro-Evidence’ statements differently from her. To be more precise, when the question that Trish mentioned was asked, I think it was misunderstood...(clear now?:-)

since this went down like a stone, I’ll explain more fully.
    Trish said:
3.6. PROBLEM REJECTED

At one stage Janet asked them a specific question - what sort of problem is this approach relevant to? And the reply came loud and clear: “all problems”.

I understood this to mean something like the following: it was asked with the implication that ‘relevant to all problems’ implied such things as ‘can this assist in making non-evidentiary decisions (like policy-making)?’. But what was understood and answered was ‘can this contribute something in its specific role of dealing with evidence to any problem?’ ie it was asserted that where the approach works (to examine the treatment of evidence), it can contribute. In my assumed example then, rather than assist in policy-making judgements, the approach could assist in checking the conclusions about those things on which judgements could be made.

I am struggling to think of useful analogies. Perhaps each ‘side’ could say—whether I am on the right track?

3.6.5 Emma Byrne wrote with title Re: Challenge feedback:
email post

addressing Jason Davies and Evidence Personnel (in direct response to the previous message) (Monday, June 13, 2005 10:44:57 am)

Hi Jason (et al.)
A useful attempt at synthesis. Certainly we’d be among the first to argue that where such approaches are of use then they should of course be adopted. However, their use in policy making is largely restricted to only those questions in which the definition of the set of evidence to be considered is uncontentious, and that some single or competing hypotheses or questions already exist.

With regard to your example, below - it would be excellent if this were the case, however policy decisions are not easily verifiable by some objective standards. Reasons include:

The fact that one can’t clearly identify the effect of the policy decision taken. The policy is not implemented in isolation nor is it implemented against a static background. The effect of the policy cannot easily be separated from the effects of changes in society for example.

Nor can one verify against any meaningful standard such as “the best” policy (best on what grounds? according to whose values?)

The definition of the “problem” to be solved may be as contentious as the solution, so there is no one “hypothesis” being tested (eg, is sex education for the under 16s a policy to: delay sexual activity, make such activity safer, sell more
condoms, prevent teenage pregnancies, destroy the nuclear family unit, undermine religious dogma, all (or none) of the above... apologies for the effect on spamfilters!)

What we would argue instead is that rather than attempting to verify the outcome (the policy itself) we should be bringing reflective practice, support tools, frameworks or whatever else is in our arsenal, to improving the policy making /process/. This is where we come back to the thorny problem of using evidence. Part of that process should ideally be a conscious and well understood attempt to “make sense” of the evidence. Some of that evidence could be a randomised controlled trial for example, or some other such evidence that allows us to adopt a probabilistic or formal-logical way of deciding what that evidence “points to”.

But there are other types of evidence, and other questions about evidence, that require some further, rational, “healthy” techniques in order to incorporate that evidence into the process. For example: how do we incorporate questions of ethics, values or moral good in our considerations? How do we account for competing frames such as those who view sex ed as state interference in a private matter vs those who see sex ed as a necessary tool to prevent some social ill.

This is where we feel that “healthy” rhetoric has its part to play. A judicious understanding of Rhetoric would allow us to differentiate between unhealthy influences (as baculum arguments, the framing of false dichotomies, fallacious ad hominem) from healthy arguments (the informative use of analogical reasoning, the finer graining of concepts by dissociation). We feel that this would incorporate the /reality/ of how policy decisions need to be made in practice with a striving for the ideal - a transparent and healthy process of policy making.

I’d like to attempt a synthesis here - though I await correction, as it’s based on a tentative understanding of your presentation. It’s early days yet but I feel that judicious use of “healthy” rhetoric would not be a prescriptive extension to institutionalised, routinised decision making. Rather, it could provide groups of policy makers with tools to recognise and make the wise and discretionary judgements that are characteristic of the charismatic style, whilst to some extent inoculating those same groups against the capricious tendencies that are among the charismatic style’s less desirable traits.

Emma Byrne

3.6.6 Jason Davies wrote with title Re: Challenge feedback: email post

addressing Emma Byrne and Evidence Personnel (Monday, June 13, 2005 11:09:43 am)

I’d like to attempt a synthesis here - though I await correction, as
3.6. PROBLEM REJECTED

it’s based on a tentative understanding of your presentation. It’s early
days yet but I feel that judicious use of “healthy” rhetoric would not be
a prescriptive extension to institutionalised, routinised decision mak-
ing. Rather, it could provide groups of policy makers with tools to
recognise and make the wise and discretionary judgements that are
characteristic of the charismatic style, whilst to some extent inocu-
lating those same groups against the capricious tendencies that are
among the charismatic style’s less desirable traits.

I still maintain (mildly!) that this would not be contested by anyone who spoke
last week. I am not in any way suggesting that there is a way of ‘checking’ that
‘the right policy’ has been implemented (which *seems* to be what you think I
am saying). But i’m caught in the middle here, as it were, trying to speak for both
sides of this. What do others make of this discussion?

3.6.7 Emma Byrne wrote with title Re: Challenge feedback:
email post

addressing Jason Davies and Evidence Personnel (Monday, June 13, 2005 12:11:01 pm)

Mea culpa: I misunderstood your statement “[This] approach could assist in
checking the conclusions about those things on which judgements could be made.”
– I took that as a suggestion that they could be used to validate the judgements
themselves, rather than the basis on which those judgements may be founded - my
apologies.

If there is agreement on (and understanding of) our position then that’s ex-
tremely heartening, and quite surprising. However I wasn’t sure last week that
we’d successfully conveyed our position, to the point where it was universally
understood. I’d rather reach a position of disagreement based on perfect under-
standing than agreement based on “false consciousness”: however it seems that
our position is becoming clearer - perhaps this on-line wrangling has helped! Nev-
ertheless - I still welcome well founded (or even ill founded) disagreement...

That said - thanks for the surveying of the common ground,
Emma
3.7 Framing the Problem More Widely

3.7.1 Hasok Chang wrote with title Re: Challenge feedback: email post

addressing Emma Byrne and Evidence Personnel (Tuesday, June 14, 2005 4:56:13 pm)

Dear all,

Here are some of my thoughts following the evidence meeting on the 7th. Apologies for this late entry – only now I’ve managed to get enough time and mental space away from exams, etc.

First of all, I’d like to add to the chorus of thanks to the organisers and speakers for a most enjoyable and stimulating day.

Many people have already pointed out the importance of distinguishing evidence and its use. Once the distinction is made, I don’t hear anyone in the group saying that there is a substance-blind method of finding the best of use of evidence (for social policy, etc.). So there we seem to have agreement. Still, we are left with an interesting and difficult problem of project-definition for our programme as a whole, since some of us clearly are concerned about use and others prefer to focus on evidential reasoning itself for now.

Setting that question aside (not because it’s unimportant, but because I don’t have anything bright to say about it at the moment), let me turn to another issue that seems to deserve further discussion.

Does evidential reasoning always involve dealing with hypotheses? I am inclined to say yes, though I have to think more about this. (But I don’t disagree with Trish and Stephen – read on!) If we say something is “evidence”, it has to be evidence FOR (or against) something; what could that something be, except a hypothesis? So, I think the talk of evidence only makes sense in the context of the testing/confirmation of hypotheses. (Saying that is not the same thing as saying that a substance-blind approach is possible. In fact, I was just debating about this with the philosopher Peter Achinstein several weeks ago at the LSE: Achinstein, who definitely believes that there is a substance-blind concept of evidence, does not agree that evidential reasoning always happens in the context of hypothesis-testing.)

Now, this is NOT to say that all reasoning or thinking or learning has to involve hypothesis-testing. Therefore, I would certainly not answer the following question (as put by Trish) in the affirmative: “does all intellectual enquiry have to be highly focussed, hypothesis driven...?” And I have no objections to Stephen’s approach, “which tends to be towards illuminative enquiry rather than hypothesis testing”.

The broader point is this: not all good processes of enquiry, judgement or inference are evidence-based, even if they are experience-based or fact-based.
3.7. FRAMING THE PROBLEM MORE WIDELY

One possible response: “Yes, but why should we care? We are a research programme concerned about evidence. Lots of jolly good things are not about evidence, but that shouldn’t impact on what we are trying to do. Particularly, to say that some things aren’t and shouldn’t be evidence-based is no argument against the possibility or desirability of crafting an integrated science of dealing with situations that ARE about evidence.”

A different response (quoting Stephen again): “I think it is important that the programme can raise a thoughtful critical response to the rhetoric of evidence-based practice/policy/etc... If it is to do this, it will need to cast its interests much wider than simply forms of evidential reasoning, to consider wider questions.”

It is not clear to me how we can reconcile these two positions in a way that gives our programme overall coherence. But here is one suggestion: we could take it as part of the remit of an “integrated science of evidence” to clarify its own limits. In other words, we could try to answer this question: given that this is what we think evidence is, what kind of role can it play, where? That could be at least one part of the overall aim of our programme.

Sorry to ramble on a bit. I hope this is constructive in moving our big-picture debate forward.

best,
Hasok

3.7.2 David Lagnado wrote with title Re: Challenge feedback:

email post

addressing Hasok Chang and Evidence Personnel (Tuesday, June 14, 2005 5:18:41 pm)

At 16:56 14/06/2005, Hasok Chang wrote:
Dear all,
Here are some of my thoughts following the evidence meeting on the 7th. Apologies for this late entry – only now I’ve managed to get enough time and mental space away from exams, etc.
First of all, I’d like to add to the chorus of thanks to the organisers and speakers for a most enjoyable and stimulating day . . .

Hi Hasok
Thanks for some illuminating comments. I agree that it is important for any purported ‘integrated theory of evidence’ to clarify its own limits. I would just add that there at least two (non-exclusive) ways to proceed:
(1) to actually forge ahead and construct something, then investigate its scope wrt
3.7. FRAMING THE PROBLEM MORE WIDELY

a variety of real problems
(2) to debate its scope/limits beforehand

   I can’t help feeling that at this point (1) is more promising - not just because something (however limited) will actually be developed, but also because it will provide a better target for sceptics to aim at (and the goal-posts are less easily shifted).

   cheers
   Dave

3.7.3 Stephen Rowland wrote with title Re: Challenge feedback: email post

addressing Hasok Chang and Evidence Personnel (Tuesday, June 14, 2005 6:09:46 pm)

   At 16:56 14/06/2005 +0100, Hasok Chang wrote:
   Here are some of my thoughts following the evidence meeting on the 7th. Apologies for this late entry – only now I’ve managed to get enough time and mental space away from exams, etc.
   First of all, I’d like to add to the chorus of thanks to the organisers and speakers for a most enjoyable and stimulating day.
   Many people have already pointed out the importance of distinguishing evidence and its use. . .

   Dear All
   Just a very obvious point in response to Hasok’s seemingly obvious point that evidence must be evidence for something, ie for a hypothesis.
   This is of course true. All instances of evidence relate to a hypothesis.
   But it does NOT follow from this that the truth value of all meaningful propositions are to be judged on the basis of evidence. In particular, moral statements (like ‘one should be nice to people’) may or may not be true, but its truth does not depend upon evidence (although, of course, whether or not people believe it to be true is based upon evidence).
   Policy statements (and educational ones) are invariably a complex mixture of factual and moral assertions. Only the former can be based upon evidence. Therefore such statements cannot normally be justified solely on the grounds of evidence.
   Unless, that is, you believe (as some philosophers do) that all moral propositions can be reduced to factual ones. (ie that ‘ought’ can be reduced to ‘is’)
   Sure this is all very obvious to everyone, but perhaps needs saying
   best wishes
   Stephen Rowland
3.7. **FRAMING THE PROBLEM MORE WIDELY**

### 3.7.4 Trish Greenhalgh wrote with title *Re: Challenge feedback: email post*

addressing Hasok Chang and Evidence Personnel (Tuesday, June 14, 2005 8:02:00 pm)

Hasok: “If we say something is ‘evidence’, it has to be evidence FOR (or against) something; what could that something be, except a hypothesis?”

Answer: an argument....???

trish

Prof Trish Greenhalgh

### 3.7.5 Jill Russell wrote with title *Re: Challenge feedback: email post*

addressing Hasok Chang and Evidence Personnel (Wednesday, June 15, 2005 7:42:59 am)

At 4:56 pm +0100 14/6/05, Hasok Chang wrote:

Many people have already pointed out the importance of distinguishing evidence and its use. Once the distinction is made, I don’t hear anyone in the group saying that there is a substance-blind method of finding the best of use of evidence (for social policy, etc.). So there we seem to have agreement. Still, we are left with an interesting and difficult problem of project-definition for our programme as a whole, since some of us clearly are concerned about use and others prefer to focus on evidential reasoning itself for now.

But is it as unproblematic as you imply to distinguish evidence and its use? How do we decide where the construction of evidence (‘evidential reasoning itself’) ends and its use begins?

If the boundary between these entities is seen as more indeterminate then the two positions you identify (‘we are a research programme concerned about evidence’, and ‘the programme needs to cast its interests much wider than simply forms of evidential reasoning’) are perhaps not so difficult to reconcile.

I think there’s a danger here that theory (in the sense of ‘evidential reasoning itself’) gets placed outside practice. Evidential reasoning and use of evidence are both social practices, and both can (need to?) be explored as such.

Jill
3.7. FRAMING THE PROBLEM MORE WIDELY

3.7.6 Philip Dawid wrote with title Re: Challenge feedback: email post

addressing Trish Greenhalgh and Evidence Personnel (Thursday, June 16, 2005 1:21:47 pm)

On Tue, 14 Jun 2005, Trish Greenhalgh wrote:

Hasok: “If we say something is ‘evidence’, it has to be evidence FOR (or against) something; what could that something be, except a hypothesis?”

Answer: an argument....???
trish

surely not! when we call on evidence as part of an argument, it is because we consider it relevant, directly or indirectly, to the purpose of that argument, which is to support/undermine/persuade others of some conclusion (i.e. hypothesis). the argument itself is an instrument, not an end.

i myself would slightly widen hasok’s claim, and regard the scope of “evidence” as relating to *discrimination* between contending hypotheses (including as yet unframed hypotheses), rather than focused on a single hypothesis. police aiming to gather “evidence” against their prime suspect might do well to remember this...
philip

3.7.7 Tony Gardner-Medwin wrote with title Re: Challenge feedback: email post

addressing Philip Dawid and Evidence Personnel (Friday, June 17, 2005 10:41:51 am)

On Tue, 14 Jun 2005, Trish Greenhalgh wrote:

Hasok: “If we say something is ‘evidence’, it has to be evidence FOR (or against) something; what could that something be, except a hypothesis?”

Answer: an argument....???
trish

surely not! when we call on evidence as part of an argument, it is because we consider it relevant, directly or indirectly, to the purpose
3.7. **FRAMING THE PROBLEM MORE WIDELY**

of that argument, which is to support/undermine/persuade others of some conclusion (i.e. hypothesis). The argument itself is an instrument, not an end.

I myself would slightly widen Hasok’s claim, and regard the scope of “evidence” as relating to *discrimination* between contending hypotheses (including as yet unframed hypotheses), rather than focused on a single hypothesis. Police aiming to gather “evidence” against their prime suspect might do well to remember this...

Philip

Dear evidencers,

I’ve been away as those of you who have posted will probably realise from auto-replies. My memories of the Schum day, though almost washed away by the fantastic sights, tastes and rains of Hong Kong, are also very positive. I haven’t read the correspondence as well as I should, but felt I could maybe try to make a clear distinction that could help to clarify some issues.

Arguments may be intended either (i) to arrive at a belief, or to alter the probability (degree of belief) that people assign to a factual proposition, or (ii) to arrive at a decision, or action, or policy.

Type 1 (the commonest type in academia) is the simpler. A fact is “relevant evidence” if it alters the belief, either directly or indirectly (i.e. conditional on some other fact or belief). Type 1 doesn’t depend on values or utilities. Examples would be “Is climate change a fact?”, “Does regime Y have WMDs?”, “Does this parent abuse his children?”, “Could three siblings die of SIDS?”, “Will it rain here tomorrow?”, “Do students prefer lectures given by active researchers?”, “Do specialist nurses lead to better outcomes in circumstances X?”, “Does being nice to people make them be nice in return?”

Type 2 (probably the commonest in Tricia’s life, and the real world generally!) is more complex. Unlike type 1, it depends on utilities (values placed on outcomes). Evidence may be relevant either (or both) because it affects beliefs related to the policy or because it affects utilities, or beliefs about utilities. To run through the list of beliefs above and ask some related decision questions: “Shall we raise US fuel tax?”, “Shall we go to war with Y?”, “Shall we remove this child from his father?”, “Shall we acquit the defendant?”, “Shall we take the umbrella?”, “Shall we have teaching-only universities?”, “Shall we allocate the budget this way or that?”, “Shall we be nice to people?”.

Just to take the horny old second issue here, suppose facts increase belief that regime Y has a secret programme producing deliverable nuclear weapons. This evidence may increase belief (type 1) that Y has intent to attack other countries, which itself may be the foundation of the argument (type 2) for going to war. But it may also increase greatly the estimates of the negative utility of going to war.
3.7. **FRAMING THE PROBLEM MORE WIDELY**

with Y (you will get nuked), perhaps so much that it may act against a decision to go to war.

I’m not sure whether anyone is really confused between these different kinds of argument and applications of evidence, but it seems possibly useful to set them out quite explicitly. Any disagreements?

Tony GM
3.8 The Final Word: An Outsider’s Synthesis

3.8.1 Deirdre Dwyer wrote with title [Evidence] Challenge feedback

The following message is posted on behalf of Deirdre Dwyer (by Philip Dawid)

Dear evidence list,

I should like to thank very much everyone involved for a very stimulating workshop, and for the maintenance of your website since 2003, which has a lot of good information, and which has helped me to follow your project from Oxford.

As someone who is researching the use of expert evidence in legal decision making, I thought that I would offer a couple of suggestions on where I think that your project might go.

There seem to be some fundamental issues of definition. In particular, there seem to be ambiguities over (1) are you trying to produce a science of evidence, a theory of evidence, or something else, (2) what do you mean by evidence, (3) for what types of decision is your science/theory of evidence intended to work?

I think a ‘science’ may be more useful than a ‘theory’. The former suggests a network of concepts and theories, while the latter tends towards a single approach to evidence, which I think would unnecessarily limit the field. I admit that ‘science’ has negative connotations in some disciplines, but I think that this is outweighed by its overall utility as a concept.

David Schum’s theory about substance blind probabilistic reasoning, for example, has a very narrow application in the overall scope of human decisions involving evidence, but is very relevant to particular types of questions, that can be expressed in terms of facts and probabilistic inferences. I am very interested in the Schumian thesis, which is very helpful to me in looking at expert evidence in fact finding, but I think it may have much less application in more normative questions, such as ‘what is in the best interests of the child’. So perhaps the Schumian thesis should form part of a science of evidence, but there is a need for a framework into which that thesis fits.

I am of course biased by my research interests, but I think that there would be a lot of value in seeing how different disciplines approach the same evidential questions, to see not only how the mechanics of inference are similar/different, but also how decisions are made about things such as relevance and weight. Legal evidence, for example, has developed a whole set of slightly arcane rules in this area, that are to do not only with how one makes an accurate finding of fact, but also with enforcing ideas such as how far the state, as judge or criminal prose-
cutor, might interfere in the life of the individual. You may, for example, have seen Professor Zahler’s review of Twining and Hampsher-Monk’s ‘Evidence in History and Law’ in the International Journal of Evidence and Proof, in which he suggests that comparing law and history might be deceptively straightforward: “If evidence law and practice are to benefit from experiences in other disciplines, the next step might be to look towards other practical professions, for example clinical psychology, medicine, engineering, business, management, or military and political decision making.” (2004) 8 IJEP 211.

Best wishes Deirdre
Deirdre Dwyer DPhil Student in Law Pembroke College Oxford

Dear Tony,

I would agree with your distinction, which I think underlies a lot of the differences of opinion that have been expressed at the workshop and since about what a science of evidence should encompass.

Theories of legal decision making since at least the 17th century have largely sought to maintain that determining accurately the facts of a case (‘questions of fact’) is a distinct activity from determing what law should be applied to those facts (‘questions of law’). Lawyers are keen to insist that legal decisions are not case specific. Hume similarly insisted that the ‘is’ (the facts) is distinct from the ‘ought’ (what one should do). For him, one was a question of empirical method (evidence?) while the other was one of morality.

So although it may be true that ius oritur ex facto - law (justice) arises from the facts - how one moves in legal reasoning from your type 1 argument to your type 2 argument remains poorly understood. Courts have an additional problem that the legally ‘correct’ decision of what action to take may sometimes go strongly counter to ‘common sense’ ideas of what should be done. I think that the study of how different disciplines move from type 1 to type 2 arguments certainly warrants study - the example in your paper of how race might be used in medical, police and trial decisions. One of the difficulties of the use of expert evidence in government decision making (eg BSE) is that politicians tend to hope that the policy decisions will arise naturally from the evidence, which of course they do not.

Deirdre

3.8.2 Philip Dawid posted a compilation of the emails up to this point

on Fri Jun 10 13:00:42 BST 2005.

They can be viewed at http://tinyurl.com/9tdyx.
3.8. THE FINAL WORD: AN OUTSIDER’S SYNTHESIS

3.8.3 Terence J. Anderson wrote with title [Evidence] Belated Evidence Challenge feedback

addressing Philip Dawid (Wednesday, June 29, 2005 8:23:11 pm)

Philip and all,

I enjoyed the conference and meeting many of the participants. I have read the feedback and thought it might be helpful to clarify a bit further what Wigmore’s chart method is (and is not). I have also sketched what a “Wigmorean challenge exercise” might look like. I hope it is useful.

Notes on Wigmore’s Chart Method in Response to Challenge Feedback

The chart method of analysis is a tool. It is the most rigorous method of analysis known to Schum, Twining, Anderson, and others for depicting the relationships between the “evidence” and a hypothesis to be tested or an argument to be made. It can be used at any stage of an inquiry, from initial investigation through critiquing the arguments supporting or attacking a hypothesis that has been accepted or rejected. Its principal advantage stems from the fact that it requires the analyst to be very precise in identifying every step in an inferential argument that might be open to doubt and to produce a product that others can readily critique and refine.

The chart method is, however, only a tool for analysis and for argument testing. It is not a tool that magically identifies solutions. (Nor is it the only tool. For example, Bayes’ Theorem and Bayes Nets are also tools that are useful in many contexts.) Given a body of evidence and a proposition to be advanced, an analyst can depict in charted form her view of the strongest way in which the available evidence and inferences can be marshaled to support the proposition and the strongest way in which her opponent might marshal the evidence and inferences to undermine or oppose the proposition. This is, in essence, what I did in charting the necessary steps that had to be considered in the Iamblichus phase of Mark’s argument that cuneiform remained a living script until the third century A.D. It is also what I attempted, at a macroscopic level, at the conference in structuring the epilepsy dialog as a Wigmorean might.

The macroscopic structuring is important because it identifies the principal propositions being advanced so that the relevance of particular item of evidential data can be identified. Significant insights, however, ordinarily come from the microscopic analysis necessary to identify, inference-by-inference, the propositions necessary to demonstrate the relevance of the particular items and, by so doing, to identify the potential sources of doubt or weaknesses in the argument. The Iamblichus analysis in Evidence and Inference in History and Law: Interdis-

For purposes of the project, a Wigmorean “challenge” could be: Can the chart method of analysis be usefully employed across and within disciplines and other contexts to any inquiry that endeavors to discover and rationally demonstrate the strengths and weaknesses in the arguments that provide a logical justification for a hypothesis, be it a proposition to be proved in a legal case, a claim to be advanced in a historical context, a policy proposed or adopted in health care context, etc. We claim that it can or, with Popperian caution, we claim that we have not yet been confronted with a problem to which the tool could not be usefully applied.

A full test of the Wigmore challenge may exceed the interest and available resources of those involved in the Leverhume evidence projects. The learning barriers that a Wigmorean analyst trained in one discipline, such as law, must overcome in order to apply the method to a problem in another, such as Assyriology, almost certainly precludes a cross-disciplinary exercises. (A point illustrated and developed in Evidence and Inference at 140.) Trained Wigmorean analysts could be found, but the project has little to offer that which might attract them to make the investment necessary to apply the method in a foreign discipline. The only realistic alternative would be to train post docs from the several projects in Wigmorean analysis and to assign a project from within the discipline or context in which they are working. That would be feasible, but the question whether responding to the challenge warrants the diversion of post doc time required is one that only those involved could answer.

Terence J. Anderson
Professor of Law
University of Miami, School of Law
3.9 Later Discussion

The following emails were mostly sent offlist to a variety of individuals (though they began on the list): they show the results of the time elapsed and the chance to reflect on the discussion and thus seemed to merit being treated as a separate section. These posts were contributed for the compilation by Philip Dawid and James Edwards.

3.9.1 Michael Joffe wrote with title [Evidence] Belated Evidence Challenge feedback

addressing the Evidence List (Sunday, July 3, 2005 9:30:40 pm)

Dear all

Sorry to respond so late, but I was away and then after I came back I thought, maybe it’s too late now.

The “challenge day” and the subsequent discussion were very interesting, and I think led in useful directions. My first frustration, though, is that I’m not sure where it has left the programme. Do we all accept Tony’s analysis of type 1 and type 2? I do, and was going to try and formulate something along the same lines until I read his email (but I’m glad I didn’t, as he gave better examples than I would have done).

If we do accept Tony’s distinction, what is the scope of the programme to be? I would favour explicitly restricting it to focus on Type 1, which I would characterise as being about knowledge, and exclude Type 2, which is about decisions. Type 2 involves values, and inputs of many different kinds – and with policy decisions and legal decisions, other complications arise, as Emma, Deirdre and others have said. If we try and include all this, we are in danger of our discussions becoming too diffuse to be useful.

So maybe we should focus on Type 1 only, and just consider Type 2 insofar as it influences Type 1. Up to now, we’ve placed no restrictions on discussion, other than the implicit ones set by the various contributions, and by our own personal perceptions of the meaning of the word “evidence” – but maybe it is time to decide what is in and what is out?

My second frustration is that we have not properly done justice to the Type 1 discussion. In one of the first contributions, Stephen talked about illuminative enquiry rather than hypothesis generation and testing. Although this has been referred to at various times since, it has got mixed up in the decisions/values debate, which is a separate issue. In situations where hypotheses do not play a major role – Stephen’s illuminative enquiry – does this integrated discipline of evidence (or
3.9. LATER DISCUSSION

whatever we’re going to call it) have nothing to contribute? Or maybe it is relevant in certain ways if adapted? Hasok suggests that we address this by setting limits – restricting the scope of an integrated discipline of evidence to the hypothesis situation. I would prefer to try and engage with the type of work Stephen and other members of the group do, to see what useful connections can be made. In fact I think we’ve already done some of this, but perhaps it has not been made very explicit.

Also, returning to situations where hypotheses are being used, “evidence” applies to the testing but not the generation of hypotheses, and there are philosophical issues here that I think we’ve been avoiding – is it possible to have an integrated discipline of evidence without addressing these? – concerning what I might loosely describe as the two-way relationship between theory and empirical reality.

Best wishes
Mike

3.9.2 Richard Leary wrote with title [Evidence] Belated Evidence Challenge feedback

addressing Mike Joffe (Monday, July 4 2005 09:10 am)

Mike
What do you mean by “evidence” applies to the testing but not the generation of hypotheses”. This is taken from the text below. I would then like to comment to you.
Thanks.
Richard.

3.9.3 Mike Joffe wrote with title [Evidence] Belated Evidence Challenge feedback

addressing Richard Leary (Monday, July 4 2005 11:47 am)

Richard
Is it not clear? The rules of using evidence have been described in various contexts, and I think these tend to take the hypotheses to be tested as given. I haven’t checked systematically through all the contexts, but in the context of criminal law the hypothesis is whether (a) particular defendant(s) is/are guilty. In science, with which I am more familiar, it applies to the testing of a hypothesis that already exists. So that leaves various questions, including “Where does this hypothesis

\footnote{Due to the reorganisation for this compilation, this text is now the message above.}
3.9. LATER DISCUSSION

come from?” “Why is it privileged over others?” And so on. I am no philosopher, but it’s clear that the relationship between the attainment of “knowledge” and the way these processes operate has been discussed at least since Hume, forms an important part of the philosophy of science, and I believe I’m right in saying is still not “solved”.

I hope that helps. Best wishes

Mike

3.9.4 Richard Leary wrote with title [Evidence] Belated Evidence Challenge feedback

addressing Michael Joffe, Philip Dawid and one other (Monday, July 4 2005 13:20:19 pm)

Mike

Thank you for the clarification. Here are some quick and ‘straight into the word processor thoughts’ that could be better refined with the luxury of time. You will no doubt get the general point. I am promoting a very wider view of evidence.

I believe that evidence applies to both the testing and the generation of hypotheses and cannot possibly see how this could be any other way.

I not only believe this I use it in the work I do routinely in investigating large collections of information that may contain ‘evidence’. Evidence for me means ‘material that may be relevant to proving whether some fact known or contemplated is true or, whether it leads to another fact not or not contemplated’. Let me explain what I mean:

My standpoint is that of an investigator and so I shall speak as one. Investigation and intelligence analysis (finding out what we do not know) are concerned with the provision of explanations or arguments that provide for some level of certainty where once there was doubt. The only rational reason for investigating anything is the presence of doubt. If we already know ‘something’ to a high degree of proof then why investigate it? (This provides the rationale for my pursuit).

If we have doubt there must rationally be some reason for it. It may be concerned with the relevance, credibility or weight of the evidence we already possess or on the other hand it may be that someone else has introduced a new line of reasoning or a different perspective on the evidence we already have. If this new explanation touches in any way upon the evidence we already possess or the hypotheses we are already contemplating then it is (to some degree) relevant and therefore I claim ‘evidence’ in itself. I am not making any claims about how persuasive or weighty the evidence is but I am claiming it is ‘evidence’ because it is information relevant to the proof of some fact now in issue. Every item of evidence always (yes always) points to another item of evidence let alone a hy-
3.9. LATER DISCUSSION

hypothesis. There is a very good reason this to be the case. It concerns the nature of evidence itself. One interesting characteristic about evidence is that it can be decomposed infinitely. I can demonstrate this to degrees of boredom you should avoid but what I mean is that one can continually decompose the elements of evidence to sub-elements ad infinitum. (Note this is not a legal view of evidence it is more concerned with science of proof). The important point is that the evidence (existing or newly introduced) always says something about itself that goes beyond the description of what it is. For example, blood and a knife: How long is the blade? What is it made from? Who’s initials are on the handle (who’s are not) The knife owned by whom? Blood emanating from whom? What is the blood type? What is the DNA Profile of the blood? Is there a body with a stab wound that fits this blade? and so on, and so on . . . . . . Similarly with a Fingerprint: What surface is the mark attached to? Is the mark made of Amino Acids or is it ink or blood or something else or all of them? What DNA profile can be found in the Amino Acid (if present)? and so on, and so on . . .

Investigation therefore concerns discovery and the pursuit of lines of enquiry based not only on the evidence we possess but also about evidence linked to evidence about which we are yet to be satisfied. This is as well as the linking of evidence to hypotheses. Investigation (finding out what we do not know) benefits from the careful analysis of evidence we possess and evidence we contemplate could exist if certain explanations run true. In this sense, evidence is not limited to testing hypotheses. If it were, how could we ever raise a rational hypothesis? It would be done so without recourse to evidence and that would amount to the search of everything and anything in the hope of finding anything or something. Not a very fruitful or profitable activity!

The search for evidence and hypotheses for testing is cyclical. One thing leads to another in a continuous process of trial and error, suggestion, rejection, construction and deconstruction. It is a highly intellectual process regardless of what we may believe from crime fiction or the media about policing. (This does not mean that all investigators are intelligent it simply means the process demands high degrees of intellectual processing). Furthermore and most importantly, this activity exists in a world that is indeed highly inter-connected in terms of cause and effect. We are surrounded by naturally and humanly created occurrences, activities and systems. These are so complex in reality that we truly cannot keep track of things that take place or exist in any given time frame. Imagine trying to monitor or gain knowledge of every simultaneous act that takes place at a given time on a given day in a given locality? Furthermore, imagine trying to place a series of events that occur sequentially between given time slots in a given locality? Our world is a highly complex place yet we know (piecemeal) that things are continually happening and are continually causing other things to happen. The quest for the human being to need to survive as well as our inability to be able to
3.9. LATER DISCUSSION

keep track of all this has helped us to develop strategies to evaluate ‘best courses of action’ based on the evidence we have or, when challenged, to use the evidence we have to raise new lines of enquiry about evidence we need to make decisions. This has stood us in good stead until now because human kind has thrived as a semi-intelligent (pardon the pun but I mean it) data (perhaps evidence) processor intent on survival. However, technological, psychological and neurological advances (they are linked) are providing us with a future in which we can truly begin to make effective steps at monitoring many things at the same time and understand many things in complex sequences. (Beginning for the first time to understand causation). This is proving to teach humans new lessons in our development: That the world is indeed a highly inter-connected place where ‘one thing leads to another’ and there are real relationships between close and distant events. We are coming to terms with understanding ‘remorenss’.

This says much more about evidence than the traditional view in which evidence is seen within small, focussed strings or reasoning operating within a closed system. The legal view of evidence promotes this but of course historically for good reason.

Best wishes . . .
Richard

3.9.5 Mike Joffe wrote with title [Evidence] Belated Evidence Challenge feedback

addressing the Evidence list (Thursday, July 7, 2005 15:16 pm)

Richard

I am glad you agree that evidence is used in generating hypotheses – often we have had discussions that seem to assume that we are just discriminating between hypotheses, or testing individual hypotheses, which leaves the question of where the hypotheses come from – and whether the process by which they are generated involves considering evidence. To me, it seems intuitively obvious that it does.

Some philosophers have stressed the imaginative, conjectural nature of hypothesis generation – and this can be important as well (interacting with the scientist’s knowledge of the already-existing evidence) – at least it is in science, it may not be so important in the context of criminal law. I think we should explore these similarities and differences more systematically, as they are illuminating.

And then, if evidence is relevant both to hypothesis generation and testing/discrimination – what you call the cyclical process – then why should it not be so in the illuminative enquiry that Tony briefly described, in which hypotheses do not play a central role? – it seems to me that this also involves some kind of cyclical
knowledge-generating process, with somewhat different rules or traditions, so that the way that a potential integrated discipline of evidence relates to it will have its own specific features.

I like your examples about the stabbing. To be honest, I get lost in some of the more abstract formulations, not just in your letter, but more generally. I think it would be useful if people could always amplify their general statements with the sorts of examples they are thinking of – I think we’ve already agreed that “substance-blind” doesn’t mean that the subject matter is unimportant. And sometimes the “general” statement may clearly apply to one type of application (criminal law, neurophysiology, ancient history, economics, quantum electrodynamics) but not others. As we progress, it is important that we see what generalisations can be validly applied in which domains. I think that this was one of the most useful features of the evidence day last month.

Best wishes
Mike

3.9.6 Richard Leary wrote with title [Evidence] Belated Evidence Challenge feedback

addressing Michael Joffe, Philip Dawid and Dave Schum (Thursday, July 7 2005 16:21:34 pm)

Mike

Thanks for the considered response. Many of the lessons I have learnt originated from the works of Dave Schum, William Twining, Terry Anderson, Charles Sanders Peirce, John Henry Wigmore and more recently, the Science of Complexity. (A more modern theory of Chaos). These Guys have taught me lessons that in hindsight were obvious yet remained invisible to my conscious understanding until they were studied in the context of evidence and proof. However, we are yet to really get stuck into understanding (perhaps with the exception of Dave Schum) the complexity of and uses of evidence. We have been ‘dogged’ for too long in the literature by theories of remoteness in terms of the relevance of evidence (especially in law) purely on the basis in my view that it protects against the travesty created by the use of unsafe evidence. However, my view is that remoteness (especially in law) has survived largely because we do not understand complexity and the nature of causality.

Regards
Richard
3.9.7 James H. Edwards wrote with title Evidence Digest: Hypothesis Generation

addressing Mike Joffe (Thursday, July 7, 2005 8:32 pm)

Dear Professor Joffe,

I have been a long-time “lurker” on the Evidence Digest mailing list, and find the formative communications most fascinating. I apologize if my unsolicited comments are unwelcomed.

Might I suggest that *one* possible source of hypothesis generation be based on Hume’s general conception of “cause and effect,” and that hypotheses generated in this fashion are attempts to conceive relationships that explain, in better fashion that those presently conceived, impressions of cause and effect.

And, pressing my luck with your indulgence of an unsophisticated layman, I pray that you might also consider:

1) Acknowledging the complete disconnection between evidence as it relates to matter of fact and evidence as it relates to matter of fact as adjudicated in a court of law. While there may be some resemblance between the two, careful consideration might show they are distinct. Moreover, there seems to be an even further additional step that moves from evidence of fact to attribution of the motive force as it relates to a quality of actor character.

2) A useful tool might be to construct an interactive Glossary through which consensus understandings of basic terms might be argued and arrived at. In such a mission as embarked on by the Evidence group, definitions of terms seem of the essence.

Sincerely,

James H. Edwards, Publisher

3.9.8 Mike Joffe wrote with title Evidence Digest: Hypothesis Generation

addressing the Evidence List (Friday, Jul 8 17:23:56 BST 2005)

James, I’m intrigued by what you say in your point (1), which seems to be saying that the courts’ use of evidence is not something that the rest of the world would accept - have I interpreted this correctly? In this programme, we have tended to accept that courts’ procedures are as valid as any others, and even if they’re not readily generalisable outside the specific context of a court of law, at

---

3This was originally sent privately, then appeared on the list when a response to it was made. I have restored it to the sequence that will make best sense for a reader of this compilation.
3.9. LATER DISCUSSION

least we can learn something important from the modes of reasoning that they employ.

I’m not competent to comment on how useful Hume would be. I do agree though with one of the early comments that we cannot ignore what philosophers have said on these issues.

Best wishes

Mike

3.9.9 James H. Edwards wrote with title Evidence – Richard Leary

addressing Richard Leary (Sunday, July 10, 2005 15:49:38)

Dear Mr. Leary:

Thank you for the privilege of your communication. Please accept this reply as a personal communication, for while I have great interest in the subject matter of the Evidence group, I have an even greater sense of uneasiness about inserting myself any further into the focused and evolutionary processes of the group; nor am I capable of sustaining such a dialogue if I were to do so. I have, however, some comments that I hope are of interest to you.

Yes, I believe I received both Professor Joffe’s post and yours more or less simultaneously; and was overtaken with an urge to reply to Professor Joffe without having read yours in its entirety. If I had, I would have seen that you had already, and in greater detail, addressed the points I raised. Moreover, your response included many additional points, on which I also concur.

For the most part I agree with everything you said; and am pleased that it resonates well with my own personal perspective, from which I ask that you consider:

1) That, within some limited number of domains, it is important to distinguish between a fact (def. as the quality of something being actual – or real, and by extension “true”) and purported facts (def. as something claimed to be real, a truth claimed). While acknowledging that facts are, in the strictest sense, only beliefs, and hence subject to reconsideration and hence understood most properly as contingently settled or satisfactorily settled, I submit that in discussions relating to evidence, such a distinction may be important. So, your statement “Evidence to me means material that may be relevant to proving whether some fact known

\[4\]The following two messages were sent privately and are included here with permission of the sender. There is some overlap between these and other messages by Mike Joffe and Richard Leary to the list – so there is some inevitable lack of cohesion in the sequence of ideas in the different messages.

57
or contemplated is true ...” causes me to wonder in which sense you meant fact (actual or purported, or both), for in common language, it seems, a known fact – an actual, real object or relation of the real world, as we understand it – would need no (additional?) proof. But, based upon your notion of “a fact at issue,” I think our differences, if any, are semantic.

2) I am convinced, however, that in some number of domains such a distinction is not trivial; and that while adding an entirely new dimension of complexity to the study, pains should be taken to keep it in mind. Further, I believe that in some limited domains the distinction is of paramount importance. In particular, those domains where potential deception is inextricably tied to the credibility of evidence (human systems). And, whether one defaults to falsus in uno, falsus in omnibus, or at the least calls into question the entire construction of a rational scenario and the concomitant factual beliefs, detected deception must always be recognized as a particularly important new evidential item. So, I submit here that the detection of the discrepancy between a contingently settled fact and a purported fact is of great moment to any investigation or hypotheses testing in general, and particularly so in limited domains. First, there seems to be the promise of efficacy in testing for such discrepancies; an area I feel is under-studied. Further, even without the detection of this discrepancy per se, recognition of this distinction leads to, among other things, an alternative form of hypotheses testing on the whole; as perhaps you recognize when you say “and evidence we contemplate could exist if certain explanations run true.” I would be surprised if such an alternative hypotheses testing technique – that is, for example, allowing the logical conclusions of the purported facts to run, so to speak, in their entirety, and to evaluate the known and indicated corroborative evidence which might reconcile such conclusions – has not already been given a name within the investigatory discipline; and I am curious as to what this name might be.

3) For in some sense, I believe, there needs to be some facts that are contingently settled (accepted as true) to allow a logical reasoning to proceed that delivers a sense of reliability (even if one, as stated previously, allows that what is settled now may not be so settled forever). I also acknowledge that it may be possible to construct digital networks that make what now seems remote less-so, and probability networks that demonstrate some level of reliability where many objects or relations are not so “known.” But even with these, the more objects and relations that can be ascertained to have truth probabilities of 1, whether near or remote, the more reliable the whole network. Especially when, as you say, much of investigation is to lower the risk of decision making under conditions of uncertainty. And, unless such probabilistic networks become ubiquitous (which they may, but not in our lifetimes), there will be a need to make ad-hoc decisions outside of such networks, which would require reliance upon objects and/or relations contingently accepted as true, i.e., facts.
4) There is, of course, the additional strategy of accepting contrary objects and relations (a modified “what-if” approach), and considering the overall “best fit” of the whole.

5) Also, in adopting what might be called a “systems approach” to the study of evidence, I have concluded that, depending on your view of the system, motivation to investigate can arise from either “doubt,” or more expansively, perception of risk or “risk minimization.” I believe we have no quarrel over this, if we were to properly understand each other, but mention it only because it may, in certain circumstances, frame things differently and perhaps offer new insights.

Consider the following contingently settled (agreed upon?) facts:

Circumstances find you outside the door to a room. You see Mr. Smith exit the room, closing the door behind him, and saying “All’s well.” You’ve known Mr. Smith your whole life, share common bonds of loyalty and fidelity, and have the utmost trust in him and his skills, ability, and wherewithal. You do not doubt Mr. Smith. With this perspective there would be no reason to seek any evidence to either prove (support?) or disprove Mr. Smith’s statement of fact, which, while quite broad, is accepted as “All’s well.” No doubt, no need for evidence; as you say.

Now, from another perspective (obtained with knowledge of an additional fact):

The room Mr. Smith has just exited contains a button that, if pushed, begins a countdown to nuclear Armageddon. I submit that this additional fact (or as stated, technically a belief), changes much. There is maximum risk. Now, decision (human?) factors such as prudence, circumspection, duty, preservation of self and family, etc. all come into play. Surely you would not be unreasonable to seek evidence that the button had not been pushed.

I’m sure you see, as do I, that this story resolves, strictly speaking, into a question of “doubt,” but I present it to you to highlight the risk minimization (reduction of uncertainty) perspective that may, and likely does, and perhaps even more often should, motivate some investigations. While one might justify seeking additional evidence in the above scenario by the fact (belief) that people are often mistaken, make mistakes of which they are unaware, etc. (i.e., doubt), the motivating objective is risk minimization. Again, it’s not that I suspect you disagree, but only that there may be something more we can all learn if this perspective is considered. A systems perspective, I believe, more properly captures the complexity of the real world, seeks additional information (evidence?) to address known risks and to uncover unknown risks; and to my mind this encapsulates what it means to be alive and to seek an understanding of the world – so that our decisions are made under conditions of everless uncertainty (evermore certainty) – and that to live rationally is to investigate the evidence of life.

6) Lastly, a minor detail. Perhaps you might consider the iterative “search for
3.9. LATER DISCUSSION

evidence and hypotheses for testing” to be spiral rather than cyclical. In addition
to coming back on to itself, spiral carries with it the connotation of moving in a
direction. I refer to 3) above in that, for useful purposes, it seems, some facts must
be settled for the belief to take hold, and perhaps even for the logic to progress.
The direction of the spiral would represent those facts left behind, so to speak, as
settled. And, of course, the sense in which they are so settled is a conventional one,
to the purpose of the task, given all that may then be “known,” subject to revision
(whether by a new item of evidence or by a restructuring of existing evidence) and
not in an absolute sense. **

Sincerely,
James H. Edwards, Publisher

* I am aware that some believe no real world object or relation can be proven
true, but only that the truths are the understandings that remain after alternatives
have been disproven.

* I do not wish to imply that I believe all knowledge discovery is strictly “object-
vive based;” for on the contrary I believe strongly in (and have run informal tests
to demonstrate) the role of serendipity in evidence-based knowledge acquisition.
Instead, I would think that knowledge discovery is objective motivated.

3.9.10 James H. Edwards wrote with title Evidence – Richard
Leary 2

addressing Richard Leary (Monday, Mon, 11 Jul 2005 20:23:42 -0400)

Dear Mr. Leary:

Spiral was an insufficient recommendation. Perhaps a conical spiral, or heli-
cal spring-like shape, or something; just that cycle would seem to not make any
"progress," unless, of course you might mean it’s iterative until the clock stops
and what’s best at that time is it.

Sincerely,
James H. Edwards, Publisher

3.9.11 Terence J. Anderson wrote with title Glossary

addressing Philip Dawid, David Schum and William Twyning (Monday, 11 Jul 2005
10:11:12 -0400)

If you want to pursue Mr. Edward’s idea, that a glossary of evidence termin-
ology be formulated, the entries in the Glossary in Analysis (2d) might provide
3.9. LATER DISCUSSION

a useful starting point.

Terence J. Anderson
Professor of Law

3.9.12 Terence J. Anderson wrote with title Thought for the day: A Glossary

directing Philip Dawid, David Schum and William Twyning (Tuesday, Tue, 12 Jul 2005 10:44:17 -0400)

Phillip,

Supplementing my note, I think you should give serious thought to establishing a mechanism through which the evidence group could develop a glossary of terms that might be useful in an “Integrated Science of Evidence.” The construction of a set of agreed terms and definitions would test (and I suspect, support) the “Twining hypothesis.” One of the underlying frustrations at the Schum Challenge was, at least for me, a sense that language was the most significant barrier to developing a common framework. I would ask Dave and William to comment on this idea.

Terence J. Anderson
Professor of Law

3.9.13 Deirdre Dwyer wrote with title Evidence Challenge Feedback

directing Philip Dawid and Richard Leary (Fri, 15 Jul 2005 15:56:34 +0100 (BST))

Dear Richard,

I am very interested in your discussion with Michael Joffe which seems to bring us into clear philosophical waters. I am in a DPhil student in Law, about to submit, with a first degree in Philosophy. I would be grateful for any assistance/clarification that you may be able to give specifically on how and why you think evidence might be infinitely divisible.

More generally, your discussion with Michael Joffe appears to turn on how we are using the word ‘hypothesis’, and I thought I might make a suggestion on what, I think, is going on, to see whether that takes is any further, or stimulates any more thoughts. In particular, I think it touches on the fundamental question of how evidence may be used differently for different types of task.
When Michael Joffe says ‘hypothesis’ he appears to mean it in a sense that is deductivist in general, and Popperian in particular. The deductivist part is to say that we have a formal statement about how the world might operate, and we test the validity of that statement against real data about how the world is operating (broadly, for a Hempelian this will verify the validity of our hypothesis, while for a Popperian it will only show that our hypothesis has some validity, because it remains to be disproved). This, as I understand it, is the sense in which statisticians use ‘hypothesis.’ The Popperian part is to say that it does not matter where that hypothesis came from, e.g. we might wake up in the morning after a vivid dream, with a wonderful (but completely weird) new hypothesis, which we can then go off and test. I think that Popper’s point is that whether the hypothesis is true or not is in no way affected by how the hypothesis is arrived at. In practice, we tend to run with hypotheses that are similar to previous hypotheses, and conform to an already recognized (if not agreed) explanatory framework. This makes our investigations more efficient, but it also constrains their potential (hence the perceived value of ‘thinking outside the box’).

You, on the other hand, would appear to be an inductivist. I could take your view to be that the hypotheses spring from the intrinsic nature of the data in question. I would have difficulties with that view, and think, from what you say, that he you might mean that our existing model of how the world works, and possibly what is going on in the current situation, provides us with questions to ask. Thus, when a police officer arrests a woman charged with murder, the officer does not as a matter of routine search the suspect’s body for unusual marks, such as moles, as evidence of a pactum diabolum, nor does the officer make enquiries of neighbours about black cats or toads seen in the vicinity, which would have been credible lines of enquiry into the 17th century.

One could, if trying to find common ground between yours and Michael Joffe’s approach, say that you are proceeding by taking much smaller hypothesis-testing steps, so that for each fact you encounter, you come up with a hypothesis, and say ‘if this (a) were true, then I would expect to find b’ and go off and look for b. But I suspect that something more fundamental may be going on. You appear to be concerned with finding an explanation for a specific set of events, while Michael Joffe appears to be looking for an explanation that holds true for recurring sets of events. The former is a common problem in law (broadly defined) and security work, while the latter is more the world of the scientist. The inferential differences between those two types of endeavour would, I suggest, at least partly explain some of the difficulties that courts have with experts (and vice versa). For a discussion of some of these differences, see the start of Hart and Honor’s Causation in the Law. One of the features of a specific event is that the data are finite in number, so having generated an hypothesis using those data, there are then no separate data with which to test the hypothesis. It has also been suggested
3.9. LATER DISCUSSION

(eg myself passim in Int J Evidence and Proof 2005), particularly with regard to why lawyers should most certainly not defer to police and intelligence services on questions of fact, is that fact investigators tend to (but not always do) take a highly subjective view of data, looking for the information that supports their hypothesis, closing off other avenues of enquiry and not looking for evidence to refute their hypothesis. If research scientists are using evidence in a different way to fact finders (including forensic scientists), then how (if at all) does this affect a theory of substance-blind inferential reasoning? Schum’s model appears flexible enough to handle both approaches, and I am sure that there are common elements between them, but perhaps we should also identify any specific differences, to understand how these two approaches might fit together in a common science/theory of evidence.

Any thoughts that you might have give would be greatly appreciated.
Very best wishes,
Deirdre Dwyer

This seems to have been the last post on the subject
Chapter 4

Schum’s response

Some time after the email exchange had ended, posted his response to the general exchange, the papers and the presentations of the day. Many of the last are available in some form at http://tinyurl.com/b7854.

A REPLY TO THE “SCHUM CHALLENGE” AT UCL D. Schum 6 September, 2005

4.1 Introductory Comments

After several unpromising starts this summer, I have finally written this belated response to what Phil Dawid called the “Schum Challenge”. I thoroughly enjoyed our discussions when we met at UCL on 7 June, 2005 to consider this challenge. And I certainly learned very much about your own interests in evidence and inference. I was also most grateful for your patience as I was trying to cope with speaking difficulties I encountered as a result of the cortical insult I suffered last April. Our discussions turned out to involve a variety of inferential and decisional issues that weren’t necessarily associated with anything I have said or written. So I looked upon our discussions as involving a challenge we all face in trying to say useful things about evidence in the Leverhulme-sponsored research to which many of us are committed. Mark you, I mention this not to avoid taking heat for any ideas I have generated. I have said certain things about evidence that I fear have been misleading to some of you. This is where I will begin my comments, which are based upon the specific presentations some of you made on 7 June and in the subsequent electronic dialogue that has taken place since we met.
4.2 Substance-Blindness

My use of the term “substance-blindness” has caused some trouble for which I am partly to blame. However, some of you have used this term in ways I have never used myself. But I believe I can clear up some of the vagueness in my own use of this term. I coined this term some years ago with reference to a classification of recurrent forms and combinations of evidence we all encounter in our inferences based on evidence. The problem I faced was the following: Given the virtually unlimited substantive variety of evidence, how can we ever say anything general about evidence? Being able to say some general things about evidence is useful for many purposes and was something I needed to do in much of my work. I add here that Phil has also encountered this same problem. In various ways we have both been involved in abstract studies of evidence in the process of identifying general and recurrent evidential patterns for study and analysis.

For various reasons I decided that evidence could be usefully categorized on inferential grounds rather than on any grounds related to the substance or content of evidence. I was initially encouraged to adopt this approach as a result of works I found in the field of law. However, I became convinced that further evidential distinctions could be made, especially when we consider the various forms testimonial evidence might take and when we also consider basic recurrent combinations of evidence. Evidence has three major credentials known to me: relevance, credibility and inferential [probative] force, strength or weight. Two of these three credentials form the basis for my classification of individual items of evidence: credibility and relevance. For several reasons, inferential force or weight does not provide a useful basis for categorizing evidence. The force of evidence depends on judgments made by the person evaluating the evidence and involves consideration of relevance and credibility matters. An additional problem is that the force of evidence can be interpreted, assessed and graded in several different ways.

I pause here to reflect on a comment made during our meeting. Someone said that that relevance, credibility and inferential force are “vacuous” matters for us to be concerned about in our work on the Leverhulme research. I couldn’t agree LESS with this comment. Studies of these credentials of evidence are basic matters that are the subject of lingering controversy. There is still so much to be learned about these credentials; the final book has not yet been written on any of them. Perhaps we can write a few new chapters in this book as a result of our joint efforts.

In several works, and on two occasions at UCL, I showed a two-dimensional categorization table for individual items of evidence. The rows identify common forms of evidence and are based on how the user of the evidence stands in relation to it. I illustrated how this issue involves the manner in which the credibility of
the evidence is established. For example, there are different questions to be asked about the credibility of tangible evidence [directly observable to the user of the evidence] than there are about the credibility of testimonial evidence [coming to the user from someone else]. We have to take account of how this other person obtained the information she/he reports; and we also have to allow for various ways in which people hedge their testimonies. Either tangible or testimonial evidence can reveal the occurrence of an event [in which case it’s said to be positive evidence]; or it may reveal the nonoccurrence of an event [in which case it is called negative evidence]. Missing evidence, either tangible or testimonial, can itself be evidence under certain conditions. Finally, we all encounter evidence of various sorts whose credibility we would never be asked to establish. For example, we make frequent use of tabled information, such as the times of high and low tides at certain locations, that would be widely accepted as being credible information.

I believe this credibility-related part of my categorization of evidence is fairly uncomplicated except that we often observe mixtures of these forms of evidence. When these combinations occur we often have some very difficult credibility assessment tasks on our hands. For example, we may have a tangible document whose authenticity and accuracy is of concern. But this document records the testimony of a human observer whose credibility is at issue. So we have this person’s veracity, objectivity and observational sensitivity to consider.

I think the columns in my categorization table that concern relevance may be the major source of confusion about what I have meant by the term “substance-blindness”. The question here is: How does the evidence stand in relation to matters at issue in your inference? I’ll try my best to give a clearer account of what this dimension entails. In the table I showed you, I first distinguished between directly relevant and indirectly relevant [ancillary] evidence. In the discussion of relevance that took place this summer, I side with Hasok who argued that relevance concerns the bearing of information on hypotheses we entertain. Hasok will be pleased to learn that I have often used a quote by Charles Darwin that I found in a logic text written by I. M. Copi [Introduction to Logic, 6th ed. 1982]. Darwin said that information must be for or against some view if it is to be of any service at all. In other words, information becomes evidence when its relevance is established on some matter at issue.

But there are two species of relevance that need to be identified. An item of information is directly relevant evidence if it can be linked to some proposition [hypotheses] you are trying to prove or disprove by a defensible argument or chain of reasoning. Ancillary evidence concerns the strength or weakness of links in such reasoning chains. In short, ancillary evidence is evidence about other evidence and its credibility and the strength of its bearing on hypotheses. In the table I showed you I made a distinction between two forms of directly relevant evidence: direct evidence and circumstantial evidence. I made this distinction
4.3 WHEN SUBSTANCE IS NEVER BLIND

in earlier works I have written. But Terry Anderson and William Twining have urged me to drop the distinction between direct and circumstantial evidence in our recent work on evidence. The basic reason is that direct evidence is said to go in one step from the evidence to some proposition at issue. If the evidence is credible, that settles it and so the proposition is true. Circumstantial evidence, on the other hand, is always inconclusive regarding this proposition, even if the evidence is credible. Trouble arises in situations in which we can decompose the linkage between the evidence and the proposition at issue. What the evidence is direct upon then depends on how we have decomposed this linkage.

Finally, my categorization of the ways in which combinations of evidence can exhibit patterns of harmony, dissonance, and redundancy are readily observable in any context and should be quite uncontroversial. Now, as far as substance-blindness in my categorization of evidence is concerned, we reach a FULL STOP. I have said nothing about what is involved in the process of establishing the relevance and the inferential force of evidence. What is involved here is NOT substance-blind, as I will now try to explain.

4.3 When Substance is Never Blind

Clearly, what evidence you gather and the hypotheses you generate during the discovery or investigative stages of your work, and the nature of the arguments you construct that link evidence to what you are trying to prove or disprove from it, are all matters that directly concern the substance of your inferential work. Notice that my use of the term “substance-blindness” in categorizing evidence says nothing at all about the probabilistic issues you encounter in assessing the force of your evidence on hypotheses you entertain and in reaching final or tentative conclusions from your assembled evidence. I have been surprised on several occasions in some of your writings, and in our discussion, by your suggestions that I have somehow constructed a substance-blind theory of inference, or as Deirdre Dwyer said in a recent e-mail, that I have constructed a “substance-blind theory of probabilistic reasoning”. I have never made such claims. Trying to develop a taxonomy of all the mixtures of the complex intellectual reasoning that such activities involve would be a fruitless and a silly task.

In short, all I have said is that evidence is either directly or indirectly relevant on matters of concern to you; here is where the full stop occurs as far as what my categorization scheme says. What means you choose to establish such relevance

4.3. WHEN SUBSTANCE IS NEVER BLIND

are entirely up to you and will depend on the substantive nature of your work. As you know, Terry, William and I have been arguing for a long time that the methods for argument construction Wigmore developed years ago in 1913 are very useful is allowing us to make sense, in a systematic way, out of masses of different forms and combinations of evidence. Wigmore’s methods are very similar to the argument construction methods described years later by the philosopher/logician Steven Toulmin. But Wigmore captured more of the complexity of evidence-based reasoning than Toulmin has ever done.

Though Wigmore wrote for an audience of practitioners and scholars in the field of law, his methods apply regardless of context. So we might say that Wigmore’s, and Toulmin’s, methods for argument construction are substance-blind. However, the exact nature of the arguments you construct, using either of these methods or any other method you choose, will depend on the substance of your evidence and how you believe this evidence is linked, directly or indirectly, to hypotheses or propositions of interest to you. Argument construction in any context is an exercise in imaginative and critical reasoning. No one can tell you about the exact nature of the linkage between your items of evidence and propositions of interest to you. You first must generate or imagine what the steps in your argument might be; this is the imaginative part of the task. Some will say that this involves abductive reasoning. But then you must subject your arguments to careful analysis in order to detect any disconnects or non sequiturs in them. This calls for critical reasoning on your part.

I come now to probabilistic considerations that are surely dependent on the substantive nature of our work. In some instances the events of concern in our inferences arise as the result of processes that are replicable or repeatable. In such instances we may obtain relative frequency or statistical estimates of the probability of these events under various conditions. But in so many other situations the events of concern to us are unique, singular or one-of-a-kind for which no statistical estimates are ever available. In such instances our only recourse is to make subjective judgments of these probabilities. In all views known to me, the inferential force of evidence is expressed in probabilistic terms. But, as I noted above, there is considerable controversy about what the force of evidence means and how it should be assessed and graded. I devoted an entire chapter of my book on evidence to a discussion of alternative views of the inferential or probative force of evidence. In the work you do, you may of course never even consider what these specific formal views tell us and instead adopt your own interpretations of evidential force or weight and choose your methods for assessing and grading it.

4See note 1 at pages 200-269.
Conclusions based on evidence are necessarily probabilistic for the following five reasons, which may occur in various mixtures depending on the substance of your work. Our evidence is never complete and is usually inconclusive in the sense that it is to some degree consistent with more than one proposition of concern to you. Further, our evidence is often ambiguous or imprecise; we can’t determine exactly what it is telling us. Evidence is commonly dissonant to varying degrees; some of it points in one direction and some of it points in another. Finally evidence comes to us from sources [including our own senses] that have any possible gradation of credibility shy of perfection. In coping with these characteristics of evidence, that generate doubts or uncertainty, we may or may not make use of any one of several formal probabilistic systems such as Bayes’ rule in conventional probability, Dempster’s rule for non-additive probabilities, Jonathan Cohen’s Baconian probabilities for eliminative and variative inferences, or Lotfi Zadeh’s system of fuzzy probabilities for coping with imprecision in our evidence, hypotheses, arguments and in our frequent verbal expressions of probabilities. Though all of these formal systems are substance-blind, their application in particular instances depends entirely on the substantive details of the inferences of concern to us.

I think I have said enough about what I have said, and not said, about substance-blindness. I’ll now consider some things you have said about evidential and inferential matters of concern to you in your substantive areas of interests.

### 4.4 Epilepsy Services in Ferndale

Trish, Jill, Janet and Emma first posed some very interesting and difficult problems in the cover story they invented to illustrate their concerns about “evidence-based” medicine and other forms of evidence available to health care professionals. I was a bit surprised, however, at their assertion during our meeting that they never entertain hypotheses. As I remember, this is what occasioned the subsequent dialogue concerning evidential relevance being linked to hypotheses being entertained. But in their story about Dr. Bob, Director of Health Improvement in Ferndale and chair of the Priorities Forum, there seem to be hypotheses latent all over the place. This occasioned the Wigmorean exercise that Terry Anderson generated at our meeting that involved an entire hierarchy of linked hypotheses. Some hypotheses concern what actions Dr. Bob will take or recommend during the meeting of this forum. Other hypotheses concern reasons why, as Dr. Bob says, members of the forum seem to be hard-nosed in their preference for evidence-based studies of the costs and benefits of various forms of health services and their apparent sneering at what they regard as anecdotal evidence obtained from individual patients. Pauline and Fran, nurse professionals who treat
epilepsy patients in Ferndale, are important stake-holders, as are the epilepsy patients themselves, in the deliberations and choices made by the Priorities Forum.

I had thoughts about several matters suggested by the Ferndale example. The first concerns one of the most difficult inferential problems I know about; examples of it abound in law and intelligence analysis and it appears in the medical example that Trish and her colleagues generated. This problem occurs when we attempt to draw conclusions about the inferences or decisions others will make or have made. In some sense this involves our making inferences or choosing courses of action by proxy for someone else. Imagining how others will think or act is commonly a difficult exercise. For example, attorneys will attempt to infer the conclusions a magistrate will reach from initial evidence in reaching his/her decision about whether there is “probable cause” for holding a person suspected of committing a crime. Following the tragic events in London reported on BBC news, some so near UCL, I read so many different accounts of possible reasons why the terrorists chose the particular times and places to cause so much misery in London. Pauline and Fran have certainly tried to put themselves in the minds of members of the Priorities Forum in their attempts to predict what decisions they will make concerning the allocation of funds for methods of treating epilepsy patients in Ferndale.

In some of the discussions that followed our meeting on 7 June, that concerned the Ferndale example, mention was made of the matter of distinguishing between inferential judgments and value [or utility] judgments. Here is one interesting issue that I think has not been discussed as much as it should be. To what extent, if any, should the value we place on a decision consequence be related to our assessment, based on evidence, of the probability that this consequence will occur? Stated the other way around, should our probability that this consequence will occur be dependent in any way upon the value we place on this consequence? Phil and I have had several discussions concerning what decision theorists have said about whether their canons for “rational” decision-making assume the independence of probability and value judgments. It’s not clear what many theorists believe about this issue. We can all think of examples in which persons are pessimists who often believe that unfortunate consequences are the ones most likely to occur. And, we encounter optimists who often believe that desirable consequences are the one most likely to occur. I’ll bet Nigel Harvey has many interesting things to say about this issue. I bring all of this up because it bears upon what inferences can be made by Pauline and Fran concerning what decision the Priorities Forum will make concerning programs for the care of epilepsy patients in Ferndale.

This brings up my final point about the medical care example; it is one that allows me to revisit substance-blindness and how it cant be related to matters concerning the force of evidence and how it is assessed and graded. Dr. Bob tells
4.5. ACADEMICS TEACH BETTER IF THEY ARE ALSO RESEARCH ACTIVE

us about the “proper weighing up of the evidence”. He proceeds by asserting how the Priorities Forum will grade randomized controlled trials as the best evidence, followed by cohort studies. At the bottom of the evidence value hierarchy comes anecdotal “stuff”. We would all be quite at home debating the force of evidence on any matter. This immediately raises the very important matter of one’s standpoint, as has been so carefully and extensively discussed by Terry Anderson and William Twining. What roles we play, what stages in what process we find ourselves, what objectives we entertain, and what information we have access to will certainly influence our judgments about the relevance, credibility and force of any kind of evidence in any situation. There is no substance-blindness here. We aren’t told what standpoints are represented by any of the members of the Priorities Forum, but we can easily infer that they are different from the standpoints of Pauline and Fran. It seems that persons who daily treat the epilepsy patients in Ferndale, such as Pauline and Fran, have a job on their hands in trying to make all members of the Priorities Forum, including Dr. Bob, rethink how they have ordered the force of evidence concerning patient care in Ferndale.

4.5 Academics Teach Better if They are Also Research Active

This is the proposition or hypothesis, I’ll call it H, that Steve Rowland asked us all to consider when we met on 7 June. He reviewed some of the past studies in the UK that have been conducted to test this hypothesis. The results of these studies seem to have been as inconclusive as those obtained in similar studies in Australia and the USA. Then Steve considered a variety of inferential and methodological problems encountered in research on this hypothesis. He argued that even if we had uncontroversial measures of teaching effectiveness \( T \) and research accomplishments \( R \), we would still encounter difficulties in associating \( T \) and \( R \) evidence in tests of \( H \). The basic question he then asked was: Can a science of evidence that is “substance blind” ever get a foothold on the question of the relationships between teaching and research?

In my remarks following Steve’s presentation I was obliged to note that I have never argued that any “science of evidence” could be substance blind. Entirely abstract studies of evidence would certainly be unable to capture all of the richness of evidential reasoning in our different disciplines. New evidential and inferential subtleties or complexities are continually revealed in the substantively different research we all do. William Twining has been very convincing in his arguments that any science of evidence must be interdisciplinary in nature. Later, in my closing comments in this reply I’ll return briefly to some things I will say in a
subsequent paper on a science of evidence.

I am almost finished reading Steve’s excellent book: *The Inquiring University Teacher* [The Society for Research into Higher Education & Open University Press, 2002]. I wish this book had been available 40 years ago when I began my academic career. The question posed by hypothesis H must have occurred to all of us at some time or another during our academic careers. To what extent do our students benefit from all the hours we spend trying to advance knowledge in areas that interest us? I was no more successful than anyone else in supplying instant answers to the questions Steve raised about how we could use measures such as T and R [even if they existed] in inferences about H. But I have thought quite a bit about Steve’s questions and have a few suggestions to offer him.

As stated above, hypothesis H is a fine example of a vague, imprecise, undifferentiated or fuzzy hypothesis. This imprecision is revealed by the words “teach better”. Natural questions arise such as: How much better?, and: In what ways is it better? Steve is not to be faulted by the fuzzy nature of his hypotheses since, if we had direct reasonable measures of how much better and in what ways it is better, Steve would have been able to assert a more precise hypothesis. Dealing with vagueness is something frequently encountered in decision analysis. I’ll first give an example that comes from the assessment of value or utility and then to relate it to Steve’s inferential problem involving the vagueness of hypothesis H.

In decision analysis we are concerned about what attributes we should identify for the consequences of actions we contemplate. Another way of describing these attributes is that they concern the value dimensions of consequences. The natural question is: Where do these attributes come from? One answer is to say that the attributes of consequences are measurable indications of the extent to which they correspond to our decision objectives. To put a value on any consequence we first have to say what we want as a result of the choice we have to make. This involves a statement of our goals or objectives. A very useful device comes in the form of what is called an objectives hierarchy. In most instances our major stated objectives are vague in the same way as is hypothesis H. What is done is to deduce levels of subsidiary objectives from these major objectives until we find lower level objectives, consistent with our major objectives, that are sufficiently precise and that are observable or measurable. These *measurable* indications then become attributes of our consequences. There is a very large literature on objectives hierarchies.

But there is a great similarity between an objectives hierarchy and Wigmore’s hierarchies of hypotheses in what are called today inference networks. His methods are usually employed for the purpose of making sense out of accumulating masses of evidence. But in discovery and inferences we frequently encounter hypotheses in search of evidence at the same time we encounter potential evidence in search of hypotheses. I will dwell on the first of these searches in Steve’s case.
What Steve has asserted as H is what Wigmore would call an ultimate probandum [U], a proposition or hypothesis of basic interest to us. From this ultimate probandum we attempt to deduce penultimate probanda [P] that form our major lines of argument on \( U = H \), in Steve’s case. The field of law is the only one known to me in which in which the P’s are specified for any given U. Stated another way, law is the only area that specifies materiality criteria in the process of proof. But Steve has no such advantage. To prove H at some acceptable level he must deduce what he believes to be necessary and sufficient main lines of argument on H. Each of these main lines of argument will have multiple subsidiary convergent lines of argument in which there will be many unobservable propositions or hypotheses; Wigmore would refer to them as interim probanda. This hierarchy proceeds downward until we find propositions that have observable evidential states.

If Steve can begin by coming up with main lines of argument on H and then imagine what lines of evidence would bear on each of these arguments It’s just possible that he might be able to make H more precise. One frequently overlooked virtue of the structural methods described by Wigmore is that they can enhance the process of discovery or investigation. It occurs to me that there are many potential lines of evidence and arguments in what Steve has already written on this subject. But this suggestion is entirely structural in nature. It requires the substantive expertise Steve possesses in order to bring life to it.

### 4.6 A Historian’s Perspectives on Evidence

I read with particular interests two documents that Jason Davies sent to me before our meeting on 7 June. One was entitled: *Why not a substance-blind approach to evidence*. The other, which I presume is to be the first chapter in a longer work Jason is writing, is entitled: *The Failure of Knowledge: The Place of Ignorancesa*. One reason why I have taken Jason’s work so seriously is that I am on record as having paid the greatest attention to what historians have told us about the properties, uses and discovery of evidence. In my work with Jay Kadane on the Sacco and Vanzetti evidence, we were obliged to consider elements of a historical standpoint in our work. The reason is that this case was decided years ago in 1927 when Sacco and Vanzetti were executed for a crime many persons believe they did not commit. You can see for yourself how seriously we took the work of historians in our analysis of the evidence in this case. In the act of writing this work I learned so much from a variety of noted historians whose works should be read by anyone having concerns about the evidential foundations of inference.

---

I agree with so many of the remarks Jason has made concerning the uses and misuses of information and evidence. He is certainly correct in his focus on judgmental issues, though I did not follow all of the distinctions he mentioned in his “charismatic-institutional” account. I also agree with his accounts of how opinions about the relevance and force of evidence depend on context and may change over time. As far as I know, Queen Elizabeth II has never touched any sufferer of what used to be called “The King’s Evil”. It was thought in Samuel Johnson’s time that the reigning monarch could cure victims, having the disease that also used to be called scrofula, by her/his simple touch. We now understand that this disease is the result of tuberculosis of the lymph glands.

My major quibble with what Jason has said involves how he has linked the matters of concern to him with what he calls the “substance blind approach to evidence”, which he links to me. I’ll say again for the record that my use of the term “substance-blind” only concerns what kind of evidence you have. It does not concern what use you are making of it. I hope I made this clear in the section above that I titled “When Substance is Never Blind”. Jason apparently has some concern that adherence to what he calls a substance-blind approach to evidence will lead to a slavish form of inference in which any hope of obtaining wisdom will be lost. I hope that in his subsequent thoughts and writing about evidence he will no longer relate his concerns to anything I have said about evidence. My hope is that, if we are better aware of what kinds of evidence we have, we might make better use of it in our efforts to obtain understanding and, possibly, wisdom.

4.7 Ethnicity Issues in the Use of Evidence

Tony Gardner-Medwin has raised some issues in his presentation that are currently of great substantive importance in many contexts. Suppose a person is identified as being a member of a reference class whose attributes involve racial or ethnic details. Is evidence about these details always relevant and admissible regardless of the context in which inference are made concerning this person? The answer to this question seems to be “no”; context does matter. In inferences and decisions concerning this person’s health, information concerning this reference class can often be relevant and of significant value in treating this person and would in most cases be perceived as admissible by her/his physician. Members of this reference class may show a high statistical incidence of having a certain disease that is also consistent with many of this person’s symptoms.

But suppose a person faces a criminal trial and belongs to a racial or ethnic reference class whose members have been statistically associated with crimes of the sort with which this person is now charged. Should this reference class information be viewed as relevant and admissible evidence when this person comes to
4.7. ETHNICITY ISSUES IN THE USE OF EVIDENCE

In our courts in the USA there is a Federal Rule of Evidence [FRE 403] which outlines conditions under which relevant evidence may not be admissible. I would guess that there are similar restrictions in English law [William Twining can help us here]. The point is that perceived relevance does not guarantee its admissibility. I have two points to make regarding Tony’s comments on these matters. The first is a minor quibble about what he says about Bayes’ rule, which he uses to make several points is his arguments. The second comes from the writings of some prominent legal scholars in the USA how argue that reference class information is neither relevant nor admissible in trials at law.

My minor quibble concerns Tony’s description of what is called the log-odds form of Bayes’ rule. He says that the log of posterior odds favoring some hypothesis are equal to the log of the prior odds on this hypothesis plus the sum of independent evidence weights. I have underlined the term independent since this is what my quibble concerns. Bayes’ rule contains within it perhaps the most powerful mechanism I know about for capturing evidential subtleties or complexities in probabilistic reasoning. It is a concept called conditional dependence. Among the things this concept allows us to do is to show how the weight of any item of evidence may be dependent on other items we have. Using expanded forms of Bayes’ rule we can show the nature of these dependencies and what consequences they have in our inferences. In short, Bayes’ rule works even when the weights of evidence are not independent. This is where it allows us to capture so many evidence subtleties that lie just below that surface of even the “simplest” probabilistic reasoning tasks [if there are any such simple tasks].

Paul Bergman and Al Moore have written a very perceptive paper on matters of the sort that Tony has mentioned. This paper is also highly entertaining, unless you are a hardened and shameless Bayesian. The F-Word in the title of the their paper does not stand for what you might be thinking it does, it stands for frequency. Here briefly is what Bergman and Moore argue. Any form of reference class information resting on statistical relative frequencies, however they are determined, are not relevant and, therefore, not admissible in criminal charge against defendant in a court of law. The only thing at issue is what this defendant did or did not do. What all the other members of this reference class did or did not do to is totally irrelevant to what defendant did or did not do. This single defendant is the one now facing trial and the only relevant evidence concerns him/her and not any members of any reference class to which this person might belong. For reasons not having anything to do with reference classes Jay Kadane and I had to respond to Bergman and Moore in our Sacco and Vanzetti book. We took issue

---


7See note 5 at pages 261-265.
with how they argued that the force of evidence might be judged in the absence of any statistical information.

4.8 In Conclusion

Phil gave me eight pages to write this reply. I am now in the process of writing a paper on a science of evidence, even though some of you might object to use of the term science. We have told the Leverhulme Foundation that we are at work on a science of evidence. It would be somewhat embarrassing for us to claim that there is no such thing as a science of evidence. I have certainly learned very much from the enjoyable interactions we have had; but they have not made my next task any easier. I will await your comments when I finish this paper.
addressing David Schum and his response (Tuesday, October 4, 2005 17:19:15)

Dear David,

Thanks so much for your report here. I’ve given it quite a bit of thought, but it needs more.

Can I just ask you about one point?

In your comments on my ‘challenge’ – about teaching and research – you note that my ‘Hypothesis’ (that academics teach better if they are also research active) is “vague” and your helpful suggestions indicate how it might be possible to come up with a more “precise” rendering of H.

I am pondering on what is meant by ‘vagueness’ and ‘precision’ in this context, and whether there may be some circularity in these meanings.

When I articulate the meaning of T (good teaching) I am able to be quite articulate about this. It is something about which I have given a lot of thought (although much more, I’m sure, is needed!). But an articulate, well developed, conception of T is not the same as a ‘precise’ one if, by ‘precise’, we mean something that is describable in terms of measurable, observable features. A Wigmorian kind of inferential reasoning requires precision in this sense. From my understanding, a precise understanding of T in this sense would be a reductive one which failed to capture the subtleties of the situation.

The situation is further compounded if H concerns a conception of T which is to reach general agreement. For, even of all those whose agreement is sought are also articulate, their articulations are likely to be different. If we are to reach an agreement about T, the likelihood as that what is agreed will lack articulation and thus be vague in the sense of amounting to a ‘lowest common denominator’ of consent. I would be disparaging about this kind of vagueness

Thus, in trying to get somewhere in my exploration of H, I am looking for articulacy rather than precision. It is, perhaps more like (but not quite like) the poet who is trying to articulate the nature of (say) love, rather than the psychologist who is trying to reduce it to observable features. I think ‘vagueness’ is not a property of the good poem even though it lacks ‘precision’ in the sense you mean. Its precision consists in quite different features which include, for example, its ability to convey ambiguity and confusion with clarity.

This leads me to wonder whether the notions of vagueness and precision that are appropriate for a Wigmorian kind of analysis may be unable to capture the richness of the ways in which we do, in fact, draw upon evidence in reaching judgements. But this is, perhaps, taking a very large leap. A circularity may
arise if the conception of ‘precision’ and ‘vagueness’ are merely a function of the analytic method, which demands measurability, rather than more appropriate attributes.

I’m not sure if I have made my problem here clear. Nor am I sure if it affects the suggestions you make which, of course, fully acknowledge that such conceptions as T may be very important while still be (in your terms) ‘vague’.

I will leave it there for the moment, but will welcome any thoughts you may have.

Very best wishes

stephen
Chapter 5

Appendix: Email Conversations

5.1 Formatting Email in Print

5.1.1 Difficulties of Genre

There are many difficulties in compiling a set of conversations carried out on email and this document inevitably reflects some of those: it is always difficult to represent properly overlapping conversations – even a chronological transcript cannot do justice to the extent of the overlaps when people refer back to earlier messages than were being replied to. Chronology is much less reliable than in a conversation. Thus some indication of precisely what a poster is addressing needs to be included. Other factors include the accidental omission of ‘reply to all’ – some emailers will do this automatically, others do not. Thus several conversations spiralled off the main body and became ‘one on one’ conversations, which were themselves often (deliberately) brought back to the main group (which itself varied: sometimes it was the full Evidence list, but at others was the Personnel actively involved in the Challenge). This document reflects those conversations that Jason Davies, James Edwards and Philip Dawid had ‘off-list’ but the fact remains that many (potentially more interesting) conversations may have taken place off-list (people may have spoken more frankly, especially about each other.) Hopefully what is here is intelligible to readers – where it seemed important, missing messages were requested (and obtained) to make sense of the various threads.

5.1.2 Technical Difficulties

Top-posting

On the web as a whole, ‘top-posting’ is not encouraged. This is where a poster includes the previous message (usually in its entirety) and puts their answer at the
5.1. FORMATTING EMAIL IN PRINT

top. This breaks down when conversations became complex. Unfortunately, most emalers default to top-posting. Thus, it is impossible to tell how much of the quoted message is being addressed. It also becomes necessary to strip out large amounts of material that would otherwise be duplicated repeatedly.

The standard way of representing the problem of top-posting is like the following paragraph:

Because it messes up the flow of reading.
>How come?
>>>I prefer to reply inline.
>>>>>What do you do instead?
>>>>>>No.
>>>>>>>Do you like top-posting?

Some of the conversations actually ended up more like this:

Because it messes up the flow of reading.
>>>>>What do you do instead?
>>>>>>No.
>How come?
>>>I prefer to reply inline.
>>>>>>>Do you like top-posting?

Or even (where internet quotes were not used by emalers):

Because it messes up the flow of reading.
What do you do instead?
>>>No.
>How come?
>>>I prefer to reply inline.
>>>>>>>Do you like top-posting?

This requires careful comparison with earlier messages to ascertain who said what, and when. This record is thought to be reasonably accurate – there are no obvious non-sequiturs and the original messages were identified, and the quote levels adjusted accordingly. In addition, to make the document readable, messages are represented as if ‘bottom-posting’ had been used – ie both the original message was trimmed to become indicative rather than exhaustive in reference
5.1. FORMATTING EMAIL IN PRINT

to what was being responded to and the quoted material was moved so that it was above the reply. This was necessary to avoid forcing the reader to go back and forth: the document should make sense in the order in which it is presented, though the reader may have to refer back to get a fuller version of what is being addressed.
Chapter 6

Appendix: Evidence Personnel

There follows the list of those who comprise the list referred to as ‘Evidence Personnel’

John Adams
Terry Anderson
Hasok Chang
Andrew Chesher
Phil Dawid
John Fox
Tony Gardner-Medwin
Mark Geller
Hidehiko Ichimura
Jason Davies
Jill Russell
David Lagnado
Richard Leary
Simon Lee
Janet McDonnell
Nigel Harvey
Stephen Rowland
Roy Bailey
David Schum
Bob Sharples
Trisha Greenhalgh
William Twining

This collection edited by Jason Davies.
Index

Emma Byrne, 27, 36, 38
Hasok Chang, 39
Jason Davies, 20, 24, 30, 34, 37
Philip Dawid, 5, 8, 10, 11, 13-15, 18, 28, 29, 32, 35, 43, 47
Deirdre Dwyer, 46, 61
Tony Gardner-Medwin, 6, 8, 10, 43
Trisha Greenhalgh, 22, 28, 34, 42
James H. Edwards, 56, 57, 60
Nigel Harvey, 21
Terence J. Anderson, 48, 60, 61
Michael Joffe, 50, 51, 54, 56
David Lagnado, 23, 25, 40
Richard Leary, 51, 52, 55
Janet McDonnell, 9, 26, 30
Stephen Rowland, 7, 19, 30, 41, 77
Jill Russell, 42
David Schum, 64