Paper 1: Eliciting Experts’ Knowledge: A Comparison of Two Methods

FABIANA SCAPOLO1 AND IAN MILES2

Abstract

Foresight and technological forecasting often involve applying qualitative methods for eliciting and using experts’ judgement, e.g. in the form of much scenario workshop activity. Quantitative methods using expert judgement are also very common in these exercises, with Delphi and cross-impact being among the best-known. But there is a marked shortage of clear guidance as to when and where - and how - particular methods can be useful, what their costs and benefits are. Much of the literature is very superficial, presenting results of studies with little discussion of the practical choices that were made. The more detailed accounts tend to examine single methods, usually from a partisan position. And it emerges that even detailed accounts typically leave a lot to be desired.

Different forecasting methods naturally will not offer the same outputs, even when they are drawing on evidence and perspectives that should be yielding broadly similar results. (We would hope at a minimum that the results will not be contradictory.) Nor do they involve the same inputs in terms of data requirements, forecaster time and skills, technology support, etc. With a great deal at stake in the decision to undertake one or other method, understanding which criteria can be used to select among methods?

This question underpins the present paper. It presents and discusses the results of a comparative analysis and evaluative implementation of two methods, when they are applied on the same area. These two methods - Delphi and cross-impact analysis - both use expertise, and apply questionnaire approaches to elaborate their views of essentially the same topic (future European transport systems). What appeared to be as a straightforward task of comparing and contrasting methods turned out to be extremely complicated. While numerous challenges had to be confronted in implementing the methods so that they would be comparable, the key issues revolved around the immense lacunae in documentation of these two methods. What is good practice, in terms of all sorts of issues in design and implementation of the methods, has been very poorly explicated. It remains largely a matter of tacit knowledge, so that newcomers to the use of either method are liable to find themselves adrift, unless they can gain apprenticeship with experienced practitioners.

The study generated a great deal of insight as to the requirements of the methods through the exercise of attempting to give each of them a “good run” at addressing the issue of future transport systems. Assessment and evaluation of the results of the implementation of the two techniques were also accomplished, mainly through two evaluation questionnaires completed by the experts that participated in the study. This evaluation examined their views of the usefulness of the results for the field of application, and obtained their critical views of the methods as implemented. The study cannot provide a definitive test of Delphi vis-à-vis cross-impact analysis, because all we have been able to assess here is one particular implementation of each method –

1 Dr. Fabiana Scapolo, European Commission Directorate General Joint Research Centre – Institute for Prospective Technological Studies (DG JRC-IPTS), Seville (Spain) e-mail: fabiana.scapolo@jrc.es
2 Prof. Ian Miles, University of Manchester PREST, Institute of Innovation Research IoIR, Manchester (UK) e-mail: ian.miles@man.ac.uk
though we sought to construct equally good practice implementations, to avoid weighting the dice in favour of one or other approach. The paper will describe these encounters with methodology and evaluation, present some of the results of the forecasting study and draw lessons as to methodological good practice in general (e.g. the need for systematic documentation, the scope for debate about established practices) as well as good practice in use of these specific methods.

I. INTRODUCTION

Expert opinion is taken into consideration in policy making, on the grounds that such views should give policy makers added information and knowledge in the fields where they lack sufficient knowledge. Policy makers will often have insufficient knowledge to comprehend the complex domains they are dealing with – especially (but by no means exclusively) areas of science and technology. In these cases policy makers are highly dependent on the quality and reliability of information they have at disposal, and they depend on, the advice and guidance of experts (often relying upon particular experts to tell them about what expert communities think).

Technological forecasting and Foresight can be seen as aiming in part to produce information on the supply of and demand for new technologies, in the context of evolving capabilities and contexts in the social and physical environment.

Expert opinion can be solicited by including the experts in meetings or by interviewing them; in-depth interviews are particularly valuable. But such techniques are restricted to the inputs of relatively few people, and there are opportunities for the loudest or most prestigious voices (which are not necessarily the most informed ones) to dominate affairs – though new Information technology assisted methods of group interaction allow for a great many more views to be generated and reviewed than traditional methods permit (There do tend to be problems of information overload when it comes to the use of large volumes of qualitative information captured in these ways, however.) Techniques such as the Delphi method were developed to circumvent such problems.

In the past decades, the literature on such methodologies has elaborated how they should be implemented (Glenn and Gordon 2003), but provides less evidence on how to make an informed choice on which technique is more appropriate to use when undertaking a technological forecasting. The present study was undertaken with the aim of elucidating how the most appropriate methodology to apply when undertaking a technological forecasting study could be selected. The starting point was the idea of a comparative evaluation of two methods applied to the same topic, to determine how the processes of gathering information differ, and to compare the types of knowledge developed from the application of the technique. However, a simple comparative evaluation proved very difficult to undertake.

The two techniques applied here were the Delphi method (using the standard anonymous questionnaire survey, with two rounds of iteration), and the SMIC (Cross Impact Systems and Matrices) method of cross-impact analysis. Applying the techniques to the field of future transport technologies raised questions about how we can elicit what sorts of knowledge from expert how the results of this elicitation process can be applied so as to (hopefully) improve the quality of information inputs into for the decision making process. The study thus casts light on the requirements for the implementation of these techniques, both in terms of practical steps and content necessary for their applications.
II IMPLEMENTATION OF THE TWO FORECASTING TECHNIQUES

The objectives of this study were to establish how far are there, what are there, and how important are there:

a) Generic issues to do with elicitation of expert opinion (e.g. selection of experts, boundaries of tasks, etc.);

b) Specific issues posed by this class of methods – based on questionnaire, and with qualitative and quantitative inputs and outputs – (e.g. selection of questions, coding of questions and responses, responses rates, etc.)

c) Specific issues in the specific techniques if it is possible to overcome some shortcomings of the two techniques, and

d) Specific issues related to implementation (e.g. possible items of good practice of implementation) and whether if the outcomes of this assessment produce a level of knowledge suitable to be used in the process of decision making.

The two forecasting techniques selected, the Delphi method and the SMIC method, are both based on eliciting expert knowledge. Experts’ contribution is seen as a help in areas of research where an explicit conceptual framework may not exist (i.e. where formal methodologies which make the use of any existing theory and data are not available, or underdeveloped, or not widely accepted). For this reason experts on a particular problem area may possess unstated ‘mental models’ and knowledge of the causal structure of a particular system. This approach attempts to externalise, materialise and manipulate such informalised ‘expert’ opinion as a basis of forecast.

The first serious attempt to systematically implement this approach - beyond traditional means of group discussion - was the Delphi technique. In the course of the years, the Delphi method has seen a tremendous number of applications in almost every field, this due to an unquestionable characteristic of this method which is its high level of flexibility. A conventional Delphi application asks when, if ever, a series of events will occur and perhaps require estimates on the desirability of events being forecast. One particular variation of this implementation is where the experts are asked to give probability estimates of an event’s occurrence by a certain date in the future, rather than just guessing when an event might occur. This particular variation of the Delphi technique may be most suitable when it is necessary to estimate how much likely is the occurrence of an event, when mutually exclusive events may occur. Moreover, it gives a picture of a particular year, rather than a future history.

Nevertheless, one of the basic observations on the Delphi methodology has been, that with this technique it is possible to obtain forecasts for many events, without having information as to the respondents’ assumptions whether questioned events are interrelated. Thus, it is possible that the outcomes of a Delphi implementation could produce a forecast of events mutually reinforcing or exclusive of each other, and that an artificial ‘consensus’ may be reached.

The initial works of Helmer and Gordon on cross-impact analysis aimed to determine interdependence between events. The cross-impact technique attempts to provide the probabilities of occurrence of an item, adjusted in view of the occurrence of related items with potential interactions on its occurrence. It thus tries to overcome a disadvantage of the Delphi method.

Even if the developments and analysis of cross-impact methods have concentrated their attention on the manipulation and refinement of probability estimates, it is doubtless
that these techniques can have a clear role in understanding and clarifying the complex causality of dynamic socio-economic systems. Cross-impact methods, applied together with other available techniques, like the Delphi method, produce as added value a better comprehension and evaluation of hypotheses about the way a system works. If applied in this context, cross-impact methods can be seen not as a technique able of providing exact and neat solutions to the future, but as an aid to learning, a heuristic device which might be able to improve the understanding of complex problem areas, which are already tackled by forecasting. In this study we tried to establish whether cross-impact analysis employed in such a manner, would be able to underline the differences between the ‘mental models’ used by experts.

In this way, the implementation of cross-impact models can be relevant to the process of decision making. Cross-impact analysis has a great potential because it provides its users with a method for synthesising a wide range of beliefs and opinions concerning future developments. In fact, the final outcome of the cross-impact method used for this specific implementation, the SMIC method, is a cardinal sequence of possible scenarios, which provide an additional dimension for problem evaluation to the decision maker.

In spite of that, it is not possible to say that scenarios which stem from cross-impact analysis, no matter how accurate probabilities might be, are going to become the reality of the future. They are only possible images of how the future could look like, if the considered events do, or do not occur. However, if all the possible combinations are generated, then the right one should be present, even if it is possible that it might be ruled out through low probability.

1. Sequence of procedural steps for the implementation of the Delphi method and the SMIC method

In order to be able to answer the questions set at the beginning and to meet the objectives of this study, a whole series of decisions had to be made in order to implement the two forecasting techniques in such a way to allow any type of comparison. Figure1, describes the main steps undertaken in the course of this study. Once the objectives of the study were defined, we had to take a number of decisions related to the technologies to forecast, but also a whole series of decisions related to the implementation of the two techniques including elements to allow a comparison between them.
Figure 1: Sequential procedural of the main steps for the Delphi and SMIC implementation
• **Choice of the subject**

One of the first decisions had to do with the technological area whose future was to be examined. The thematic area of this forecasting study was related to the future evolution of the urban transport sector in Europe. The focus was on the role and development of Transport Telematics (ATT) technologies, that are being developed to help to improve the current situation of traffic in European cities.

Rather than poll a set of experts as to important emerging technologies in this area, we were able to identify specific technologies to forecast from the European Commission programme on ATT (Advanced road Transport telematics), undertaken in the course of the 3rd Framework Programme of the European Union. This programme piloted projects addressing seven different areas of major operational interest. (The areas were addressing both the optimisation, integration and adoption of telematics technologies for private and public transport, and the collection, processing and distribution of travel and traffic information including services such as traveller information, trip planning and route guidance. In addition technologies to assist the driver and communicate information between the vehicles were included in the programme.) The pilot projects of the Programme were validated in several European cities and/or arteries - small parts of the transport network, indeed, and the challenge for the future lies in their application to much wider portions of the transport network.

The study was constituted so as to examine views of the level of development of these applications, the forces that facilitate or impede this development, and the consequences for the normative goals that they are aimed at. Thus we determined to address:

a) **the likely level of use of transport telematics technologies and their impact on the transport system in the future in Europe; and**

b) **the consequential direct impact of these technologies on the level of congestion and environmental issues.**

c) **How far factors of economic, social, legislative and political nature may impede the development of ATT technologies.**

Specification of this general content of the forecasting exercise is only the first of an entire series of decisions that proved necessary for the implementation of the two forecasting tools. In order to enable comparison between the two techniques, it would be most appropriate useful to maintain the same structure for the implementations of the Delphi and the SMIC method. But the precise characteristics of the two methods imposes limits on these common elements.

First, one major constraint is the number of questions (or events) that can readily be included in the two implementations. The Delphi method is fairly flexible in this respect. It could in principle contain as many questions as the moderator would like to ask, though a lengthy questionnaire will reduce participation levels. Thus one authority (Parenté and Anderson Parenté 1987)suggests that twenty-five topics (questions) should determine the upper limit to make sure that the exercise can be manageable, both for the moderator and for the experts. (This is very much a matter of precise organization and use of topics, we would suggest.) The SMIC method has a limitation in the number of events handled, generally a maximum of six, so as to constrain the maximum number of

---

questions that can be reasonably put to the experts, that grow exponentially as more topics are added (Godet 1993).

This constraint on the number of questions that can be used in the implementation of the two techniques immediately leads to problems when we try to apply them to an area such as the future of transport telematics technologies. The areas of the ATT Programme are very broad, including not only various telematics technologies, but also policy measures helping to reduce traffic, and organisational and social considerations related to the provision of information and the response of users. It is hard to capture these in 25 topic items, let alone 6!

How could we manage to institute a SMIC study given this diversity of applications? In order to have some ground to compare results, and to allow SMIC to be applied to the areas of most significance, the decision was made to select, in effect, the events to be included in the SMIC method on the basis of the results of the first round of the Delphi exercise. The SMIC was used to consider the most prominent transport telematics technologies, along with two elements of the “impact” of those technologies (on congestion and environmental issues).

Thus at the very outset we confronted decisions about selection of topics that were highly constrained by the specificities of the techniques. We were able to work in much more common ways in respect of certain other facets of the study.

Where it came to the time horizon to be studied, we oriented both methods around developments by a particular point in time. Given that ATTs are at quite different stages of development, it was reasonable to enquire as to the degree of penetration in the market of different systems at a particular point in time (about twenty years ahead), a how far question rather than one asking when a certain technology will be available on the market. A second rationale for doing this was to have a comparable structure of the Delphi method toward the SMIC method, which asks experts to give probability estimates of occurrence of events by a certain date in the future. We could thus directly compare the replies given by experts in the Delphi and SMIC exercises - assuming that formulated questions are similar and comparable.

The selection of experts for the implementation of the two techniques was made on similar grounds, drawing on

- participants to latest international conferences on Advanced Transport Telematics Technologies;
- experts co-nominated by others;
- experts selected from national research centres and academia on transport, (also with the support of Internet);
- contacts made in the course of the transport research.

At the end of this procedure, we were able to identify 300 experts to involve in the exercise. This population experts was divided into three sub-groups: the aim was to send one group of experts the Delphi method only, another one the SMIC method only, and to the last group both the Delphi and the SMIC methods. This would enable us to examine how far results are coherent; whether expert responses and participation levels are affected by receiving only one or both questionnaires, and how far they perceive the issue in a different manner, and apply different degrees of effort to it, in the different contexts. They might also assess the methods and their outputs differentially according to whether they had participated in them.
III. Delphi operationalisation

The background information of the exercise

The Delphi survey was conducted in two rounds. The structure of the questionnaire was divided in three stages corresponding to those indicated above by points (a), (b), and (c), and also included a general background discussion which embodied a scenario and a definition of European medium size city. This general background was the same for the Delphi and for the SMIC implementation.

This background information included a baseline scenario to help experts while replying to place themselves in a common context. The selected scenario can neither be considered a ‘business as usual scenario’ nor a scenario of ‘wider innovations’, but something in between (i.e. it could be defined as a scenario of ‘moderated change’). The starting point has been the past economic, social and political trends of the last 20 years, which have favoured the use of personal transport and encouraged the use of cars, to the detriment of use and development of Public Transport.

We attempted to build a ‘plausible’ scenario, in order not to impose too much on the experts working with it. Following the same line of thinking, experts were provided with a definition of medium-size city.

The selection of transport telematics technologies from the DRIVE/ATT Programme

The selection of the transport telematics technologies to include in the questionnaire, as well as the formulation of the questions to the experts, was one of the most crucial points of the design of the questionnaire.

It is possible to define this implementation of the Delphi method as a ‘technology-oriented’ questionnaire, where the forecasting of the development of different ATT systems has been the primary goal.

The examination of the Transport telematics Programme considered relevant for the purpose of this study has allowed us to list in the first place 42 ATT systems, which can be applied in the context of the urban transport. The list of 42 ATT systems was reduced to 24 systems for us to include in the Delphi exercise. This reduction of the list of technologies was achieved by eliminating technologies that were implemented in different pilot projects for distinct applications and with different purposes.

In a final step, the questionnaire should ideally be piloted before sending it out to the experts. In this study, this process permitted as described later, the achievement of a further reduction of the technologies to be forecasted from 24 to 19 ATT systems.

Before reaching the final version of the Delphi exercise it was necessary to take many other small decisions, in order to have the clearest possible questionnaire, and especially in view of the desired comparison with the SMIC method. These decisions referred mainly to the topic of the forecast in a way to ensure that the information generated was actually going to provide new ‘knowledge’ and information on the development of the technologies. The challenge of ATT systems lies in their widespread use on a large part of the existing transport infrastructure. At the time of the designing of the inquiry, such use was foreseen, but it was still very hard to predict when such systems might be implemented on the widest scale, and not only on a limited area in the framework of a pilot project. This is due to different reasons and constraints, especially because ATT
systems are passing now from the research and development step to that of experimentation and in some cases commercialisation. These constraints are not only of a technological nature. The complexity of inserting ATT technologies into the existing institutional and legal frameworks, and the considerable diversity that is present at national level in the European countries, where the roles of the authorities vary enormously in term of competencies, can delay the homogeneous adoption of these technologies.

For these reasons, experts were asked, in the first step of the Delphi questionnaire, what will be in their opinion the likely level of use of the selected ATT systems applied to urban transport, in European medium size cities in the year 2015. In order to facilitate this task, experts had to indicate, for each technology, what proportion of European medium size cities will adopt the technology by the time horizon considered (i.e. they had to indicate the level of diffusion of the forecasted technologies).

The second and third steps of the Delphi inquiry aim to achieve merely qualitative answers from experts. The aim here was to gain some knowledge on the direct effects and implications of the implementation and development of this range of technologies on the transport system.

In the third step of the Delphi exercise, experts were asked to qualitatively assess factors which might restrict the occurrence of the widespread use of ATT technologies.

Table 1 gives an overview of some decisions taken at the level of Delphi implementation for this study, and distinguishes from typical implementation of the method.

**Table 1: Summary of some decisions for the Delphi implementation**

<table>
<thead>
<tr>
<th>ISSUE IN STRUCTURE OF DELPHI</th>
<th>‘TYPICAL’ IMPLEMENTATION OR CHOICES AVAILABLE</th>
<th>IMPLEMENTATION IN THIS STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall framework</td>
<td>Often diffuse</td>
<td>Background scenario + medium-sized city defined</td>
</tr>
<tr>
<td>Nature of experts</td>
<td>Typically national</td>
<td>Cross-national</td>
</tr>
<tr>
<td>Topic forecast</td>
<td>Few words (recommended = most 25)</td>
<td>Relatively many words (average = from 7 to 29 words)</td>
</tr>
<tr>
<td>Nature of the forecast</td>
<td>When ‘X’ will happen? To a given extent</td>
<td>How far ‘X’ will happen in a given time?</td>
</tr>
<tr>
<td>Time horizon</td>
<td>Different time horizons</td>
<td>Fixed time horizon</td>
</tr>
</tbody>
</table>

1. **Main criticisms of Delphi methodology in the literature**

In this section, we would like to discuss elements of the implementation decisions by recalling the main criticisms of the Delphi method encountered in the existing literature. We examine the most relevant issues already analysed by the literature, taking into consideration the problems the operator has to face in the implementation phase of the methodology, and the problems related to the evaluation phase of the outcomes achieved.
In the course of the years, much research has been devoted to the Delphi method, especially if compared to the amount of literature related to any other forecasting methodology.

- **The issue of the participation of experts in a Delphi exercise**
  A first point refers to the lack of precise indications on the number of experts to involve in a Delphi inquiry. (Marbach, Mazziotta et al. 1991) There are indications, in the literature, that the minimal size of a panel of experts to involve in a Delphi exercise, should be no less than eight to ten members. (Mitchell 1991) However, usually the number of experts to involve in the exercise is a choice left completely to the operator. This procedure includes as an inherent risk, that not all the competencies required to investigate the inquired subject may be represented in an adequate manner. For our specific implementation we have decided upon having a substantial number of transport experts, rather than a small sample. In doing so, we assume that we increase the possibility of representing in a balanced way, the wide range of competencies on the future of ATT technologies in European cities by the year 2015. However, we can confirm that the lack of precise criteria on the number of experts to include in the Delphi survey - and the lack of criteria on how to determine the number of experts in a specific case - was perceived as a drawback of the methodology.

Second point related to the use of experts in the Delphi method, refers to the validity of expert judgement compared to non-expert judgement. In the literature it is possible to encounter controversy about this. Sackman (Sackman 1975) argues that, a panel of experts composed by people with similar backgrounds and interests, may tend to comprise an elite with a vested interest in promoting the area under Delphi investigation. Then, he believes that the participation of prestigious individuals does not provide evidence of improved accuracy. On the contrary, it can be counterproductive because they usually do not provide carefully considered responses, if compared to those stemming from younger panellists. In relation to this point, the Delphi literature provides examples of research where experts and non-experts were used, without any substantial underlying differences in the results. (Bedford 1972) Certainly, it is not possible to say that experts always provide more accurate predictions than non-experts do. Our implementation of the Delphi method on the future of ATT systems in European medium size cities in the year 2015 resulted in a largely technological oriented inquiry. The uncertainties about the future and rapid development of transport telematics technologies – technologies that are often not known to non experts – led us to select a panel of experts with at least some knowledge on ATT system. It was not a requirement to be a specialist in this field, but it was a requirement to work at least indirectly in the transport field. For this specific implementation of the Delphi method, with Sackman’s theory that ‘the concept of expert is only marginal and virtually worthless from a practical point of view’ (Sackman 1975 p. 43) did not apply.

- **The construction of Delphi event statements**
  The problem of construction of Delphi event statements was raised by Salancik et al., (Salancik, Wenger et al. 1971) and further examined by Martino. (Martino 1993) Salancik started from the tasks of the operator of a Delphi exercise. Among these tasks there is the description of the events to be forecast. The role of the operator must be to
facilitate agreement. His/her effort is to describe the potential event so that all respondents interpret it exactly in the same way.

The description of the event statement is crucial, since the expert has first to decide what interpretation to give to the event, and then try to translate this interpretation into a prediction. To the extent that there is more or less conflict in the information describing the event, the operator could expect that the expert would have to add or subtract information to derive his interpretation. This process varies, and is very subjective.

Our implementation of the Delphi method, paid a lot of care in the formulation of the questionnaire, and of the event statements in particular. As well as being concerned with validity, we were aware that a complicated questionnaire could only increase the dropout rate. Given that, the event statements in this implementation of the Delphi method were technologically oriented (i.e., basically a description of various ATT technologies and their possible application), clarity was required in the description of the events. We have decided that it was better to have shorter rather than long descriptions, and for this reason we have constructed the event statements excluding as much as possible any superfluous word. Out of the nineteen technologies considered in our inquiry, only one event statement was composed of more than 30 words (i.e. event 15). The formulation of all other events counted from 7 to 29 words. Nevertheless, the concise description of the ATT systems was criticised by some experts because gave the questionnaire a too technical orientation.

It was more difficult to avoid similar events. This was mainly due to the fact that some ATT systems, based on different technology, serve similar applications. Since it is not clear yet which system is going to lead the market in the future, it was of interest to assess the opinion of experts on the likely level of use of those systems.

- The problem of dropout rate
  A problem with the Delphi method discussed in the literature, is the tendency for high panel attrition. (Hill and Fowles 1975) The influence of and reasons for dropouts of the Delphi method have not been widely investigated in the literature. Sackman (Sackman 1975)suggests that, there are factors that operate to determine a hard-core group that sticks with a Delphi study through all iterations. The positive reasons can be identified as strong motivation and interest in the target area, the negative ones can be identified in a strong disagreement with the design and content of the questionnaire. Another motivation for dropout can be due to a critical attitude towards the utility and methodology applied of the study, and to ‘design content’ of the technique applied.

In our implementation of the Delphi method, we had only one expert who sent back the first round of the Delphi method uncompleted, explaining that he does not believe in results of forecasting exercise. He refused to participate in the study arguing that, from previous experiences he learned that these exercises are a waste of time.

The first round of the Delphi exercise was sent to 224 experts. Out of the 224 inquiries, 87 experts (more than one third) returned their questionnaires completed. We do not have any information on the 110 experts contacted who did not answer to the first round of the Delphi exercise. We assume that, a proportion of these experts did not reply because they were in disagreement with the content and methodology applied for this study. However, we believe that there are other factors that have contributed, to this amount of non-response questionnaires this study, such as lack of time. The second round was sent to 87 experts, and obtained 58 responses. From the first to the second
round the dropout rate was 33%. We again do not have information on the reasons for dropouts, which might correspond to the factors described above in this section. We would like to add that the repetitiveness of the Delphi procedure could generate the weariness of respondents, who might think that the procedure of the Delphi method is too demanding.

- **The achievement and stability of consensus**

The Delphi method was designed as a tool to overcome the biasing effects of face-to-face discussion in a group. Its goal is to reach, without such psychological interference, a summary, which expresses common opinions of the group. This summary should be achieved as a result of convergence of opinions (when achieved), which are the expression of consensus of the interrogated group of experts.

The iterative process with controlled feedback, should assures the experts a way of autonomous judgement on their opinions over successive rounds. However, the pressure to reconsider their opinion over successive rounds, can constitute a pressure for the convergence of opinion. In the literature, there are controversy assumptions to this regard. The most critical reviewers of the Delphi technique, believe that the method manipulates responses towards the minimum dispersion of opinions so as to achieve consensus. The supporters of the technique believe that, the method does not force the achievement of consensus, but seeks to facilitate consensus. (Marbach, Mazziotta et al. 1991) Nevertheless, Martino (Martino 1993) who is favourable to the Delphi method, has pointed out that there is ample evidence, from a number of experiments, that if the panellists feel that the questionnaires are an imposition to them, or if they feel they are rushed and do not have time to give adequate thought to the questions, they will agree with the majority simply to avoid having to explain differences. Therefore, in this respect the Delphi procedure is not an absolute guarantee against the influences of the ‘bandwagon effect’ and fatigue.

Actually, the information at each round of the median value, or of any other measure of central tendency, might be interpreted from the participants as the ‘correct’ answer. The experts may consider, when they examine their previous estimates, that it would be appropriate to get closer to the median value.

In our implementation, we restricted the number of rounds to two, in order to limit as much as possible the ‘bandwagon effect’ and the weariness of respondents.

- **The accuracy of Delphi results**

A critical issue for the Delphi method concerns with the quality of agreement that takes place in the successive rounds. Some of the literature is critical of the idea that convergence of opinions improves the accuracy of the forecasts. Others argue that the exchange of information, arguments and hypotheses can help to improve accuracy.

It is impossible to assess whether the outcomes of our implementation of the Delphi method on the future of ATT technologies in European cities in the year 2015 are accurate, since we cannot check the future. As a consequence of this, we cannot determine whether the results achieved have produced a correct prediction. Nevertheless, the point of the method is to determine the collective knowledge of expert views on the issues. This should in any case have some utility, regardless of accuracy.
Additionally, we were able to collate some information and indications, by mean of the evaluation questionnaire on the results of the Delphi exercise, which can be considered as a worthwhile piece of research, at least to understand how the experts involved in the study have perceived the forecast obtained.

Once the content of the questionnaire has been established, other features can be designed - and other decisions can be made - which can facilitate the experts in their task of replying to the Delphi exercise. In fact, it is not enough to define the content of the questionnaire. Other elements such as guidance on how to answer the questionnaire, are also extremely important. All the material that accompanies the questionnaire can influence the experts. A very clear explanation on how to complete the questionnaire, can help respondents to better understand the aim of the study. If the questionnaire is presented in a clear way, the dropout rate can be reduced - the respondent is liable to feel that the designer has taken care, so care is required in return. In general, if a questionnaire is complicated, experts see it as too time consuming, and this may evoke the decision not to participate in the exercise.

IV. SMIC OPERATIONALISATION

The implementation of the SMIC method followed the computation of the results of the first round of the Delphi questionnaire. One goal of the SMIC assessment was to validate the information, stemming from the results of the Delphi study, on the future evolution of ATT technologies. One approach followed was to send the SMIC questionnaire to a group of experts that were also involved in the Delphi inquiry. This was done to check the reliability and consistency of the outcomes of the forecast on the future of ATT in European cities in the year 2015.

The SMIC method usually takes the form of a postal inquiry. The SMIC method, and its dedicated software, provides an established way for structuring the questionnaire. This implies that, the person applying the methodology has to design the sequence of the decisions involved for the implementation in a less flexible way than in the Delphi method.

As for the Delphi method, this specific implementation of the SMIC method was applied to the future of ATT technologies in European cities in the year 2015. The SMIC method has a limitation in the handling of events, which in general are reduced to six. In order to have some common points for the comparison of the results achieved by the two methodologies, it was decided to design the questionnaire including four events concerning the adoption of ATT technologies by European medium size cities in the year 2015. The remaining two events addressed the effect of these technologies on two transport problems, which were again the same as those included in the Delphi questionnaire – that is level of congestion and level of traffic volume of passenger vehicle in urban areas. It will be noted that there were no events involving factors which could influence the widespread use of ATT technologies in European cities. This reflected space limitations.

The background information provided with the SMIC questionnaire were the ‘background’ scenario and the definition of medium size cities that were identical to those used for the implementation of the Delphi questionnaire.5

1. Structure of the SMIC questionnaire

5 We will not here repeat the description and justification for the adoption of the baseline scenario and definition of medium sized city. The descriptions given for the implementation of the Delphi method apply also for the SMIC method.
The questionnaire of the SMIC method involves two steps. In the first stage, experts are asked to estimate the simple probability of occurrence of the selected events, through means of a probability scale from 1 (i.e. very low probability) to 5 (i.e. very high probability). In the second stage, experts are asked to estimate, in the form of conditional probability, the likelihood of an event coming true as a function of another event. The events are considered in pairs. The experts express their opinion through means of the same probability scale as used in the first stage, but an extra option is available that allows them to indicate the independence of two events. Basically, in the second stage experts have to fill in two matrices for each event. The first matrix was referring to the probability that each of the other five events listed would occur given that event (En) turn out to be true. In the second matrix they had to indicate the probability that each of the other five events listed will occur if event (En) does not turn out to be true.

Table 2, summarises some of the characteristics of the design of the Delphi and the SMIC inquiries, underlines the common items used for the implementation of the two methodologies.

Table 2: Characteristics of the Delphi and SMIC inquiries

<table>
<thead>
<tr>
<th></th>
<th>DELPHI METHOD</th>
<th>SMIC METHOD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Background scenario</td>
<td>same</td>
<td>same</td>
</tr>
<tr>
<td>Description of the city</td>
<td>same</td>
<td>same</td>
</tr>
<tr>
<td>Number of topics</td>
<td>19</td>
<td>6 (4 taken from 1st round Delphi)</td>
</tr>
<tr>
<td>Inputs to questionnaire</td>
<td>same</td>
<td>same</td>
</tr>
<tr>
<td>Number of rounds</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Probability</td>
<td>occurrence by date</td>
<td>occurrence by date conditional probability</td>
</tr>
<tr>
<td>Number of experts</td>
<td>224 (first round) 87 (second round)</td>
<td>126 77 (only SMIC) 49 (SMIC and Delphi)</td>
</tr>
</tbody>
</table>

Criteria for the selection of the technologies

The selection of the four technologies to include in the SMIC method was accomplished on the basis of the results achieved on the first round of the Delphi questionnaire. This was done especially to allow some comparison of results the two techniques were able to generate on the same subject. One criterion used was to select four technologies amongst the most representative, where the level of convergence of opinions achieved in the course of the first round of the Delphi was relatively high on all the questions asked. This was done also to see if the results achieved in the context of the Delphi method could be confirmed by the SMIC implementation.

A second criterion was to include four ATT technologies belonging to the three different main groups considered for the implementation of the Delphi method, which were:
• technologies able to measure the traffic level on the network;
• technologies that, on the basis of the existing traffic level, are able to adjust traffic signals to give priority to Public Transport vehicles;
• those technologies that can assess the level of congestion in real-time, and can communicate with the vehicle (i.e. re-routing of vehicles).

Formulation of the event statements
Among others, the main goals of the implementation of the SMIC method were to:
• assess the validity of the results achieved from the Delphi exercise;
• assess if and how far the application of two different forecasting methodologies on the same subject produces similar or different results;
• examine the interdependency among questions.
Experts had to assess, the simple probability of the events in the first stage of the SMIC inquiry, whereas in the second stage they estimated the probability of pairs of events.

In designing these questions, the researcher was aware that, the selected formulation of the event statements would have led to some loss of information. In fact, the decision to include a percentage range of European medium size cities in the formulation of the event statement involved posing a clear threshold. The reaction of experts to this formulation of event statements may vary from one expert to another. If one expert disagreed strongly with the threshold posed, his response might be to decide not to answer the questionnaire. We suspect that this option is most likely, if the expert thinks that the indicated percentage range is too optimistic.

There are always some drawbacks to consider while designing a questionnaire. Any formulation of questions or events can be criticised on the ground that they restrict the information that can be elicited. Unfortunately, there is no clear indication on what is the ‘best practice’ on how to implement this type of forecasting tool, based on eliciting expert knowledge. Moreover, it is sheer utopianism to determine a ‘general best practice’ on how to implement forecasting tools, since they are applied to complete different topics.

The decisions taken by the person applying a forecasting and or prospective methodology will have some drawbacks, or pitfalls. Nevertheless, there is little existing advice in the literature, or indications of how overcome this problem. The arbitrary components of any implementation of this range of methodologies, and related consequences, should be taken into consideration, both from the person conducting the forecast, and from the person participating in such exercise.

2. Presentation of the SMIC inquiry to the experts and main criticisms of SMIC in the literature
In order to complete the implementation of the SMIC method, material such as an explanation to the experts on how to complete the questionnaire, and an accompanying letter with a description of the aim of the study, was necessary. The presentation of the SMIC questionnaire to the experts, once the content has been established, was articulated in the same manner as for the Delphi implementation.

The explanation of how to use the SMIC method was important, since this methodology has been less widely applied than the Delphi method. The guide to filling in the questionnaire, indicated that it was divided in two steps. Examples of answers were given, with clarification of the meaning of the probability scale that experts had to use to complete the exercise.
The SMIC inquiry was sent to two different expert groups. The first group of experts was not contacted beforehand, and received the SMIC inquiry only. The second group of experts had already been contacted in the framework of the Delphi inquiry. This group of experts received both the Delphi and the SMIC inquiries. The main reasons for sending two different inquiries to experts was to assess:

- if results are consistent on an individual basis;
- if the attitude of the experts is different when they receive only one as opposed to both questionnaires (e.g. whether there is a higher level of drop out from the exercise);
- whether using both methods leads to a different response to one or other of them (i.e. for example, thinking about interdependency in SMIC might change judgement on one single issue in Delphi).

Compared with the literature on the Delphi method, much less has been written about the way these techniques should be implemented. However, especially in relation to the SMIC method, it is possible to identify some common points, which deal with implementation that has been considered in the framework of the Delphi method. Since, the two methodologies are both based on eliciting experts’ knowledge through mean of inquiry systems, there are some common procedures, such as the selection of the experts, which can also influence the outcomes of this technique.

The critical literature on the SMIC method was most active from the 1974, when the first application of the SMIC method was presented, to the year 1976. The critical literature on the SMIC method was principally concerned with two major issues:

- the problem of inconsistency in cross-impact probabilities; and
- the use of cross-impact analysis, the SMIC method in particular, to forecast the future of social systems.

Furthermore, is not the aim of the application of cross-impact analysis to provide exact forecasts of how the future could be. Rather these techniques should be seen as tools to improve the understanding of complex problem areas. Cross-impact analysis can be used, together with other forecasting techniques, as a heuristic device to clarify and increase our knowledge for the selection and evaluation of hypotheses about how events may develop in the future, and thus helping the decision making process.

Our application of the SMIC method aimed to deepen and extend the knowledge achieved within the Delphi exercise. We were especially keen to check the interdependencies among events and to examine scenarios, which were not considered in the framework of the Delphi method. The literature reviewed gives little insight into the ‘best practice’ for implementing cross-impact analysis. The emphasis has often been on how the various techniques refine and manipulate probability estimates, underlying the computation aspect rather than the conceptualisation one.

There is little indication in the literature on a number of issues discussed in the context of the Delphi study, which can be applied also to the cross-impact analysis. There are some issues, discussed in the context of Delphi, which are applicable also to the SMIC method. There has been little deliberation on the SMIC method on issues related to the

---

role of experts, the influence of the size of the panel on the results, and the issue of the validity of experts’ judgement.

In relation to the number of experts to whom the inquiry should be sent, Godet (Godet 1993) indicates that the group should be composed by 40 to 60 experts, it is not clear if there is a rationale behind this indication. However, little indication is given on how to adequately represent all the competencies of the inquired subject. As with the Delphi method, there is a lack of criteria on how to compose the group of experts as a drawback of the methodology.

Concerning the involvement of experts in Delphi exercises, the literature has been critical of the use of experts’ judgements, since these can be biased. Similar discussion of the possibility of obtaining biased forecasts, when a cross-impact method is applied, is absent. But, if this has been considered in the framework of the Delphi method, it should be considered also for cross-impact analysis, since the support of mathematical-statistical computation cannot certainly eliminate biased judgements. We believe that, if the arguments noticed for the Delphi method – already described above in this text – have to be considered valid, the same should be applied for the SMIC method, and cross-impact analysis more in general.

In addition, it was not possible to locate any reference discussing the issue of how to construct SMIC event statements, as exists for the Delphi method. Yet, it is evident that for the SMIC method, which can be composed of a maximum of six events, the construction of the event statements plays an essential role. For example, in the Delphi method the expert without implying a serious distortion for the final outcomes can skip an ambiguous event. But for the SMIC method the implications will be completely different. In fact, an ambiguous description of an event statement does not allow the expert to reply to the cross-impact questions. It will thus invalidate the entire questionnaire.

For our implementation of the SMIC method we have followed the criteria suggested in the framework of the Delphi method, derived from the literature review and discussed above in the text. Thus, we have tried to construct a ‘best practice’ approach. However, the lack of specific criteria for the SMIC method has been considered as a drawback, in particularly the lack of indication addressing which type of event statements it is preferable not to have combined in the same inquiry. In fact, it is not specified whether it is better to avoid mixing among the six events, some that are technological oriented with some that are more economical or social oriented.

Finally, the fact that the SMIC method takes form of a single questionnaire to the experts, has exonerated it from the debate which involves the Delphi method in relation to the improved accuracy of forecasts over successive rounds. Of course, this would not necessarily imply that the forecasts, which stem from cross-impact analysis – including the sometimes ‘mysterious’ mathematical and statistical computation and manipulation of data – could be considered more accurate than those ones which stems from a Delphi exercise.
V. CONCLUSIONS

The work described in this study was initiated by a set of questions on the role of forecasting in the policy decision making process, more precisely on the role of methods based on soliciting experts opinions in the policy decision process. The basic questions concerned:

1. ‘how’ to select the best methodology among the existing ones;
2. ‘how’ to reach and define ‘best practice’ implementation of these techniques;
3. ‘whether’ and in ‘how far’ it is possible to achieve a different level of knowledge and information, when two different techniques – based on soliciting experts opinions – are applied to the same topic, but having a different process of gathering and analysing information. Finally,
4. ‘whether’ and in ‘how far’ the process and decision sequence required for the implementation of these techniques have an influence on the usefulness of the generated forecast.

In order to answer to these questions, and in absence of a precise technical template that can help the implementation of methods based on soliciting experts’ knowledge, the study was organised in a way that provides explicitly:

- the decision sequence for the implementation of the two selected techniques;
- the actual operationalisation of the methodologies;
- the analysis, reporting and comparison of the results of the implementation of the two techniques;
- an evaluation of the final results achieved by the implementation of the two methods, provided directly from the experts participating in the various phases of the study.

In this final section we will try to synthesise what it can be generalised from the experience of the specific implementation of the two techniques, especially by focusing on:

- what the evaluation shows about how useful each method was, about the relative merits of each approach;
- what the practical experience of the operator was of using the two methods;

In addition some discussions and general conclusions will be provided on:

- what further research is needed on the methods, after using and evaluating them;
- practical implications for would-be users, how to implement the methodologies, concerning how to use and disseminate results.

Evaluation on the implementation of the two forecasting methods

The implementation of the two techniques was based on using a particular form of the Delphi method, and followed the strict ground-rules associated with a pre-given cross-impact technique, the SMIC method (Cross Impact Systems and Matrices). Inevitably, this study is evaluating and comparing the implementations and results of particular versions of Delphi and cross-impact analysis, but hopefully this evaluation is on ‘best practice’ within these versions. For this reason, the discussions and conclusions of this study have greatest bearing on the particular versions applied. It is not possible to evaluate and compare all versions of the Delphi method and of cross-impact methods in such detail, of course. But we can consider how far the conclusions reached here are liable to apply to other versions and implementations of these methods, and to methods based on eliciting experts’ knowledge more generally.
In the study, there were different levels of decisions involved with the implementation of the two techniques, which can be summarised as follows:

a) decisions about the operationalisation of the inquiry;

b) decisions about the analysis of the collected data;

c) decisions about the presentation and evaluation of results;

d) decisions for the dissemination of the final results.

The four levels of decisions are related to each other, with decisions taken in the first level have implications on the other levels.

**Evaluation of the usefulness of the methods**

In general terms, the experts involved in the study have judged the study in a positive manner. In terms of active participation from experts, the results achieved exceeded the initial expectations.

In respect to the Delphi method, none of the experts reported many gaps in this implementation of the Delphi method. Thus, the experts assessed that the operationalisation of the Delphi inquiry, and how the inquiry was elaborated and managed did not have many gaps. This positive assessment indicates that this implementation and the results achieved were rather comprehensive. Furthermore, the experts assessed the inquiry as a valuable mean of communication for exchanging opinions on the topic. This was especially valid for the Delphi method that provides results through statistics of group responses, where the opinions of the entire sample of experts are represented. The experts found that the Delphi method gives the possibility to assess the existence of broad views on the future of a topic, increases credibility and acts as a tool that help to reinforce personal opinions. Finally, such results can encourage action and investments in a field (i.e. in our case on ATT technologies).

On the other hand, the experts felt constrained – in this particular implementation – since there was no space to forecast other topics than the ones selected by the designer of the method. (This would not be the case in a more traditional form of Delphi, where topics for study were initially predefined by the pool of experts). Then, not all participants appreciated one of the characteristics of the Delphi, which is that the inquiry takes part over successive rounds; this was seen as something of a forced way to reach convergence of opinions, and might have lead some experts to drop out after the first round. However, the implementation of Delphi requires that the inquiry take part over successive rounds. Avoiding successive rounds would have meant that the designer was deliberately not following the necessary requirements of the Delphi method, thus it would have not be possible to evaluate the methodological principles of the Delphi technique. Finally, the evaluation of the Delphi demonstrates that a limitation of this exercise was seen as being the lack of consideration of possible interrelations and interdependencies among different items. This is a problem with practically all Delphi exercises: to overcome it means moving well beyond standard Delphi practice.

The evaluation of the Delphi method reports that this implementation was considered particularly useful to better understand the future trends of ATT technologies at European level rather than at National or Local level. Moreover, the experts found that the results achieved by this inquiry formed useful information to provide evidence to help transport policy decisions, and to help guide new research on ATT technologies.

A remarkable result of the Delphi evaluation was that all participants felt that results would be useful for transport policy makers. Of course, the degree of this perceived usefulness varies among experts, and it was also reported that the achievements
obtained are more useful for policy makers which deal with transport policy at European level, rather than at National and or Local level. The results were especially seen as helpful for the transport policy as a source of background information, to identify key technologies, and setting priorities in the field of ATT technologies.

The evaluation of the implementation and of the results the SMIC method was undertaken by a much smaller sample of experts. Nevertheless, the results have been considered valid for the evaluation of the method, even if there was awareness that the results could be a source of positive bias (i.e. only the experts that were favourably to the SMIC method replied to the evaluation questionnaire). Again, for the SMIC method, none of the experts reported that there were many gaps in this implementation and in the results achieved. This means that experts found rather comprehensive the implementation and the outcomes.

As a general observation, experts considered the task of completion of SMIC questionnaire more difficult than the Delphi one. This was further influenced by the formulation of the selected event statements. However, experts regarded that one of the major advantages of the technique was its precision. This can be considered a characteristic of the method itself because it follows strict ground-rules, but it was also a merit of the approach, which provided very clear instructions and explanation on the purposes of the implementation of the method. As reported for the Delphi method, the study was evaluated by experts as a useful occasion for a large sample of experts to provide and/or share an overview of their opinions on future trends of ATT technologies. Further, the presentation of results, through means of possible scenarios and the explanation of the more likely ones, was considered as an interesting way to evaluate results obtained from the method, (independently of whether the experts agreed with its outcomes).

On the other hand, one of the advantages of the method, which is the assessment of interrelationships among events, was evaluated by experts as complex and time consuming. The major problem was that sometimes experts reported that it was difficult to assess cross-impacts, because they had the feeling they were contradicting opinions expressed in previous answers. This reflects two problems of the methods. The first is cognitive: the difficulty of keeping in mind the factors that led to previous decisions. The second problem is methodological. The method only allows for twofold cross-impacts to be assessed. But more complex interactions may occur as a result of factors co-occurring. This gets hidden by the method, but probably does enter respondents’ consciousness. Moreover, due to the specific implementation of the method (i.e. that was based on the results of the first round of the Delphi inquiry), and the formulation of the event statements, answers were very much dependent on the degree of expertise of the experts. This result is interesting. In fact, whereas in the context of the Delphi method the lack of interrelations and interdependencies was seen as a limiting factor of the implementation, at the same time its presence was criticised in the context of the SMIC method because it was difficult to tackle.

The two forecasting methodologies in retrospect
A general observation is that these (and other) forecasting methods based on eliciting expert knowledge have always been treated separately in the literature, with little comparative analysis. This is somewhat reasonable, since each technique is based on different methodological principles. However we could ask the question whether these methods should be comparable or whether they are intended to provide information for
different purposes. In any case both Delphi method and the SMIC method, have among their final goals the determination (and even development) of the collective cognition, which mean that they aim to elicit and share the best knowledge of expert views on an issue.

Each of the two techniques discussed in this study use survey tools to extract expert knowledge. During the implementation of the two methods, in some respects the two techniques raised the same type of issues and questions. For example, both Delphi and SMIC confront the same problem, which is ‘how many’ experts should receive the inquiry. Then, issues such as how to build a clear event statement, and which is the ‘best’ time horizon for a reasonable forecast, have to be considered in the application of Delphi and SMIC. For these reasons, we believe there is scope for trying to harmonise these commonalities, faced in practically any implementation of this type of techniques. Further research could examine in how far it is possible to have common ‘rules’ for the implementation of experts’ based methods.

Compared to the Delphi method, relatively little has been written about the way that cross-impact method should be implemented. It is also true that compared with the Delphi method, cross-impact analysis saw less application. This might reflect the more limited flexibility to adapt topics to cross-impact analysis, than to the Delphi method. However, some of the characteristics of cross-impact analysis, such as the possibility to check interrelationships among events, can be a useful tool to apply to forecasting well-defined topics quite generally. Thus, it could be useful to have some guidelines that can help developing ‘best practice’ implementations of these methods. In the literature, there is little indication of which type of event statements is preferable to or not to combine in the same cross-impact inquiry. For example, it is not specified whether is better to avoid mixing among events that are technology-related, with some other that are more economics-oriented, because this might invalidate the outcomes by making impossible the cross-impacts among events.

Another drawback that emerged in the implementation of the SMIC method, and where there is still scope for further research, is the difficulty of establishing the consistency of the technique. It is hard for the non-statistically expert user, to understand the underlying mathematical structure of the method, and also the consequent computational routines performed by the software This difficulty is mainly due to lack of transparency about the calculations performed by the software, it is difficult to check how the software process the data. It is not clear, whether the method, presented with consistent data, is able to generate consistent results, if the assumptions about how experts make choices and judgements on events and trends are themselves not consistent. One way to check consistency in data is to apply a weight to experts’ estimations on the basis of their expertise. Even if the SMIC software allow weighting data by assigning to the experts a weight from one to five, we decided to avoid using weights for two reasons. In the Delphi method data were treated without any weight system, thus for the SMIC data the same assumption was applied. The latter reason is that – again – the explanation in the SMIC literature and in the SMIC software handbook were not clear enough to understand which were the criteria to apply in order to assign weight to experts’ replies. In a sense, the user is an onlooker that is expected to trust ‘the black box’. However, the difficulty in checking consistency may hinder would-be users, who as a result are likely to apply another and more transparent technique.
A particular matter in relation to the SMIC software, is that, if this is to be (as intended) a tool to reduce the efforts required from the user of the technique, a more user-friendly instruction handbook needs to be developed.

Notwithstanding, the numerous applications of the Delphi method, fewer studies have been undertaken to clarify aspects of implementation of this technique. Of course, there were many attempts and discussions in the literature on the advantages, drawbacks and pitfalls of this technique. But we are not aware of the existence of a technical template that provides answers to the various questions the user has to face when dealing with the implementation of this method. We are not asserting that is easy to provide such a template. To create this would be very ambitious, but some general guidance notes would ease (and harmonise) the implementation task of this technique.

For example, from the evaluation of our Delphi implementation it was apparent that the experts would have liked to weight responses. In the past, self-rating of expertise has been assessed as tool to weight responses, and some practitioners are dubious about any formal weighting scheme (preferring simply to exclude non-experts, or at least to contrast the judgements of experts with that of the whole sample). However, the literature does not provide enough evidence that self-rating is an appropriate indicator to weight responses. It is possible to further investigate whether it is possible to find a standard tool that allow to weight expert responses, and under what circumstances this might be applicable.

Another aspect of Delphi that suggests further research is related to the degree of consensus. It can be useful to have indication on ‘how much’ consensus is necessary or enough to make a forecast believable or useful. The lack of this indication forces the user of the Delphi method to establish arbitrary criteria to assess if convergence of opinions was reached. Possibly, either a theory-based or statistically derived standard to assess consensus would certainly be helpful and powerful. In any case, as already mentioned in the previous chapters the search of convergence of opinion should not be an overriding objective of a Delphi study. A lack of consensus is itself useful knowledge, since knowledge of disagreement can be valuable.

**Final conclusions and policy discussions**

On the basis of the experience gathered with this implementation of the Delphi and the SMIC methods, and the outcomes that follows from this, there is space and there is a role for forecasting in the policy decision process. Similarly, there is a role for methods based on soliciting experts’ opinion as an aid to inform the policy decision process. However, we are still far to have an optimal and automatic system that allows a systematic input (let alone use) of forecasting results in the policy decision process.

This study does not have a satisfactory reply (we have doubts if there is such reply) to the question of ‘how’ to select the best forecasting methodology from among the existing ones. It is evident that the choice of the application of a methodology depends in a first place from the topic, but especially on the objectives of a study. In principle, methodologies based on soliciting experts’ opinions are applicable to any topic. Of course the degree of flexibility of these techniques varies, and the boundary conditions of the methods can constrain the freedom of the user (as the results of the implementation of the two methods showed in this study). In order to optimise the process of eliciting the ‘best’ knowledge of experts’ view on an issue, the user of the technique has to face and implement many decisions, where a substantial degree of tacit knowledge is necessary, both on the method in terms of its mechanics, but also on the forecasted topic. It is possible that a technique by itself influences the process of best
knowledge elicitation. This is likely the input data are in some way affected by the boundary conditions of the technique, and/or the way data are processed and filtered, there might be the possibility that some information can be lost. Equally, the implementation of the same technique, to a same topic, and using the same experts but implemented from two different designers may influences the process of ‘best knowledge’ elicitation. In this case, the necessary degree of tacit knowledge (on the topic and on the technique) is likely to play a difference. Consequently different techniques, even if based on soliciting experts opinions through mean of inquiries and applied on the same topic, are likely to achieve different levels of knowledge and information. This is attributable to the fact that each technique has a different process of gathering and analysing information, but it is possible to use and combine different methodologies to confirm results, achieve different types and different degrees of detail of knowledge and information. Finally, for all the reasons discussed above, the decision sequence required for the implementation of expert-based methods, may influence the usefulness of a forecast. However, the presence of a technical template with some ground-rules on implementation can facilitate the task of the user, could avoid to run into the pitfalls of the technique, and elude useless applications of these techniques.

The outcomes of this study can be further disseminated and used. The researcher believes that the results can be considered as a catalyst to cross-fertilise ideas and knowledge on the future applications of transport telematics technologies and their implications. The results of the study can be considered a way to assess skills, expertise and requirements on Research and Development (R&D) of the area of ATT systems. For transport experts the implementation of the methods was a useful tool to share knowledge. The results of the implementation can be used as a basis of further work and research in the transport field, for example it is possible to use the results and assess them more specifically on a National level. The same inquiry could be undertaken at a National level focusing on cities of a selected country. This can be useful to confirm and compare results and especially to have a framework, which can help underlying the constraints and the requirements to the adoption of ATT systems on a National basis.

It was beyond the task of the researcher for this study, to undertake a consultation exercise directly with European transport policy makers. A consultation exercise allows to assess their opinion on the usefulness of the outcomes, both in terms of provision of knowledge in relation to the field of ATT and in terms of whether it is possible to foresee, on the basis of the results, possible actions or recommendations. This exercise might still be undertaken, and it can be conducted through a policy workshop where results can be used as a basis for debate or through personal interviews with policy makers.

As a final conclusion, we would like to use the experience and outcomes of this study to generate some discussions that aim to contribute to the discipline of Futures studies. The lesson that we have learned is that in terms of development of (published) Futures studies methodologies, not much explicit progress has been made from the 1970s on the analysis of the implementation process and its implications, and here we refer to forecasting techniques based on eliciting experts’ opinions. The field of Futures studies is lacking of this of reflections and research. In a sense, this is surprising since from the beginning of the 1990s the field of Futures studies – especially through Foresight exercises very often undertaken at very broad level – is now experiencing a very active moment. Moreover, these studies, which involve considerable amount of money, are often based on the Delphi method. The discipline of Futures studies is progressing on the practical applications of techniques, but apparently there was not a similar progress
on the more conceptual aspects of the forecasting process. In like manner, not much progress has been done on procedures for filtering, transforming and translating future images into a forecast usable in the decision making process. It is still a complicated task to translate the results of a forecasting exercise into key variables and turning points, which can be used as a basis to indicate that one or other specific intervention at one or another moment, should have a specific impact on the possible future.

Bibliography:


Role of experts in FTA

- Expert opinion and judgments are useful to provide policymakers with knowledge on complex domains especially dealing with S&T
  - when no historical data exist
  - impact of external factors is more important than the factors that could govern the development of the technology
- Forecasting techniques based on expert knowledge aims to overcome constraints of more traditional means of consulting experts

Questions

- How to make an informed choice when applying a forecasting method?

- Do different methods provide similar results when applied to the same area? Is comparison of techniques possible (desirable?)
Implementing the techniques

The implementation of the techniques requires to face a number of issues and decisions

- Generic issue to do with elicitation of expert opinion
- Specific issues posed by this class of methods
- Specific issues related to the specific method
- Specific issues related to the application of the method to a specific area

*How can we reach and define that good practice is applied when this type of techniques are implemented?*

The process

- Objective of the study
- Choice of the subject
- Selection of the experts
- 1st round Delphi method
- Analysis of results
- SMIC questionnaire
- Evaluation questionnaire
- Analysis of results
- 2nd round Delphi method
- Evaluation questionnaire
- Final outcomes
**Delphi implementation**

<table>
<thead>
<tr>
<th>ISSUE IN STRUCTURE OF DELPHI</th>
<th>‘TYPICAL’ IMPLEMENTATION OR CHOICES AVAILABLE</th>
<th>IMPLEMENTATION IN THIS STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall framework</td>
<td>Often diffuse</td>
<td>Background scenario + medium-sized city defined</td>
</tr>
<tr>
<td>Nature of experts</td>
<td>Typically national</td>
<td>Cross-national</td>
</tr>
<tr>
<td>Topic forecast</td>
<td>Few words (recommended = most 25)</td>
<td>Relatively many words (average = from 7 to 29 words)</td>
</tr>
<tr>
<td>Nature of the forecast</td>
<td>When ‘X’ will happen?</td>
<td>How far ‘X’ will happen in a given time?</td>
</tr>
<tr>
<td>Time horizon</td>
<td>Different time horizons</td>
<td>Fixed time horizon</td>
</tr>
</tbody>
</table>

**Characteristics of Delphi and SMIC inquires**

<table>
<thead>
<tr>
<th></th>
<th>DELPHI METHOD</th>
<th>SMIC METHOD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Background scenario</td>
<td>same</td>
<td>same</td>
</tr>
<tr>
<td>Description of the city</td>
<td>same</td>
<td>same</td>
</tr>
<tr>
<td>Number of topics</td>
<td>19</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>(4 taken from 1st round Delphi)</td>
<td></td>
</tr>
<tr>
<td>Inputs to questionnaire</td>
<td>same</td>
<td>same</td>
</tr>
<tr>
<td>Number of rounds</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Probability</td>
<td>occurrence by date</td>
<td>occurrence by date conditional probability</td>
</tr>
<tr>
<td>Number of experts</td>
<td>224 (first round) 87 (second round)</td>
<td>126 77 (only SMIC) 49 (SMIC and Delphi)</td>
</tr>
</tbody>
</table>
Some important issues for the implementation

- Number of experts to involve in an inquiry
  - How to ensure that relevant competencies are represented?

- The construction of event statements
  - Especially important for SMIC. Experts assess simple probability of occurrence and probability of pairs of events

- Accuracy and validity of results
  - Convergence of opinions in Delphi
  - Mathematical and statistical computation and manipulation of data in SMIC

Conclusions

- Comparison was on one particular version of Delphi and SMIC

- Little systematic attention to conceptual development of the field (i.e. how to implement methods)
  - It is necessary a substantial degree of tacit knowledge both in terms of mechanics but also on the forecasted topic

- Further research on methods is desirable
  - To examine if it possible to harmonise commonalities of experts’ based methods using inquiries
  - To develop a technical template with some guidance notes on implementation of this type of techniques