Technological Forecasting & Social Change 105 (2016) 63–76

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

Biases in future-oriented Delphi studies: A cognitive perspective

Jens Winkler [a,*], Roger Moser [b]

[a] Research Institute for International Management (FIM-HSG), University of St. Gallen (HSG), Dufourstr. 40a, 9000 St. Gallen, Switzerland
[b] Research Institute for International Management (FIM-HSG), University of St. Gallen (HSG), Dufourstr. 40a, 9000 St. Gallen, Switzerland & Adjunct Professor IIM Udaipur, India

ARTICLE INFO

Article history:
Received 20 August 2015
Received in revised form 12 January 2016
Accepted 27 January 2016
Available online 13 February 2016

Keywords:
Cognitive biases
Framing
Anchoring
Desirability bias
Bandwagon effect
Belief perseverance

ABSTRACT

Delphi is an established information gathering and forecasting approach that has proven to deliver valuable results in a wide variety of specialist fields. Yet, Delphi studies have also continuously been subject to critique and doubt, particularly concerning its judgmental and forecasting accuracy. To a large part this can be attributed to the substantial discretion researchers have in their design and implementation. Awkwardly designed Delphi studies may lead to severe cognitive biases that adversely affect the research results. This paper takes a cognitive perspective by investigating how different cognitive biases take effect within future-oriented Delphi studies and how their unfavorable impacts can be mitigated by thoroughly adapting specific Delphi design features. The analysis addresses cognitive biases affecting panelists' initial estimates — namely framing and anchoring as well as the desirability bias — as well as such cognitive biases taking effect during feedback and revision loops — namely the bandwagon effect and belief perseverance.

1. Introduction

Delphi processes have been used for decades in a variety of fields and methodological variations. As a structured information gathering and forecasting approach it still enjoys unabated interest as indicated by recent applications (Wester and Borders, 2014; Álvarez et al., 2014) and design considerations (e.g. Förster and von der Gracht, 2014; Gallego and Bueno, 2014). Delphi studies regularly deliver accurate and valuable results (e.g. Holmes et al., 2002; Parente and Anderson-Parente, 2011) but continue to be criticized as well. The major concern of practitioners and academics is Delphi's judgmental and forecasting accuracy (Lin et al., 2014; Parente and Anderson-Parente, 2011; Fildes and Goodwin, 2007). Researchers investigating the impact of different design features, e.g. statistical vs. argumentative feedback, on Delphi results' accuracy found contradictory results (e.g. Rowe et al., 2005; Rowe and Wright, 1996). However, these studies frequently do not apply a strong cognitive perspective on Delphi processes, i.e. they do not link the design choices to cognitive processes and biases they may cause or mitigate. Therefore, we argue that more conceptual and empirical work in this area is required as Delphi's accuracy depends on i) how researchers use (or abuse) their high degree of discretion in terms of study design and execution (Rowe and Wright, 1999; Story et al., 2001), and ii) to which extent several cognitive biases take effect at different stages of the process; the latter being to a large part dependent on the former.

As research in psychology and cognition sciences has identified hundreds of biases that could potentially take effect in some Delphi constellation as well, it is beyond the scope of a single study to provide a comprehensive overview on all cognitive biases without being overly superficial. We therefore decided to elaborate on the four cognitive biases encountered by Delphi participants that seem to be most frequent and most impactful in Delphi applications, namely framing and anchoring, the desirability bias, the bandwagon effect, and belief perseverance. We believe that researchers controlling for these biases via specific Delphi design decisions could not increase accuracy much further by controlling for additional cognitive biases. Following our focus on participants' cognitive biases, we do not address other issues such as sampling biases, i.e. the selection of proper experts, that have been studied elsewhere (e.g. Okoli and Pawlowski, 2004; Rowe and Wright, 1999). Concerning the kind of Delphi studies, our analysis is focused on expert-based future-oriented studies looking at least five years ahead from today. For researchers applying other kinds of Delphi studies, for example such including laymen assessing contemporary almanac questions, other cognitive biases might be of relevance.

We structure our analysis along the typical process steps of a Delphi study, and indicate which biases may occur at which stage of the process. As illustrated in Fig. 1, framing and anchoring as well as the desirability bias impact experts' first estimates, while the bandwagon effect and belief perseverance come into effect during stage 4 which includes feedback and potential revisions of estimates. As participants

* Corresponding author.
E-mail addresses: jens.winkler@student.unisg.ch (J. Winkler), roger.moser@unisg.ch (R. Moser).

http://dx.doi.org/10.1016/j.techfore.2016.01.021
0040-1625/

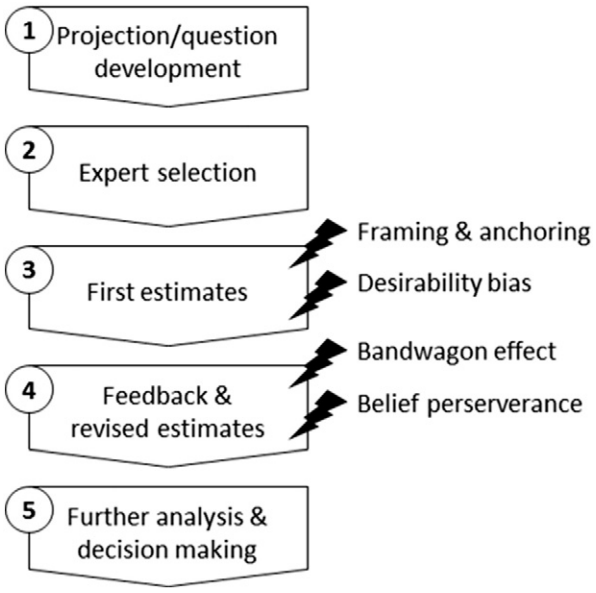

Fig. 1. Delphi process stages and selected cognitive biases.

are usually not involved in process steps 1, 2, and 5, cognitive biases on their side only occur in stages 3 and 4.

By bridging literature on Delphi research and the fields of cognition and psychology we hope to make a methodological contribution that is of value to both academics and practitioners applying Delphi studies in a variety of fields. We do so by a) discussing different cognitive biases and their modes of operation during Delphi applications, b) elaborating on the impact of certain design choices on the prevalence of cognitive biases in future-oriented Delphi processes and c) developing design recommendations that aim to mitigate or avoid the negative effects of cognitive biases and work towards increasing Delphi accuracy. Hence, our main research question is: How should future-oriented Delphi studies be designed in order mitigate (or avoid) the negative effects of participants' cognitive biases?

2. Delphi

The Delphi methodology is a structured, interactive group communication and judgmental forecasting process aiming at systematically exchanging informed opinion concerning an uncertainty-bearing field of interest among a panel of selected experts and developing consensual understanding that reduces uncertainty and finally enhances decision quality (Hallowell and Gambatese, 2010; Donohoe and Needham, 2009; Dunn, 2004; Linstone and Turoff, 1975). In future-oriented Delphi studies the field of interest may concern issues lying as far as several decades ahead. Delphi rests on the assumption that structured group approaches provide more accurate judgments than a single expert (Donohoe and Needham, 2009; Linstone and Turoff, 1975) and are more appropriate than traditional meetings (Graefe and Armstrong, 2011). Another underlying assumption is that, even in highly uncertain environments, some features of the future are predetermined and foreseeable (Walsh, 2005).

A typical Delphi proceeds as follows. After designing a survey of questions or projections (Klenk and Hickey, 2011) it is sent to a group of experts, each of whom provides individual evaluations, ratings or rankings (Chiravuri et al., 2011), e.g. concerning the probability or feasibility of the items under investigation (Klenk and Hickey, 2011). Additionally, experts may be asked to provide qualitative arguments supporting their individual estimates (Graefe and Armstrong, 2011). Having received all answers, the Delphi administrators consolidate and analyze the contributions (Klenk and Hickey, 2011) and feed the results back to the experts, sometimes with a reworked questionnaire. Respondents are asked to review the estimations (and arguments, if any) of the other anonymous participants (Hallowell and Gambatese, 2010), encouraged to reconsider their own contributions (Landeta and Barrutia, 2011; Sharfman and Shaft, 2011; Linstone and Turoff, 1975) and given the opportunity to revise their estimates (Georgantzas and Acar, 1995; Rowe and Wright, 1999). This process can be repeated several times until a pre-determined criterion, e.g. a certain level of consensus, is met (Klenk and Hickey, 2011).

Such a procedure comprising at least one round of reconsidering and possible adaption of prior estimates (Landeta and Barrutia, 2011) constitutes the iterative character of Delphi that allows for accuracy-improving social learning (Dunn, 2004; Hallowell and Gambatese, 2010) and the reduction of noise (Strauss and Zeigler, 1975) but also bears the risk of cognitive biases taking unfavorable effect as discussed below. Besides its iterative fashion the key characteristics of Delphi are controlled feedback, and the anonymity of participants (von der Gracht, 2008; Elliott et al., 2010; Story et al., 2001; Rowe and Wright, 1999; Georgantzas and Acar, 1995; Yang et al., 2012). Controlled feedback means that the Delphi administrators decide on how feedback is provided and which aspects of the group's responses are included (von der Gracht, 2008).

Anonymity of participants is probably the most controversially discussed characteristic of Delphi as it brings along a number of advantages but also drawbacks. In general it is said that Delphi uses the positive attributes of structured group interaction while mitigating or avoiding the negative social, psychological, and power effects of direct confrontation (Kauko and Palmroos, 2014; Graefe and Armstrong, 2011; Klenk and Hickey, 2011). To be more concrete, anonymity avoids experts' statements to be biased by dominant personalities, panelists from higher hierarchy level or social status ("halo effect"), or such with strong oratorical abilities (Landeta and Barrutia, 2011; Tersine and Riggs, 1976). Furthermore, anonymity creates a free thinking space that reduces the unwillingness to give estimates on uncertain issues (Häder, 2002), encourages to express and challenge unconventional opinions and alternative viewpoints (Donohoe and Needham, 2009), and offers the opportunity to change a stand once taken without losing face (von der Gracht, 2008; Rowe and Wright, 1999; Okoli and Pawlowski, 2004).

Delphi is best suited to fields and circumstances of application where objective factual data is scarce (Gray and Hovav, 2008; Daft and Lengel, 1986) and knowledge necessary to make profound decisions is incomplete (Skulmoski et al., 2007; Amos and Pearse, 2008). Delphi is highly valuable in situations of severe uncertainty stemming from rapidly unfolding, non-calculable dynamics, or uncertainty originating from large multidisciplinary problems in highly complex environments (Yang et al., 2012; Donohoe and Needham, 2009; Ziglio, 1996). In these situations precise analytical data processing techniques are not applicable (Melnyk et al., 2009; Donohoe and Needham, 2009; Ziglio, 1996) and trend extrapolation is mostly inadequate (Melnyk et al., 2009; Linstone and Turoff, 1975). Instead, information collection and knowledge must be built on informed opinion and subjective expert judgments as well as experience-based interpretations (Yang et al., 2012; Melnyk et al., 2009; Linstone and Turoff, 1975).

The major disadvantages attributed to the Delphi methodology comprise difficulties in assessing participants' level of expertise, the potential of anonymity and iteration to lead to compromise rather than consensus, and limitations in assessing result accuracy and reliability — particularly when an issue in the long-term future is investigated (Story et al., 2001).

Comparisons of Delphi and other techniques in terms of accuracy came to discordant results. Although there are several examples of Delphi studies delivering accurate results (Czinkota, 1986; Czinkota and Ronkainen, 1992, 1997, 2005; Gray and Hovav, 2008; Holmes et al., 2002; Parente and Anderson-Parente, 2011), and some researchers come to the conclusion that "Delphi's effectiveness over comparative procedures, at least in terms of judgmental accuracy, has