EDITED BY
LUIZ MOUTINHO
MLADEN SOKELE

INNOVATIVE RESEARCH METHODOLOGIES IN MANAGEMENT

VOLUME II: Futures, Biometrics and Neuroscience Research

Innovative Research Methodologies in Management

Luiz Moutinho • Mladen Sokele
Editors

Innovative Research Methodologies in Management

Volume II: Futures, Biometrics and Neuroscience Research

palgrave
macmillan

Editors
Luiz Moutinho
University of Suffolk, Suffolk, England, UK
The University of the South Pacific, Suva, Fiji

Mladen Sokele
Faculty of Electrical Engineering and Computing
University of Zagreb
Zagreb, Croatia

ISBN 978-3-319-64399-1 ISBN 978-3-319-64400-4 (eBook)
DOI 10.1007/978-3-319-64400-4

Library of Congress Control Number: 2017954805

Printed on acid-free paper

This Palgrave Macmillan imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To everyone in the world that kindly resonates with my thoughts….
Luiz Moutinho

To my dear son Leo
Mladen Sokele

Preface

I am very grateful to Palgrave for the fact that they were enthusiastic about this project. It has been one of my scholarly "pet" projects for some time. I was also extremely happy that I was able to secure the incredible collaboration of my friend Mladen Sokele as my co-editor of this book.

Methodologies are the basis for scientific research. However, scholars often focus on a very limited set of methodologies, partly due to a lack of knowledge about innovative methodologies outside their area. This is especially true in the area of management science. Providing management scholars with an education about methodologies outside their primary area of expertise is the goal of the proposal made to Palgrave Macmillan, global academic publisher, for a book providing a comprehensive presentation of innovative research methodologies: Palgrave's *Innovative Research Methodologies in Management***.**

This book is to be positioned as a seminal collection of mind-stretching and thought-provoking research methodology essays. Hopefully, these research methods and topics will greatly enhance the research methodology "armoury" of management scholars and alter the research "modus operandi" of academic research output and published work.

One of the aims of the *Innovative Research Methodologies in Management* text is to identify and foster methodological research innovation in the academic management field. This book project seeks out research practices that are not highlighted through the majority of academic research outlets and journals, that are not highlighted typical research method courses or and to have an impact on the research process in novel ways. These innovative methodologies usually entail a combination of (i) technological innovation, (ii) the use of existing theory and methods in new ways, and (iii) interdisciplinary approaches. The project's focus on innovative research practices will range across new philosophical insights into academic research, new experimental designs to new technological research contexts, and new analytical techniques. Departing from the somewhat perennial situation of academic research methodolatry and scientism, this tome will be focusing on a series of rigorous methodological advances in many areas. The primary focus relies on emerging and bleeding-edge areas of academic research methodologies in management, making the contents of this book an authoritative source on the applications of these methodologies to management science.

Volume 1 is dedicated to the coverage of innovative research methodologies within the realms of research philosophy, research measurement, and research modeling.

Volume 2 is focused on chapters dealing with Futures Research, Biometrics Research, and Neuroscience Research in Management.

Chapter 1 (Zarkada, Panigyrakis, and Tsoumaka) introduces a panoply of metamodern mixed methods in management research dealing with Web 2.0+. It discusses metamodern socioeconomic phenomena, mixed methods designs, and sampling integrity, among many other topics. It follows an interesting light analogy and is very robust in terms of theoretical content. Interesting reflections are also included.

In Chap. 2, Hackett and Foxall tackle the topic of neurophilosophy. This is a chapter with a rich theoretical background. There is an interesting section on psychological constructs and neurophysiological events. There is also a challenging exploration into mereological understandings.

Kip Jones (Chap. 3) deals with emotivity and ephemera research. The content is focused on arts-led and biographical research as well as relational aesthetics. There are some interesting insights into neo emotivism. There is a challenging section on performative social science. There are also interesting comments on experimentation and the experimental redux.

Chapter 4 presents the novel approach—Abductive Thematic Network Analysis (ATNA) using ATLAS-ti written by Komalsingh Rambaree. This chapter introduces ATNA as a methodological approach for qualitative data analysis. It starts by providing a brief description on abductive theory of method and thematic analysis method. Then, it highlights how the two methods are combined to create ATNA. Using a qualitative dataset, this chapter demonstrates the steps in undertaking ATNA with a computer-aided qualitative data analysis software—ATLAS-ti v.7.5. The chapter concludes that ATNA provides to researchers a much-needed pragmatic and logical way of reasoning, organizing, and presenting qualitative data analysis.

Sullivan, Lao, and Templin (Chap. 5) deal with diagnostic measurement. With diagnostic measurement, the aim is to identify causes or underlying properties of a problem or characteristic for the purposes of making classification-based decisions. The decisions are based on a nuanced profile of attributes or skills obtained from observable characteristics of an individual. In this chapter, the authors discuss psychometric methodologies involved in engaging in diagnostic measurement. They define basic terms in measurement, describe diagnostic classification models in the context of latent variable models, demonstrate an empirical example, and express the broad purpose of how diagnostic assessment can be useful in management and related fields.

Yang and Fong (Chap. 6) explore the issues of incremental optimization mechanism for constructing a balanced, very fast decision tree for big data. Big data is a popular topic that highly attracts attentions of researchers from all over the world. How to mine valuable information from such huge volumes of data remains an open problem. As the most widely used technology of decision tree, imperfect data stream leads to tree size explosion and detrimental accuracy problems. Over-fitting problem and the imbalanced class distribution reduce the performance of the original decision tree algorithm for stream mining. In this chapter, the authors propose an Optimized Very Fast Decision Tree (OVFDT) that possesses an optimized node-splitting control mechanism using Hoeffding bound. Accuracy, tree size, and the learning time are the significant factors influencing the algorithm's performance. Naturally, a bigger tree size takes longer computation time. OVFDT is a pioneer model equipped with an incremental optimization mechanism that seeks for a balance between accuracy and tree size for data stream mining. OVFDT operates incrementally by a test-then-train approach. Two new methods of functional tree leaves are proposed to improve the accuracy with which the tree model makes a prediction for a new data stream in the testing phase. The optimized node-splitting mechanism controls the tree model growth in the training phase. The experiment shows that OVFDT obtains an optimal tree structure in numeric and nominal datasets.

Sokele and Moutinho (Chap. 7) introduce the Bass model with explanatory parameters. Over the 45 years, the Bass model is widely used in the forecasting of new technology diffusion and growth of new products/services. The Bass model has four parameters: market capacity; time when product/service is introduced; coefficient of innovation; and coefficient of imitation. Although values of coefficient of innovation and coefficient of imitation describe the process of how new product/service gets adopted as an interaction between users and potential users, their explanatory meaning is not perceptible. These authors explore this important gap.

Chapter 8 by Volker Nissen is titled “A Brief Introduction to Evolutionary Algorithms from the Perspective of Management Science.” Summarizing, it is useful to differentiate between several perspectives. From a methodological point of view, the myriad of nature-inspired heuristics is rather confusing for casual users and definitely not helpful in creating more acceptance for metaheuristics in practical applications. Moreover, there is evidence (e.g., Weyland 2015; Sörensen 2015) that at least some of the nature-inspired concepts published recently (such as Harmony Search) are rather old wine in new skins. One should better look for over-arching and bearable concepts within evolutionary algorithms (EA) and related metaheuristics, putting together a common framework (or toolbox) that integrates different options of solution representation, search operators, selection mechanisms, constraint-handling techniques, termination criteria, and so on. Fortunately, such frameworks are available today. Then, properly choosing the components for a hybrid heuristic from such a framework requires a deep understanding of which components actually fit together and work well for certain classes of problems or search spaces.

Moreover, following the No-Free-Lunch Theorem (Wolpert and Macready 1997), problem-specific fine-tuning of heuristics remains important to achieve truly good results. Today, much of this is still more an art than a science, despite helpful textbooks like Rothlauf (2011). As a consequence, there are lots of interesting research issues to be solved along these lines. This process is indeed ongoing for several years now, but, as Sörensen et al. (2016) point out, also bypassed by useless initiatives to invent ever "new" nature-inspired heuristics.

For more than 20 years now, EA have become integrated in business software products (e.g., for production planning), so that, as a result, the end user is often unaware that an evolutionary approach to problem solving is employed (Nissen 1995). Today, large software companies like SAP use EA in their enterprise software. However, since customizing options are limited in these systems, it appears fair to say that the full power of EA is frequently not unleashed by such standardized approaches. This confronts us with a dilemma. Simple forms of EA that can be fairly easily understood and applied widely are of only limited power. If we want to use metaheuristics like EA to full extent, then this requires knowledge and experience in their design and application. Most users have neither the qualification nor the time to dive that deep into the matter. According to my own observations as an IT consultant, this unfortunately also holds for most consultants that could potentially help customers in applying modern heuristics. Thus, creating really powerful applications of EA today is frequently an issue for only a small number of highly specialized IT companies. This situation is unsatisfactory and could only be changed if the design and application of modern heuristics becomes an important topic in general management studies at universities and related higher learning institutions. The author argues that this should indeed be the case, because in today's digital era, there is strong evidence that we are entering an age of knowledge-based competition where a qualified workforce that is able to creatively use modern tools for data mining, ad-hoc reporting, heuristic optimization, artificial intelligence, and so on will make the difference in many branches of industry.

Volker Nissen also kindly contributes to Chap. 9 on applications of EA to management problems. EA (or Evolutionary Computation) represent as nature-inspired metaheuristics a branch of computational intelligence or soft computing. Their working is based on a rough abstraction of the mechanisms of natural evolution. They imitate biological principles, such as a population-based approach, the inheritance of information, the variation of solutions through crossover and mutation, and the selection of individual solutions for reproduction based on fitness. Different variants of EA exist, such as genetic algorithms, evolution strategies, and genetic programming.

This chapter reviews EA from the perspective of management applications where "management" indicates that predominantly economic targets are pursued. In general terms, the preferred area of application for EA, and other metaheuristics as well, is optimization problems that cannot be solved analytically or with efficient algorithms, such as linear programming, in a reasonable time or without making strong simplifying assumptions on the problem. Many of these problems are of a combinatorial nature, such as job shop scheduling, timetabling, nurse rostering, and vehicle routing, to name just a few. In practical settings, often the issue of "robustness" of a solution is equally important as "optimality," because the optimization context is characterized by uncertainty and changing conditions.

While the first variants of EA were already invented in the 1960s, it is in the last 15–20 years that these powerful methods of heuristic optimization have attracted broader attention also outside the scientific community. Providers of optimization software such as MathWorks with MATLAB, but also large players in enterprise resource planning (ERP), such as SAP AG with SAP APO, integrated EA in their canon of software-based optimization methods.

However, following the "No-Free-Lunch Theorem," metaheuristics such as EA always require a certain degree of adaptation to the individual problem at hand to provide good solutions. In practice, this can pose a problem for the inexperienced user of EA-based software products. On the scientific part, lately there appears to be an inflation of nature-inspired metaheuristic approaches, some of them related to EA and partly criticized for a lack of novelty.

The last chapter of Volume 1 is written by Beynon, Moutinho, and Veloutsou (Chap. 10). It is an exposition of the role of consideration sets in a Dempster-Shafer/Analytic Hierarchy Process (DS/AHP) analysis of consumer choice. Consumer behavior is often perceived through the notion of consideration sets. However, realistic modeling of consumer choice processes identifies impeding factors, including ignorance and non-specificity. In this chapter, the appeasement of these factors and the role of consideration sets are considered through the utilization of the nascent DS/AHP method of choice analysis. The central element in the DS/AHP analysis is the body of evidence (BOE), with certain BOE constructed at different stages in the analysis, then a number of different sets of results can be found. The research content is attempting to convey a more realistic approach for the individual consumer to undertake the required judgment-making process. The investigation is based on a group of consumers and their preferences on a number of cars over different criteria. The notion of consideration sets is shown to be fundamental within DS/AHP, and a novel approach to the aggregation of the preferences from the consumers is utilized. A notional approach to the identification of awareness, consideration, and choice sets is described, based on the levels of belief and plausibility in the best car existing in a group of cars, which could be compared with the algorithm developed by Gensch and Soofi (1995).

Volume 2 of *Innovative Research Methodologies in Management* starts with Chap. 1 by Simone Di Zio on the "Convergence of Experts' Opinions on the Territory: The Spatial Delphi and the Spatial Shang." The judgments of a panel of experts are of extreme usefulness when, in front of a decision-making problem, quantitative data are insufficient or completely absent. Experts' opinions are helpful in forecasting contexts, for the detection of innovative solutions or for the verification and refinement of consensus on objectives or alternative scenarios.

The way the views are collected is crucial, and without a rigorous methodology, any consultation process may become vain. In literature, there are many methods, but some are used for the ease of application rather than that for their scientific properties. Methods such as focus group, face-to-face interview, or online questionnaire are very popular but have quite important drawbacks.

Many of those disadvantages are overcome by the methods of the "Delphi family," whose prototype is the Delphi method, which involves the repeated administration of questionnaires, narrowing the range of assessment uncertainty without generating errors that result from face-to-face interactions. To date, the Delphi technique has a very high number of applications, and its success has produced a wide range of methods that are its variants.

In this chapter, we present two recent variants, called Spatial Delphi and Spatial Shang, applicable when consultations, and consequent decisions, concern matters of spatial location. The judgments of the experts are collected by means of points placed on a map, and the process of the convergence of opinions is built up through the use of simple geometric shapes (circles or rectangles). During the subsequent iterations of the procedure, the shapes become smaller and smaller, until to circumscribe a very small portion of territory that is the final solution to the research/decision problem.

After the discussion of the methods and the presentation of some practical applications, we propose some possible evolutions that most likely will produce a future increase in the application of these techniques.

Chapter 2, titled "Interactive Scenarios," is written by Theodore J. Gordon and Jerome Glenn. A scenario is a rich and detailed portrait of a plausible route to a future world, including issues and decisions that might have to be faced as the future unfolds. In this chapter, we are most interested in scenarios creation processes that involve inputs from more than a single person, involve feedback, and may use quantitative models to help establish a foundation for the scenario story. Several interactive methods and applications are described including (1) use of Delphi questionnaires to collect suggestions for scenario content and scenario axes, that is the major dimension that defines the domain of interest; (2) redrafting a scenario story based on feedback from reviewers; (3) integrating a cross-impact matrix, Futures Wheel, or other modeling system to help assure internal self-consistency and quantitative rigor; and (4) allowing the audience to determine the course of the scenario at key decision points. Also described is a clustering approach in which a multitude of scenarios are constructed, each of which differs in input assumptions resulting from uncertainties associated with the variables and the assumed policies. The chapter concludes with a brief speculation about the future of interactive scenarios: multi-mode presentations so realistic that users feel they are in the scenario world and make simulated decisions accordingly.

Chapter 3 by Raymond R. Burke introduces “Virtual Reality for Marketing Research.” Computer graphic simulations of retail shopping environments have become a popular tool for conducting marketing research, allowing manufacturers and retailers to test innovative marketing concepts with shoppers in realistic, competitive contexts. Virtual store tests can deliver more detailed behavioral data than traditional methods of consumer research and are faster and less expensive than in-store field tests. This chapter outlines the benefits and limitations of virtual reality simulations, describes the steps involved in creating and running a simulated shopping study, discusses the validity of the simulation technique, provides examples of several commercial and academic research applications, and summarizes the future prospects for using the virtual store for marketing research and other business applications.

Chapter 4 by Pestana, Wang, and Moutinho is titled “The Knowledge Domain of Affective Computing: A Scientometric Review.” The aim of this study is to investigate the bibliographical information about Affective Computing identifying advances, trends, major papers, connections, and areas of research. A scientometric analysis was applied using CiteSpace, of 5078 references about Affective Computing imported from the Web-of-Science Core Collection, covering the period of 1991–2016. The most cited, creative, bursts, and central references are displayed by areas of research, using metrics and throughout-time visualization. Interpretation is limited to references retrieved from the Web-of-Science Core Collection in the fields of management, psychology, and marketing. Nevertheless, the richness of bibliographical data obtained largely compensates this limitation. The study provides managers with a sound body of knowledge on Affective Computing, with which they can capture general public emotion in respect of their products and services, and on which they can base their marketing intelligence gathering and strategic planning. The chapter provides new opportunities for companies to enhance their capabilities in terms of customer relationships. The effect of emotions on brand recall mediated by gender using voice emotion response with optimal data analysis.

Chapter 5 is written by the same previous authors—Wang, Pestana, and Moutinho—and is titled "The Effect of Emotions on Brand Recall by Gender Using Voice Emotion Response with Optimal Data Analysis." Its purpose is to analyze the effect of emotions obtained by oral reproduction of advertising slogans established via Voice Emotion Response software on brand recall by gender, and to show the relevance for marketing communication of combining "Human-Computer-Interaction (HCI)" with "affective computing (AC)" as part of their mission. A qualitative data analysis did the review of the scientific literature retrieved from Web-of-Science Core Collection (WoSCC), using the CiteSpace's scientometric technique; the quantitative data analysis did the analysis of brand recall over a sample of Taiwanese participants by "optimal data analysis." Advertising effectiveness has a positive association with emotions; brand recall varies with gender; and "HCI" connected with "AC" is an emerging area of research. The selection of articles obtained depends on the terms used in WoSCC, and this study used only five emotions. Still the richness of the data gives some compensation. Marketers involved with brands need a body of knowledge on which to base their marketing communication intelligence gathering and strategic planning. The research provides exploratory research findings related to the use of automatic tools capable of mining emotions by gender in real time, which could enhance the feedback of customers toward their brands.

Chapter 6 by Jyrki Suomala provides an overview of "The Neuroscience Research Methods in Management." Enormous progress in understanding fundamental brain processes by using neuroscientific methods underlying management, marketing, and consumers' choice has been achieved. All thoughts and ideas of people are constituted by neural circuits. However, people have only limited conscious access to these neural circuits. As a result, an estimated 2 percent of thoughts are conscious, and the weakness of traditional research methods—like surveys and focus group interviews—is that they concentrate mainly for people's conscious part of mind. On the contrary, the main benefit by using neuroscientific methods is that there are many possibilities to uncover the unconscious brain processes, which are critical for human choice in management contexts. The chapter divides neuroscientific methods in biometrics and neuroimaging. The main biometric methods include eye tracking, face reading, skin conductance, and heart rate measurements. The main neuroimaging methods include EEG and fMRI. The description of each method is presented with examples in the management and marketing contexts. The analysis of each method includes benefits and drawbacks in these contexts. And finally, how much a method can predict the human behavior in the real context. This chapter introduces neuroscientific methods at an elementary level to the management science community. It gives basic concepts and ideas on the application of neuroscience in order to solve scientific and practical management problems through the use of specific neuroscientific methodologies.

Jyrki Suomala also introduces the topic on the "Benefits of Neuromarketing in the Product/Service Innovation Process and Creative Marketing Campaign" (Chap. 7). Most of the neuroscientific studies try to find a neural correlation between specific stimuli and brain circuits' activation patterns. However, the brain-as-prediction approach tries to find brain circuits' activation patterns during persuasive messages, which predict human behavior in future situations. The chapter describes the brain valuation circuits, whose activation pattern has been found critical for the prediction of human behavior in the future. In addition, the most common examples about the brain's valuation system will be presented from management's point of view. Whereas the academic community has produced a lot of research with neuroscientific tools, there is still a lot of room for applications of neuroscience to the innovation and marketing campaign processes. Thus, the chapter describes how different stakeholders can cooperate in order to find optimal products and optimal marketing campaign by using neuromarketing. Because the innovation process includes much uncertainty, it is very critical to encourage all participants to collaborate in the process. The chapter describes the benefits of this collaboration from management's point of view. In addition, the concrete examples of how researchers and business innovators together can apply neuromarketing in order to solve concrete innovation and management problems are presented. Finally, the future's opportunities of neuromarketing in innovation processes are also presented.

Chapter 8 of Volume 2 and of the whole book itself is by Robin Chark. It also deals with the nascent theme of neuromarketing. This chapter surveys methodological advancements in consumer research as a result of recent developments in neuroscience and behavioral genetics. This new approach is highly interdisciplinary and is termed "neuromarketing," paralleling similar developments in related fields such as neuropsychology and neuroeconomics. Researchers of this approach consider our behaviors to be the result of psychological processes embodied physiologically in the brain and nervous system. Thus, the biological influences may be rooted in our genes and shape the activities in our brain, and thus behaviors through actions on hormones and neurotransmitters. In this paradigm, our genotypes may be interpreted as a measure of individual differences, while brain activities may be observed and taken as more direct measures of the underlying psychological process. Such a neurobiological approach may take consumer research to another level and help answer some historically difficult questions. These questions are not confined to marketing scholars' interest in theory development; they encompass real-world marketing implications of concern to practitioners. To illustrate this, we review some recent findings in neuromarketing that make use of neuroimaging, twin study, and molecular genetics. We then discuss some recent trends in neighboring fields and their implications for the future of neuromarketing

In Chap. 9, Pacinelli presents the Futures Polygon. This futures research method is based on subjective probabilities and evaluations. It is a very interesting approach to the study of the future with facets that include subjective impacts. With varied degrees of probabilities, the Futures Polygon is designed for the expansion of scenarios. The methodology of the Futures Wheel is also incorporated. Methods of integration and important illustrations are introduced in the chapter.

We both hope that you like the challenging innovative research methodologies in management and that they will assist researchers in the enhancement of their scholarly research capabilities!

Our last word is a word of thanks to Liz Barlow, our editor, and Lucy Kidwell from Palgrave for all their help and understanding. Moreover, we wish to thank Liz for believing in this publishing project. We are both very grateful. We would also like to thank all the excellent contributors to the two volumes. Their reputation, mind-stretching methodologies, and collaborative stance were really very impressive. We are very indebted to all of them.

Enjoy the content of the two volumes, and we sincerely hope to have the readers' curation of their content and co-creation of future value in this domain that could help us prepare new writings……

May, 2017

Luiz Moutinho
Mladen Sokele

References

Nissen, V. (1995). An Overview of Evolutionary Algorithms in Management Applications. In J. Biethahn & V. Nissen (Eds.), *Evolutionary Algorithms in Management Applications* (pp. 44–97). Berlin: Springer.

Rothlauf, F. (2011). *Design of Modern Heuristics. Principles and Application*. Berlin: Springer.

Sörensen, K. (2015). Metaheuristics—The Metaphor Exposed. *International Transaction in Operational Research, 22*, 3–18.

Sörensen, K., Sevaux, M., & Glover, F. (2016). A History of Metaheuristics. In R. Marti, P. Pardalos, & M. Resende (Eds.), *Handbook of Heuristics*. Berlin: Springer. (to appear).

Weyland, D. (2015). A Critical Analysis of the Harmony Search Algorithm—How Not to Solve Sudoku. *Operations Research Perspectives, 2*, 97–105.

Wolpert, D. H., & Macready, W. G. (1997). No Free Lunch Theorems for Optimisation. *IEEE Transactions on Evolutionary Computation, 1*, 67–82.

Contents

Notes on Contributors

Raymond R. Burke is the E.W. Kelley Professor at Kelley School of Business, Indiana University, and founding director of the School's Customer Interface Laboratory. He has also served on the faculties of the Harvard Business School and Wharton. Burke's research focuses on understanding the influence of point-of-purchase factors on shopper behaviour. His articles have appeared in the *Harvard Business Review*, *Journal of Consumer Research*, *Journal of Marketing*, *International Journal of Research in Marketing*, and *Marketing Science*, among other journals, and he has been awarded several patents for his innovations in marketing research methodology and digitally enhanced shopping experiences.

Robin Chark is Assistant Professor in the Faculty of Business Administration at the University of Macau. Before receiving his PhD degree in marketing, he earned an MSc degree in economics and a BBA degree in finance, all from the Hong Kong University of Science and Technology. Chark's research interest lies in neuromarketing, consumer behaviour, behavioural and experimental economics, and neuroeconomics.

Simone Di Zio is a researcher in Social Statistics at the University "G. d'Annunzio", Chieti-Pescara (Italy). He is the co-chair of the Italian Node of the Millennium Project and member of the Italian Statistical Society. He has more than 40 publications, including articles and monographs. His scientific activities deal with Spatial Statistics, IRT modelling and methods for the convergence of opinions, with applications on terrorism, crime, energy, pollution and tourism. Recently he invented new methods, such as the Spatial Delphi, the Spatial Shang and the Real Time Spatial Delphi. He is an expert in geographic information systems (GIS) and fractal geometry.

Jerome C. Glenn is the CEO of the Millennium Project, a leading global participatory think tank, which produces the *State of the Future* reports for the past 20 years. He invented the "Futures Wheel", a futures assessment technique and concepts such as conscious-technology, transinstitutions, tele-nations, management by understanding, feminine brain drain, "Nodes" as a management concept for interconnecting global and local views and actions, and definitions of environmental security and collective intelligence. He wrote about information warfare in the late 1980s in *Future Mind*, sent his first email in 1973 and conduced 55 futures research studies.

Theodore J. Gordon is a futurist and management consultant, a specialist in forecasting methodology, planning and policy analysis. He is co-founder and Board member of the Millennium Project, a global think tank. He formed the consulting firm the Futures Group in 1971 and led it for 20 years. Prior to forming the Futures Group, he was one of the founders of The Institute for the Future and before that consulted for the math and policy department at RAND Corporation. He was also Chief Engineer of the McDonnell Douglas Saturn S-IV space vehicle and was in charge of the launch of early ballistic missiles and space vehicles from Cape Canaveral. He is a frequent lecturer, the author of many technical papers and author or co-author of several books dealing with space, the future, life extension, terrorism and scientific and technological developments and issues. He is the author of the *Macmillan Encyclopedia* article on the future of science and technology. He is currently on the editorial board of *Technological Forecasting* and *Social Change.*

Luiz Abel Moutinho is Full Professor of BioMarketing and Futures Research at DCU Business School, Dublin City University, Ireland. Moutinho's current research interests are mainly bio-marketing, neuroscience in marketing, evolutionary algorithms, human–computer interaction, the use of artificial neural networks in marketing, modelling consumer behaviour, Futurecast marketing, and tourism and marketing. He has written over 140 refereed academic journal articles and 29 books.

Antonio Pacinelli is a graduate in Statistical and Actuarial Sciences, University of Rome "La Sapienza". He is a contract professor of Mathematical Methods applied to Biological and Biomatematical Sciences at the University of Teramo, 1981–1988. Researcher of Statistics (1988–1998), Associated Professor of Statistics (1998–2004), and Full Professor of Social Statistics at the University

G. d'Annunzio, Pescara. He has 100 publications (volumes and articles). He has been one of the five members of Actuators Committee of the Faculty of Managerial Sciences, University G. d'Annunzio. His methodological scientific activities deal with Markovian analysis, space and space-time analysis and on the building of scenarios.

Maria Helena Pestana is Assistant Professor of Statistics and Data Analysis at ISCTE-IUL Business School, Lisbon, Portugal. Pestana's current research interests are mainly tourism, marketing, social sciences, computer interaction tools and Neuroscience.

Jyrki Suomala received a PhD degree in education from the University of Jyväskylä, Finland, 1999), and he holds positions as Principal Lecturer at the Laurea University of Applied Sciences and Adjunct Professor in the University of Oulu. He is the head of the NeuroLab at the Laurea. He has been a visiting researcher at the University of California, Santa Barbara, for a total period of 3.5 years. Suomala applies cognitive science, behavioral economics, and neuroeconomics perspectives, when he studies human behaviour in the different circumstances. He has participated as researcher and principal investigator in various research projects funded by the Academy of Finland, TEKES and Helsingin Sanomat Foundation.

Wan-Chen Wang completed her PhD at the University of Glasgow in 2010. She is an associate professor at the Department of Marketing, Feng Chia University, Taiwan. Her main research interest focuses on emotion research methods in marketing, emotion and consumer behaviour. She has written articles in *Marketing Letters* and other refereed international journals.

List of Figures

List of Tables

1

Convergence of Experts' Opinions on the Territory: The Spatial Delphi and the Spatial Shang

Simone Di Zio

Introduction

In forecasting and/or decision-making contexts, when quantitative data are insufficient or totally absent, subjective judgments are of extreme usefulness. Generally, no one individual has the sufficient expertise and knowledge to make the best forecast or to take the best decision, thus all along organizations have sought to gather the opinions of groups of individuals in an attempt to combine their skills and improve decision-making (Riggs 1983).

However, having a team of experts is not enough, because the way their views are collected is crucial, and without a rigorous methodology, any consultation process may become vain. Traditional methods of grouping experts, such as focus groups or face-to-face interviews, are very popular but have quite important drawbacks. In what we call here "interacting groups", compromise decisions are often reached, rather than

S. Di Zio (✉)
Department of Legal and Social Sciences, University "G. d'Annunzio", Chieti-Pescara, Chieti, CH, Italy

L. Moutinho, M. Sokele (eds.), *Innovative Research Methodologies in Management*,
DOI 10.1007/978-3-319-64400-4_1

consensus decisions, and the distortive factors of the interacting groups are widely discussed in the scientific literature (Van de Ven 1974; Riggs 1983; Di Zio and Staniscia 2014b; Grime and Wright 2016).

Following is a summary of such distortions:

- *The effect of the leadership*: When the highest-ranking of a hierarchy (e.g. military, political or academic) or a particularly dominant person expresses an opinion, the others usually tend to follow that judgment, in spite of contrary feelings. Thus, not everybody expresses thoughts freely, for the fear of coming into conflict with the leader. In other words, dominant personalities influence the group, and low-status members tend to go along with the opinions of high-status members (Torrance 1957; Chung and Ferris 1971).
- *The spiral of silence*: Those who agree with the ideas of the majority are more likely to feel confident in expressing their opinions, while those who are in the minority fear that manifesting their views will result in social ostracism; therefore, they are inclined to be silent. These perceptions can lead to a spiraling process, in which minority's views are increasingly restrained and, as a consequence, under-represented (Neill 2009).
- *The groupthink factor*: This distortion occurs when the pressure to conform within the group interferes with the group's analysis of the problem, to the point of producing poor decisions (Hoffman 1965). When the members of an interacting group strive for reaching a broad consensus, their motivation to assess alternative courses of action is affected, and the independent thinking is lost, in the pursuit of group cohesiveness (Hassan 2013). The expression "groupthink" indicates the situation in which, when searching consensus, in order to minimize the conflicts, the individuals renounce their ideas and opinions.

There are several ways to manage the previous distortions, for example, by avoiding face-to-face contacts and structuring the interactions anonymously. The aim is to prevent the association of the opinions to those who have expressed them, avoiding the errors arising out of the effect of the leadership and the spiral of silence. Moreover, instead of collecting the expert's judgments at one time, in the same place and within a limited time, one can structure the consultation in an iterative framework and asynchronously (i.e. at different times). By collecting the evaluations iteratively, the participants can review at least once their assessments, perhaps with the possibility of comparing them with the answers provided by the other experts of the same group (*controlled feedback*). The possibility for the members to interact at a distance and at different times eliminates the pressure to decide quickly, within a given time limit, so avoiding the groupthink bias.

The *anonymity*, the *iterative structure* and the *asynchronous communication* are the most used strategies for the elimination of the distortive effects typical of the interacting groups. All these features are present in the Delphi method, a very popular iterative technique for collecting expert's opinions, conceived to achieve consensus on a particular issue. Dalkey and his associates developed the Delphi at the RAND Corporation, a research institute founded in 1946 with the financial support of the US Department of Defense (Dalkey and Helmer 1963). Since its invention, the applications of the method were numerous, and over the past 60 years, many other methods related to it have been developed.

In this chapter, we briefly explain the origins and the evolution of the Delphi method, up to two recent variants, the *Spatial Delphi* and the *Spatial Shang*, specially designed to treat problems related to the territory. With these two methods, the judgments of the experts are collected by means of points placed on a map, and the process of the convergence of opinions is built up with simple geometric figures (circles, rectangles or strips). During the iterations, the figures on the map move and become smaller and smaller, until to circumscribe a small portion of territory that represents the final solution to the research/decision problem. After the description of the methods and the presentation of the early applications, we discuss some possible evolutions that most likely will produce a future increase in the use of these techniques. In particular, we will talk about the *Real Time Spatial Delphi*, a real-time version of the Spatial Delphi.

The Delphi Method

The Delphi method is a technique that uses responses (typically opinions) to a questionnaire by a group of experts (*panel*) or social actors to solve a problem, generally in a decision-making context and/or a forecasting framework. It consists of a number of iterations, called *rounds*, during which the administrator of the process (*facilitator*) provides statistical summaries of the answers given by the members of the panel. The experts communicate with each other anonymously, at distance and asynchronously, and, as seen above, these features resolve several problems typical of other methods of group decision-making.

This method was born in the 1950s, but only later (in the 1960s) it took the name *Delphi*, by the RAND Corporation. The first application, commissioned by the Government of the United States of America, concerned the use of expert's opinions for selecting an American industrial target and estimating the number of atomic bombs needed to reduce the American arsenal by a predetermined amount, all seen from the point of view of a Soviet strategist.

At that time, there were also alternative methodologies, but they consisted of complex mathematical models, requiring a lot of data, very long procedures and expensive computers. Curiously, in the following years, those sophisticated models were applied, but the surprise was that the results of the Delphi were better. This explains why, after more than half a century, the method remains valid in situations in which it is difficult to get quantitative data, and/or for long-range forecasting, when expert's opinions are the only source of information.

Due to the secrecy of the first study, the method was disclosed only after several years. Two papers are recognized as the basic literature of the Delphi, one by Norman Dalkey and Olaf Helmer, which describes the methodological and philosophical foundations of the method (Dalkey and Helmer 1963), and the other by Theodore J. Gordon and Olaf Helmer, entitled "Report on Long-Range Forecasting Study" (Gordon and Helmer 1964).

From that time, the Delphi has been used continuously in various substantive fields, often together with other methodologies and with a very high number of applications (Brockhaus and Mickelsen 1975).

Now, let us see how the Delphi works. The first step is the formulation of the topics, generally derived from the literature, from the available surveys or defined by a working group created specifically. Immediately after, you have to build the panel of experts (Gordon 2009a). Panelists are knowledgeable persons, generally identified through literature searches or recommendations from institutions or other experts, and must be selected based on their *expertise*. Nevertheless, the group should be heterogeneous, in order to ensure an adequate variability in the assessments and stimulate an exchange of views and knowledge (Rowe and Wright 2001). The size of the panel depends on many factors, such as the number of the topics, the fields of expertise, the expected response/acceptance rate and other issues, but considers that most applications use panels of 15–40 people. Each expert should be contacted individually, preferably by telephone, and informed about the study, the objectives of the research and the number of rounds.

The next step is to build and test the questionnaire, to find any flaws, and then it is sent to the participants, for example, by e-mail (Gordon 2009a). After collecting the answers, the facilitator summarizes the opinions expressed by the experts by means of appropriate statistical indices (Glenn 2009). The classic approach provides the calculation, for each question, of the median and/or the first and third quartiles of the distribution of the responses, resulting in an interval (called *quartile range*) that contains 50% of the estimates. This interval is the embryo of the consensus and constitutes a window of response for the second round of consultation.

The second round sees the administration of the same questionnaire, enriched with the synthetic results of the first round (e.g. the quartile range), thereby triggering the feedback process. The questions are the same of the first questionnaire, but each expert is asked to answer inside the quartile range (invited but not obligated). In this way, if some expert revises his/her previous assessment to stay inside the proposed range, already from the second consultation the process of "convergence of opinions" begins. The panelists are also invited to give written reasons, especially if they give evaluations outside the proposed range.

The results of the second round are processed as before, and then in the third consultation, the panelists are asked to answer the questions trying to stay within the new quartile ranges. Now, the facilitator circulates anonymously the reasons provided in the previous round, and counter-arguments can be provided, so enriching the debate. Proceeding iteratively, the quartile range of each question should narrow more and more, until a value small enough such that the consensus is sufficient. When a predefined stopping point is reached (for instance a preset number of rounds, a consensus threshold or the stability of answers), the facilitator concludes the consultation and proceeds to the final elaborations (Grime and Wright 2016). Finally, the whole procedure ends with the presentation and comment of the results.

Although consensus can be used as a stopping criterion, unfortunately there is no guarantee that it is reached and, above all, there is a difference between "consensus" and "stability". In a Delphi study, it is important to check first the stability of the evaluations and only after verify whether there is consensus (von der Gracht 2012). There is stability when the results of two subsequent rounds are not significantly different, and it should be used as stopping criterion. A possible measure of stability is the percent change from round to round, and a 15% change (or lower) is considered a stable situation (von der Gracht 2012; Scheibe et al. 1975; Dajani et al. 1979). If stability is achieved, there may be consensus, but also other layouts of responses, like for example majority or plurality of views (Dajani et al. 1979). Nevertheless, the lack of consensus does not mean the failure of the Delphi exercise, because "the absence of consensus is, from the perspective of data interpretation, as important as the existence of it" (von der Gracht 2012).

A Delphi Variant: The Shang Method

The Delphi has been so widely used to the point that is considered the father of many other methods. To give some examples, in 1970 Murray Turoff proposed the *Policy Delphi* (Turoff 1970), a consensus-oriented method used for the analysis of public policies, while a different version, called *Public Delphi*, is based on the participation of the citizens. After a few years the *Mini Delphi* was proposed, also known as Estimate-Talk-Estimate (ETE), a technique that speeds up the procedure, as it is applied for face-to-face meetings (Gustafson et al. 1973; Van de Ven and Delbecq 1974). In 1974, De Groot (1974) laid the theoretical fundamentals of the *Markov-Delphi,* and then Chatterjee (1975) studied an alternative solution, based on variable weights.

In the same year, David A. Ford proposed the *Shang* method (Ford 1975), an interesting technique in which some characteristics of the classical Delphi are kept, but the trouble of asking to rephrase the evaluations at each round is eliminated.

Like in the conventional Delphi, the first phase of the Shang concerns the definition of the topics and the construction of the panel. In the first questionnaire, the experts are asked to express a minimum and a maximum of the value to be estimated, so the first round produces two distributions, one for the minimum and one for the maximum. Denoting with n the number of the responses, you get a vector $\boldsymbol{m}$ with n values for the minimum, and a vector $\boldsymbol{M}$ with n values for the maximum. For each vector, the facilitator calculates a statistical synthesis (e.g. the median), and these two values (here denoted by m_0 and M_0) define an initial evaluation interval: $[m_0, M_0]$. Then, the central value of this interval is calculated:

$$C_0 = \frac{M_0 + m_0}{2}$$

This number represents the basis of the second questionnaire where, differently from the Delphi, each expert is invited to compare the central value C_0 with what he/she believes more plausible, simply answering "greater than" or "less than".

If the majority of the estimates are greater than C_0, the new interval is $[C_0, M_0]$, while if the majority are less than C_0, the new interval is $[m_0, C_0]$. Accordingly, a new central value is calculated, say C_1, as the mean of the extremes of the new interval. The method proceeds iteratively, calculating subsequent intervals with the respective central values $(\ldots, C_2, C_3, \ldots, C_k)$, until a stopping point is reached.

The first advantage of the Shang is that the speed of convergence is greater than in the Delphi, given that is known a priori that at each iteration the interval is halved, while in the Delphi the width of the interval of consensus depends on the distribution of the answers. Second, the Shang does not push the experts to change their estimates at each round, asking to give assessments within an interval. In fact, if an expert has in mind a value, during the subsequent rounds he/she is not induced to modify it, because it is only asked if that value is below or above the proposed central value. Therefore, while the iterations of the Shang proceed and the central value changes, the estimate that an expert initially thought can remain the same until the end.

From Delphi to Spatial Delphi: Looking for Geo-consensus

Many mental operations and the ensuing decisions involve a spatial reasoning, like driving a car, searching for an address or looking for a hotel on the Internet. With the expression *spatial reasoning,* we intend a way of thinking which implies the use of a map, explicit if it is a paper or digital map, or conceptual if our brain uses a map at a cognitive level. The decisions that ensue can be defined *spatial decisions*. Starting from these considerations, in recent years some authors are developing a new line of research, in which the logic of the Delphi method is integrated in the context of spatial decisions.

Dragicevic and Balram (2004) defined the concept of "Collaborative Spatial Delphi" which makes use of the Geographic Information Systems (GIS) technology, to support the convergence of opinions obtained with the Delphi method. During the consultation, the experts draw polygons and write comments on a digital map; nevertheless, the convergence of opinions is still reached using the classical Delphi technique.

A bit later, Jankowski (Jankowski et al. 2008) proposed a web-based spatial multiple criteria evaluation tool for individual and group use, which supports sketches created on a digital map, documents and the construction of statistical indicators. The system includes "a vote aggregation function to collate individual option rankings into a group ranking, and measures of agreement/disagreement to inform the participants about a group-derived desirability of specific decision options" (Jankowski et al. 2008). Nevertheless, the Delphi does not yet have a spatial form.

After a few years, Di Zio and Pacinelli (2011) expressly spoke of "Spatial Delphi", a technique in which the logic of the Delphi method is fully integrated within a map. The basic idea is the narrowing of the opinions of the experts on the territory, in order to find a location in the area of interest—or at least a narrow region—on which there is consensus among the panelists.

As discussed by the inventors of the method, there are three categories of problems where the Spatial Delphi can be applied:

1. **The present**: choose the optimal geographical location to place goods, services or buildings, when quantitative data are not available.
2. **The future**: if a future event has a given probability to happen, it is important to predict *where* it will be most likely to happen (e.g. catastrophic events such as hurricanes, earthquakes, floods, fires and so on).
3. **The underground**: the search for non-visible underground elements is nowadays supported by technology, but the human experience still plays an important role. The Spatial Delphi can be used for the consultation of experts in order to find things that are not visible (archeology, geology, oceanography).

However, how does the Spatial Delphi work? The authors have retraced all the steps of the conventional Delphi, adapting them to a spatial context (Di Zio and Pacinelli 2011). First, the localization problem must be defined, as detailed as possible. The panel of experts must be built according to the principles of expertise and heterogeneity, but an additional requirement is necessary: each expert must know well the geography of the area under study. Then, the questionnaire is implemented on a paper or a digital map; here the GIS technology is essential for setting the map(s) and any supplementary materials.

After the preparatory stage, the steps of the iterative phase of the Spatial Delphi start.

Step 1 Ask the experts to locate k_1 points (the subscript denotes the first round) on the map, representing the most suitable sites to locate goods, services or buildings, or the places where a future event will likely occur, or the most appropriate places to dig for underground materials, etcetera. The result of the first round is a map with a cloud of n_1 geo-referenced points, called *opinion-points*. If E is the number of experts in the panel, then $n_1 = E \cdot k_1$.

Step 2 In analogy with the quartile range of the conventional Delphi, a circle containing 50% of the n_1 opinion-points is constructed, called *circle of consensus* or *circle of convergence*. Excluding different properties in different directions, the circle is the natural isotropic shape in a two-dimensional space.

The problem is that with n_1 points distributed on a plane, the number of circles containing $n_1/2$ points is infinite, so constraints are necessary. The first constraint is that the circle must be centered on one of the opinion-points. In this way, it is guaranteed that the center of the circle of consensus coincides with one of the locations expressed by the panelists and so the number of circles is restricted to n_1. Second, given that the goal is the convergence of opinions around an area as small as possible, among the n_1 possible circles the algorithm requires the choice of the smallest one. Thus, there is only one circle with a minimum radius, centered on one of the n_1 points, containing $n_1/2$ opinion-points.

Let us denote with cc_i the smallest circle centered on P_i (the generic *i-th* opinion-point, $i = 1, 2, \ldots, n_1$), containing $n_1/2$ points if n_1 is even, or $(n_1 + 1)/2$ if n_1 is odd. With A_i and r_i respectively the area and the radius of cc_i. In the first round, there are n_1 of such circles and then a vector of areas $\boldsymbol{A} = \left(A_1, A_2, \ldots, A_{n_1}\right)$. From this vector you need to determine the smallest value, $A^1_{best} = \arg.\min\left(\boldsymbol{A}\right)$, which represents the first *circle of convergence*, denoted by CC^1 (the superscript indicates the first round). The correspondent opinion-point is denoted by $\boldsymbol{P}^1_{best}$. From a technical point of view, a distance matrix $\boldsymbol{D}$ between all the n_1 points is calculated (size $n_1 \times n_1$). The *i-th* row of $\boldsymbol{D}$ is a vector (say $D_{i\cdot}$) containing the distances between the *i-th* opinion-point and all the other opinion-points. The median of this vector is the radius of cc_i ($r_i = median(D_{i\cdot})$), and consequently the area of the *i-th* circle is calculated as $A_i = \pi r_i^2$. Repeating this calculation for each row of the matrix $\boldsymbol{D}$, it results the vector $\boldsymbol{A}$.

The facilitator draws on the map the first circle of convergence CC^1, with area A^1_{best}, centered on the opinion-point $\boldsymbol{P}^1_{best}$, and this constitutes the base of the questionnaire of the second round.

Step 3 The map with the circle CC^1 is circulated among the panelists for the second round. The experts re-evaluate their opinion-points with the invitation of remaining inside the circle. According to the original version of the method (Di Zio and Pacinelli 2011), each expert must provide $k_2 = k_1 - 1$ points, that is one opinion-point less in respect to the first round, but the number of points can also be kept constant throughout all the rounds. In analogy to the conventional Delphi, if an expert wants to place a point outside the circle, he/she can do so, but is invited to give a motivation. After the second round, there are $n_2 = \mathrm{E} \cdot k_2$ opinion-points drawn on the map (if there are not dropouts). Di Zio and Pacinelli (2011) explain as "reducing the number of available points may improve respondents' psychological state because revising their evaluations panelists can eliminate the points outside the circle and keep those included in the circle. [...] the possibility of maintaining a part of the previous evaluation and eliminating or moving another part increases the degrees of freedom and should produce a better tendency to revise preceding choices."

Step 4 After the second consultation, a new circle of consensus CC^2 is calculated, following the same algorithm, with center P_{best}^2 and area A_{best}^2, containing 50% of the n_2 opinion-points. The new circle is depicted on the map proposed to the panelists for the third consultation, and the procedure continues iteratively.

After a number of rounds, say s, there will be a circle CC^s representing the territory where the spatial convergence of the opinions is achieved (*geo-consensus*). From a geometrical point of view, if the procedure generates a consensus, the consecutive circles are smaller and smaller, namely $A_{best}^1 \geq A_{best}^2 \geq \ldots \geq A_{best}^s$.

Table 1.1 compares the main steps of the classical Delphi with those of the Spatial Delphi and helps to highlight similarities and differences.

When the stopping point is reached, the final circle represents the "geographical" result of the survey. However, there are also "non-geographical" results, like the comments of the experts and some measures of geo-consensus.

If you know the limits of the study region, a simple measure of geo-consensus derives from the ratio between the area of the final circle (A_{best}^s) and the surface of the study area:

Table 1.1 Delphi and Spatial Delphi in comparison

Delphi	Spatial Delphi
1. Formulation of the topics, selection of the panelists and construction of the first questionnaire	1. Formulation of the topics, selection of the panelists and construction of the map with the spatial question(s)
2. Administration of the first questionnaire	2. Administration of the first questionnaire. The panelists answer by locating k_1 opinion-points on the map
3. Calculation of the first quartile range (50% of the evaluations)	3. Calculation of the circle CC^1 (50% of the k_1 opinion-points)
4. In the second questionnaire, the panelists are asked to give assessments inside the quartile range. External evaluations should be argued	4. In the second questionnaire, the panelists are asked to locate k_2 points inside the circle CC^1. External points should be argued
5. Calculation of the second quartile range, which is shown in the third questionnaire, together with the reasonings	5. Calculation of the circle CC^2, which is shown on the third map, together with the reasonings
6. Administration of the third questionnaire. In case of evaluations external to the quartile range, the experts are invited to argue the choice	6. Administration of the third questionnaire. Panelists are asked to locate k_3 points inside CC^2. External points should be argued
7. Iterate, until the stopping point (stability, consensus)	7. Iterate, until the stopping point (stability, consensus)

$$MC_1 = 1 - \frac{A_{best}^{s}}{M}$$

where M is the surface area of the region. The closer MC_1 comes to one, the smaller the final circle is, compared to the study area, therefore indicating a high degree of geo-consensus (Di Zio et al. 2016).

Another relative measure of convergence, especially useful when the study area is not bounded, is the percentage ratio between the final circle and the initial one:

$$MC_2 = \frac{A_{best}^{s}}{A_{best}^{1}} \cdot 100$$

This quantity varies between 0 and 100, and a value close to 0 denotes a high degree of geo-consensus, because the final circle is small compared to the initial one (Di Zio et al. 2016).

In addition to the measures that consider the circles, there are also some indicators of geo-consensus based on the entire cloud of points, generated at each round. The *K function* is a measure of the spatial dependence between point-events as a function of the distance (Ripley 1976; Bailey and Gatrell 1995). By denoting with R the region of interest, with n the number of opinion-points and with d_{ij} the Euclidean distance between points i and j, the following expression provides an estimator for K (Bailey and Gatrell 1995):

$$\widehat{K}(h) = \frac{R}{n^2} \sum \sum_{i \neq j} I_h(d_{ij})$$

Here h is the distance and $I_h(d_{ij})$ is an indicator function with value 1 if $d_{ij} \leq h$ and 0 otherwise. If the n points are scattered, we have $\widehat{K}(h) < \pi h^2$, while for clustered points $\widehat{K}(h) > \pi h^2$. To interpret $\widehat{K}(h)$, the graphical representation of the following transformation shall be used:

$$\hat{L}(h) = \sqrt{\frac{\widehat{K}(h)}{\pi}} - h$$

If $\hat{L}(h) > 0$, the opinion-points are clustered; therefore, this function is an indicator of the geo-consensus. On the contrary, negative values of $\hat{L}(h)$ occur when the points are scattered, a situation where the panelists have not reached a convergence of opinions on the space. A plot of $\hat{L}(h)$ versus h for each round of the Spatial Delphi helps in evaluating the degree of geo-consensus. In Fig. 1.1, there is an example of $\hat{L}(h)$ estimated on a simulation with three rounds, where it is evident that the degree of spatial consensus increases as the rounds proceed.

Di Zio and Pacinelli (2011) proposed also another indicator of the geo-consensus that is the fractal dimension (FD). The FD estimated on a cloud of points measures how the points "cover" the surface. It varies

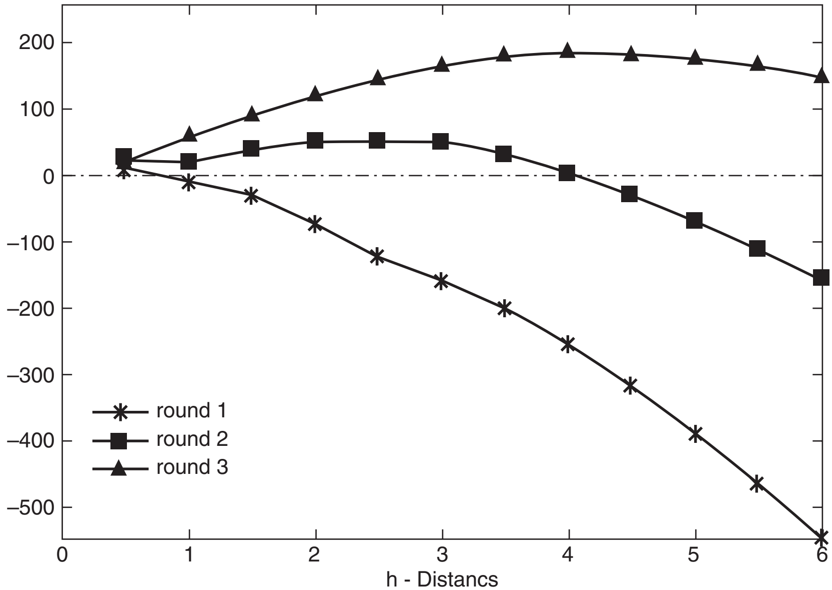

Fig. 1.1 Example of *L* functions

between zero (clustered points, high geo-consensus) and one (scattered points, low geo-consensus). For the estimation of the FD, we use the *correlation dimension* method (Hastings and Sugihara 1993). If from the first to the last round of the Spatial Delphi the FD decreases, one has an indicator of the spatial convergence of opinions.

The indices MC_1, MC_2 and FD can also be used to measure the speed of convergence and to check for stability. For example, from one round to the next, a change in the FD less than 15% can be considered an indicator of stability.

The Spatial Shang

The *Spatial Shang* is a variant of the Shang method and arises from the same considerations which have led to the Spatial Delphi. Like the Spatial Delphi, it is applicable whenever the research problem concerns the choice of a geographical location (Di Zio and Staniscia 2014a).

First, the location problem must be defined carefully, and the panel of experts (say n participants) must be built according to the principles of the expertise, the heterogeneity and the knowledge of the region under study. Like for the Spatial Delphi, the GIS technology is an excellent tool for the preparation of the maps, and any supplementary materials.

Then, the steps of the iterative phase are the following:

Step 1 The experts are invited to draw four points on the map, to delimit the area where, for example, it is assumed to occur in a future event. Two points are the extreme positions along the North-South direction and two are the extreme positions along the East-West direction (Di Zio and Staniscia 2014a). For the generic expert i^{th} a point (say N_i) is the north limit beyond which the expert considers that the event under study will never occur. On the opposite side, a second point expresses the south limit (S_i). In the direction of longitude, the other two points represent the leftmost limit (W_i) and the rightmost limit (E_i). Thus, these four points identify four imaginary lines that surround the area with a rectangle, which represents the initial solution of the spatial problem for the i^{th} expert. Of course, there will be n rectangles, one for each expert.

Step 2 The result of the first consultation consists of four vectors of n values. One vector, say $\boldsymbol{N}$, containing the n evaluations for the north limits, one for the south limits ($\boldsymbol{S}$), one for the east limits ($\boldsymbol{E}$) and one for the west limits ($\boldsymbol{W}$). These vectors generate n different rectangles on the map and, in analogy with the classical Shang, in order to have a unique solution, we have to compute a statistical synthesis for each vector (e.g. the medians or the arithmetic means). We denote these indices with N_0, S_0, E_0 and W_0 (see Fig. 1.2). Of course, N_0 and S_0 are values of latitude, while E_0 and W_0 are values of longitude.

These four values define the first rectangle of convergence (*ABCD* in Fig. 1.2), with the vertices having the following coordinates: $A(W_0, N_0)$, $B(E_0, N_0)$, $C(E_0, S_0)$ and $D(W_0, S_0)$. This rectangle, denoted by R^0, represents the average rectangle and is the analogous of the initial interval of the conventional Shang. The area of R^0 is $A_0 = (|E_0 - W_0|) \cdot (|N_0 - S_0|)$.

By using the four indices (N_0, S_0, E_0, W_0), two central values are calculated, one for the latitude: $C_{0,LAT} = (N_0 + S_0)/2$, and one for the longitude: $C_{0,LONG} = (E_0 + W_0)/2$. These are two geographical coordinates identifying the center of gravity of R^0, namely $\boldsymbol{G}_0(C_{0,LONG}, C_{0,LAT})$. Two orthogonal lines passing through $\boldsymbol{G}_0$ divide the rectangle R^0 in four sub-rectangles, each having area A_1 equal to a quarter of A_0, namely $A_1 = A_0/4$. Starting from the north-east and proceeding clockwise, we denote these sub-rectangles with *NE*, *SE*, *SW* and *NW* (see Fig. 1.2).

Step 3 The map with the rectangle R^0 and the four sub-rectangles is submitted to the panel for the second round. Each expert is asked to locate an opinion-point somewhere in one of the sub-rectangles. In other words, it is like asking the expert which sub-rectangle he/she considers the most appropriate for the solution of the spatial problem. Like in the Spatial Delphi, the answer is very easy and fast to be given.

Step 4 The sub-rectangle that received the highest number of points becomes the new rectangle of convergence (R^1) which will be divided, in its turn, in four sub-rectangles on the basis of a new center of gravity $\boldsymbol{G}_1$. The area of each new sub-rectangle (A_2) has a 16th of the area of the initial rectangle R^0, namely $A_2 = A_0/16$. Let us take an example. If the sub-rectangle *SW* receives the majority of the opinion-points, the center of gravity $\boldsymbol{G}_1$ has the following coordinates: $C_{1,LAT} = (S_0 + C_{0,LAT})/2$ and $C_{1,LONG} = (W_0 + C_{0,LONG})/2$ (see Fig. 1.2).

In case two or more sub-rectangles receive the same higher number of opinion-points, the choice will fall on the rectangle that contains the

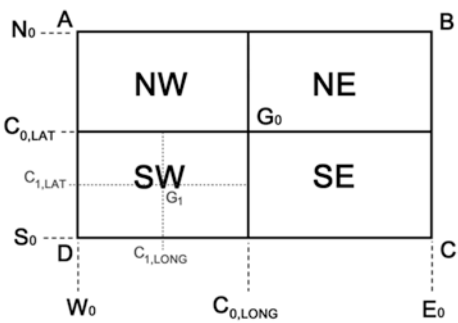

Fig. 1.2 Schematic representation of the Spatial Shang

farthest points from the center of gravity. This is the advantage of asking to locate the opinion-points and not simply the choice of one of the four sub-rectangles (Di Zio and Staniscia 2014a).

Step 5 Further rounds are performed, selecting new sub-rectangles and new centers of gravity, until a stopping point is reached, for example, after a number of rounds or when a sufficiently small portion of territory is delimited and the spatial consensus can be considered reached. For example, you can decide that the final rectangle must be smaller than a certain fraction p of the initial rectangle $(0 < p < 1)$, so the stopping criterion is $A_k \leq pA_0$.

If k is a generic round, at each iteration the area of each sub-rectangle (A_k) is reduced by a factor of $1/4^k$,namely $A_k = A_0/4^k$, and this means that the process of convergence is very fast (one of the peculiarities of the Shang method).

Benefits and Limitations of the Spatial Delphi

Even though many decision/forecast problems concern the territory, until today in almost all the Delphi applications the geographic element has been greatly overlooked. The Spatial Delphi and the Spatial Shang fill this gap. We consider these methods complementary, and not alternative, to the conventional Delphi. Our suggestion is that when in a research problem there are issues related to the territory, a conventional Delphi questionnaire can be supplemented with a number of questions of a Spatial Delphi/Spatial Shang.

Given the above, we now turn to a brief description of the advantages and limits of the Spatial Delphi that, in principle, also apply to the Spatial Shang.

In the design phase, the typical problems of the conventional Delphi about the choice of the response scales (dichotomous scales, rating scales, etc.) and about the number of response categories (three-point scales, five-point scales, etc.) are not present in the Spatial Delphi, whatever the spatial issue. The positioning of a point on a map is quick and intuitive, and does not force the participants to complex reasoning on the questions, as occurs in any conventional questionnaire. The method is easily accessible and understandable, even for a non-specialized audience, given that we are used to reading maps since we were kids, and this shortens the duration of the survey and reduces the dropouts (Di Zio and Pacinelli 2011). With the Spatial Delphi, the computation of the measures of geo-consensus and stability is easy and intuitive. Additionally, the interpretation of the feedbacks and of the geographical results is trivial, and does not require any statistical processing, unlike the other versions of the Delphi method.

Of course, in view of these advantages, there are also weaknesses. Like the conventional Delphi, the Spatial Delphi consists of a certain number of rounds, so the participants are forced to respond at any round and within the requested temporal intervals. The Spatial Delphi is based on a non-interactive map; therefore, the research team must decide the type, the scale and the extent of the map a priori. This is a limitation because respondents are not allowed to change the type of map, or explore the study area by moving the map or zooming on it, such as it happens with any interactive map (Di Zio et al. 2016).

The supporting materials are separated from the map and, inevitably, the number of maps and documents is limited. For each new application, all the necessary must be specially prepared; therefore, the time and costs for the preparation of a survey are considerable. Another disadvantage is that all the boxes for the reasonings are external to the map and, above all, are not interactive, as, for example, it is in the Real Time Delphi (Gordon and Pease 2006; Gordon 2009b), and this can discourage the participants in giving arguments.

Examples of Applications

The earliest application of the Spatial Delphi concerned the identification of the riskiest and safest areas in case of an earthquake (Di Zio and Pacinelli 2011). The study area was a surface of 2700 sq km around the city of L'Aquila, in the Abruzzo region (Italy). The aim was not to find the area where a possible earthquake is more or less probable (which is typically a geological issue), but two little areas considered the most dangerous and the most secure in case of a future seismic event. The authors, in 2010, organized a panel of 12 experts, who knew the area very well and with different expertise (geologists, seismologists, geographers, sociologists, demographers, architects and construction engineers).

The experts responded anonymously to three Spatial Delphi rounds, by placing on a map respectively three, two and one opinion-points for the places deemed riskier, and three, two and one for the safer. In the first round, the question posed to the experts was "In the event of an earthquake […], please indicate three human settlements on the map that you believe have the greatest risk and three you believe have the lowest risk" (Di Zio and Pacinelli 2011). After collecting the answers, a circle of consensus for the maximum risk and one for the minimum risk were constructed. In addition to the basic map, also supporting maps were included (road network, major resorts, a relief map, a satellite image and a seismic hazard map).

In the second round, the maps with two circles of consensus were circulated among the panelists, which were asked to locate two points for the major risk and two for the minor risk, preferably inside the two circles. With the two clouds of points, two new circles were calculated and proposed to the panel for the third round of consultation, for which only one opinion-point for each circle was asked. Some expert gave points external to the circles, giving also reasoning which circulated anonymously.

The Spatial Delphi was stopped after the third round, and in Tables 1.2 and 1.3, we report, for each round, the total number of collected points, the area of the circle of consensus (A_{best}), the indices MC_1 and MC_2, and the

Table 1.2 Results for the major risk

Round	n	A_{best} (Km^2)	MC_1	MC_2	FD
1	36	286.52	0.8939	100.00%	0.400
2	24	43.47	0.9839	15.17%	0.180
3	12	2.06	0.9992	0.72%	0.060

Table 1.3 Results for the minor risk

Round	n	A_{best} (Km^2)	MC_1	MC_2	FD
1	36	591.37	0.7810	100.00%	0.631
2	24	122.72	0.9545	20.75%	0.350
3	12	23.84	0.9912	4.03%	0.122

estimated fractal dimension (FD). It is worth noting how the size of both circles decreases rapidly during the rounds, until the identification of two small areas in respect to the study area. In fact, in the final round, the indices MC_1 are both very close to one and MC_2 very close to zero.

In addition, the estimated FDs confirm the fast convergence of the spatial opinions and the final geo-consensus. It is also evident that in the case of major risk the convergence was faster and the final geo-consensus greater compared to the minor risk. Thus, in the study region, it was easier for the experts to identify the most dangerous area than the safest. The L functions (not reported here) confirm these conclusions (Di Zio and Pacinelli 2011).

It is interesting to note that in this application there have been no abandonments of experts. In Fig. 1.3, we have all the 36 opinion-points given by the experts in the first round for the major risk question, together with the resulting first circle of consensus (CC^1). In Fig. 1.4, the three circle of consensus—CC^1, CC^2 and CC^3—are depicted without the opinion-points, in order to show as the circle moved and reduced during the survey.

We now report the first application of the Spatial Shang, made by Di Zio and Staniscia (2014a). In 2001, an Italian National Law (L. 93/2001, art. 8) established the *Costa Teatina National Park*, on the east coast, but did not define its boundaries. A number of questions have arisen: Which municipalities in the area should be included within the Park? What are the criteria to define the boundaries? What are the procedures? For more than ten years, an agreement was not reached, so a research group of the nearby "G. d'Annunzio" University (Chieti-Pescara) proposed the application of a Spatial Shang with the involvement of the local stakeholders, including local public authorities, local communities, enterprises, NGOs and associations. The spatial problem concerned the definition of a buffer small enough to help in defining the boundaries of the Park. A panel of 62 stakeholders ($n = 62$), representing the community from different perspectives, participated in the study and were selected on the basis of three criteria: (1) deep knowledge of the territory and of the conflict about the Park; (2) capacity to represent a clear position in the conflict and (3) capacity to give voice to the category they were representing (Di Zio and Staniscia 2014a).

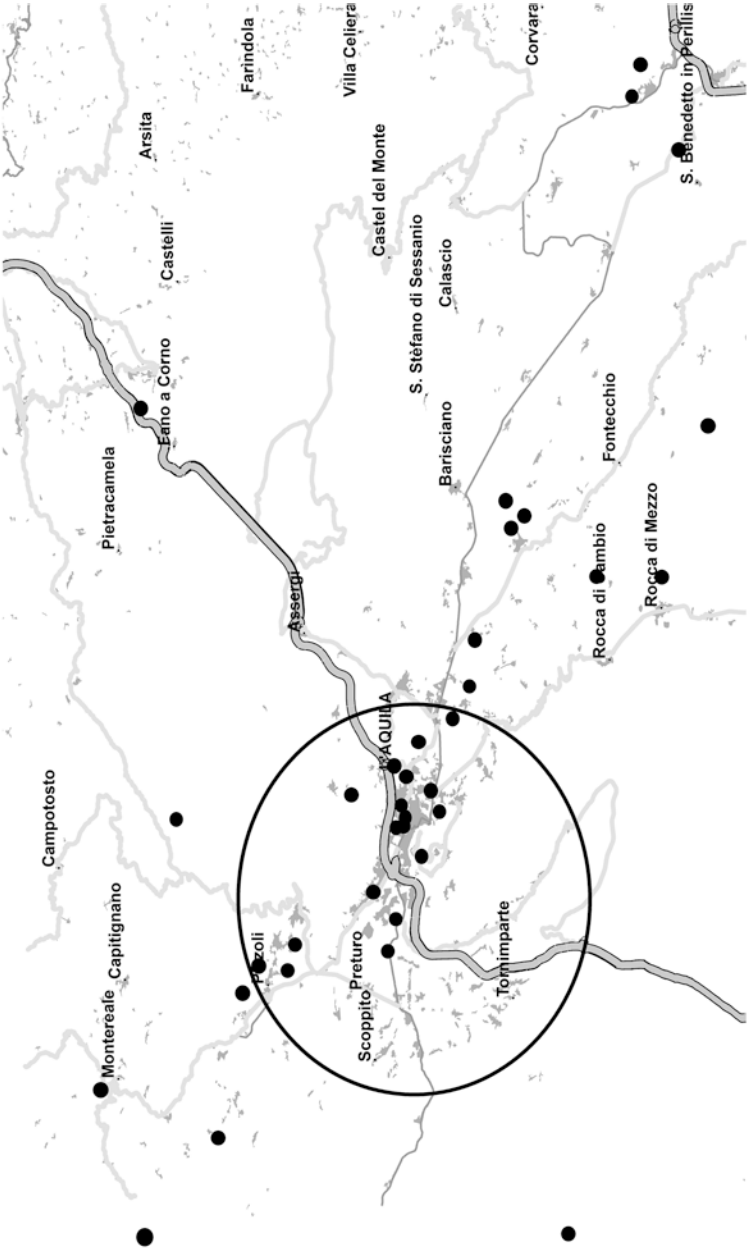

Fig. 1.3 Opinion-points and relative circle of consensus

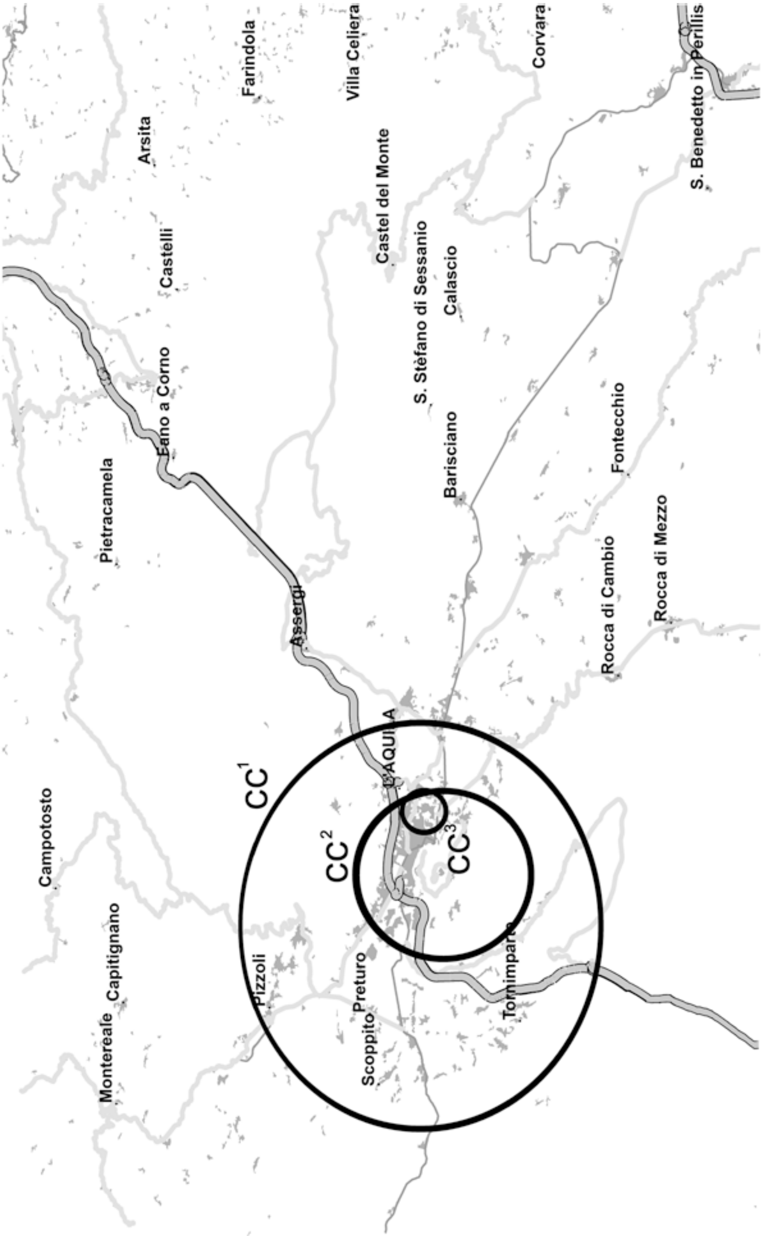

Fig. 1.4 The three circles of consensus for the major risk

Given that the problem of the boundary comes down to a line at a certain distance from the shoreline, the Spatial Shang was reduced from two dimensions to one. Therefore, the analysis was conducted only in one direction and precisely that of longitude (East-West).

On a map containing the eight municipalities involved in the park's area, the stakeholders were asked to indicate two points, representing the minimum and maximum distance of the boundary from the coastline. After the first round, the data consisted of two vectors, one for the east limits (***E***, representing the minimum evaluations), and one for the west limits (***W***, representing the maximum evaluations), both with *n* values. All the points were transformed in distances from the coastline (with a GIS software), and two arithmetic means were calculated. The results were $E_0 = 2.0$ km and $W_0 = 3.1$ km identifying two lines parallel to the coastline, which delimited a first big buffer of convergence, 1.1 km wide (see Fig. 1.5). The first central value was then $C_0 = 2.55$ km.

In the second round, the stakeholders were asked whether the limit of the park should have been back or ahead the line, drawn on the map, 2.55 km away from the coast. At the second consultation, only 35 stakeholders responded to the Spatial Shang questionnaire, and this high

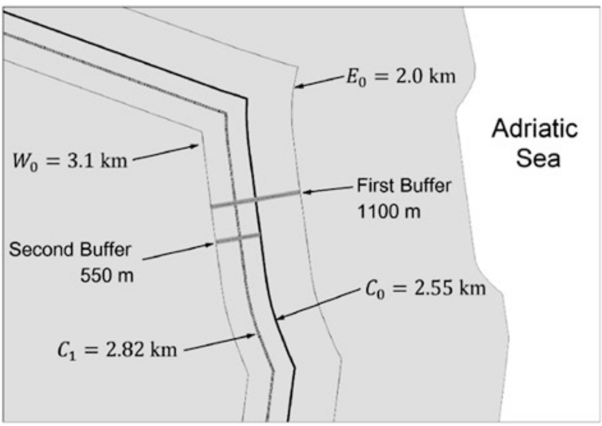

Fig. 1.5 Results of the Spatial Shang

dropout rate was due to the strong conflicts, which have arisen among the stakeholders around this issue (Di Zio and Staniscia 2014a).

Since 65.7% of the respondents declared a limit above the central line C_0, the new area of convergence resulted in a buffer of 550 m, confined between 2.55 km and 3.10 km from the coastline (Fig. 1.5). With these two extremes, the second central value was found to be $C_1 = 2.82$. Given that the area of convergence was small enough, the research group decided to conclude the iterations, and the geographical result was that the boundary of the park should be at about 2.8 km away from the coast and preferably in a strip of land between 2.55 km and 3.10 km from the coastline.

According to the research group, the application of the Spatial Shang had the advantage of stimulating a debate among the main players and represented a way to mitigate the conflicts. The geographical solution was proposed to the competent authorities as a decision support for the definition of the boundaries of the park.

Conclusions and Future Developments

We have seen how the Delphi method is extremely helpful when quantitative data are insufficient or absent, and why it is one of the most widely used techniques to overcome the distortions inherent in the interacting groups (effect of the leadership, spiral of silence and groupthink).

The structuring of the communication among the panelists according to the principles of the anonymity, the iterative structure and the asynchronicity are the main features of the Delphi technique, and all the other methods derived from it, developed in more than half a century, revolve around the same pillars.

In recent years, some scholars developed a new line of research, based on the introduction of the geographical element and the GIS technology (Dragicevic and Balram 2004; Jankowski et al. 2008; Di Zio and Pacinelli 2011). The Spatial Delphi (Di Zio and Pacinelli 2011) and the Spatial Shang (Di Zio and Staniscia 2014a) are the two methods discussed in this chapter and have been designed to treat forecast/decision problems related to the territory. The main innovation resides in the transition from the concept of "consensus" of the conventional Delphi to the new concept of "geo-consensus", in which the experts pursue a convergence of opinions on a limited geographical area.

The experts of the panel answer the spatial questions simply drawing points on a map, and the process of the convergence of opinions is leaded by means of geometric figures. The whole process is fast and easy, and allows for the calculation of a number of measures of geo-consensus and stability. The results are both geographical (the circles or the rectangles) and non-geographical (the comments of the experts and the measures of geo-consensus).

Some of the limits of the Spatial Delphi are the same of the classical Delphi, such as the presence of rounds, the manual computation of the synthesis and the non-interactive structure of any component (questionnaire, supporting materials and boxes for reasonings).

One interesting version of the Delphi method, which overcomes the previous limitations, is the Real Time Delphi invented by Gordon (Gordon and Pease 2006; Gordon 2009b). It is a computerized Delphi, which does not provide for subsequent rounds, therefore leading to a greater efficiency in terms of execution time. The absence of repeated rounds allows the simultaneous computation and delivery of the responses, the possibility of using a large number of participants, low realization costs and high efficiency with regard to the time frame needed to perform the survey. With this method, respondent are not compelled to complete the entire questionnaire in one working session, and can benefit of interactive boxes for comments and reasons.

Recently, a new version of the Spatial Delphi has been developed, from bringing together the logic of the Real Time Delphi, which is roundless and interactive, and the potential of the Spatial Delphi in the management of geographical issues. The method is called *Real Time Spatial Delphi*, which allows the consultation of experts on issues related to the territory in an efficient, real-time way, with very short times and low costs (Di Zio et al. 2016). The system automatically calculates and displays the circles of consensus that shrink and move in real-time during the survey, as well as the measures of geo-consensus and stability. The authors applied the method to the zoning of street prostitution, in Italy, identifying five areas of consensus where the experts considered the zoning most appropriate.

This method preserves most of the advantages of both the Spatial Delphi and the Real Time Delphi, minimizing the disadvantages and opening the way for a number of possible future developments. As suggested by the authors, examples of fields of applications are "architecture, landscape gardening, war games, pinpointing points of origin and potential courses of epidemics, location of future crimes, location of computer hackers, planetary exploration" and even 3D spatial application (Di Zio et al. 2016). The *Real Time Spatial Delphi* runs on a WebGIS platform, with a number of tools and functionalities that make it flexible and usable for a wide range of applications.

We are currently working on further developments of the system, like the *Real Time Spatial Shang*, which uses interactive rectangles instead of circles. Furthermore, we are working on the automatic detection of clusters of opinion-points. In fact, in real applications, starting with a single circle of convergence is a limitation, because it could happen that on the map emerges a number of clusters of points, denoting that the number of suitable places for the solution of the spatial problem is greater than one.

It is worth noting that these studies are inserted in a wider line of research, where the judgments of a panel of experts are collected via web GIS applications. We already have applications in tourism satisfaction (Sarra et al. 2015), in geo-marketing (Di Zio and Fontanella 2012), and for the perception of the risk of terrorist attacks. More precisely, in these applications, the elaborations of the results of the consultation have been performed with a statistical modeling approach known as *Item Response Theory* (de Ayala 2009; De Mars 2010).

In conclusion, the *geographical space*, the *Delphi* logic and the *real time* approach can be combined and constitute the key features of these new methods. Further potentials will be achieved when they will be used in combination with other methodologies, like scenario method techniques, technology list or others.

References

Bailey, T. C., & Gatrell, A. C. (1995). *Interactive Spatial Data Analysis*. New York: J. Wiley.

Brockhaus, W. L., & Mickelsen, J. F. (1975). An Analysis of Prior Delphi Applications. *Technological Forecasting and Social Change, 10*(1), 103–110.

Chatterjee, S. (1975). Reaching a Consensus: Some Limit Theorems. *Bulletin of the International Statistical Institute, 46*(3), 156–160.

Chung, K. H., & Ferris, M. J. (1971). An Inquiry of the Nominal Group Process. *Academy of Management Journal, 14*(4), 520–524.

Dajani, J. S., Sincoff, M. Z., & Talley, W. K. (1979). Stability and Agreement Criteria for the Termination of Delphi Studies. *Technological Forecasting and Social Change, 13*(1), 83–90.

Dalkey, N. C., & Helmer, O. (1963). An Experimental Application of Delphi Method to the Use of Experts. *Management Science, 9*(3), 458–467.

de Ayala, R. J. (2009). *The Theory and Practice of Item Response Theory*. New York: Guilford Press.

De Groot, M. H. (1974). Reaching a Consensus. *Journal American Statistical Association, 69*(345), 118–121.

De Mars, C. (2010). *Item Response Theory. Understanding Statistics Measurement*. Oxford: Oxford University Press.

Di Zio, S., & Fontanella, L. (2012). Public Geomarketing: Georeferencing IRT Models to Support Public Decision. *Statistica Applicata—Italian Journal of Applied Statistics, 24*(3), 301–320.

Di Zio, S., & Pacinelli, A. (2011). Opinion Convergence in Location: A Spatial Version of the Delphi Method. *Technological Forecasting and Social Change, 78*(9), 1565–1578.

Di Zio, S., & Staniscia, B. (2014a). A Spatial Version of the Shang Method. *Technological Forecasting and Social Change, 86*, 207–215.

Di Zio, S., & Staniscia, B. (2014b). Citizen Participation and Awareness Raising in Coastal Protected Areas. A Case Study from Italy. In A. Montanari (Ed.), *Mitigating Conflicts in Coastal Areas Through Science Dissemination: Fostering Dialogue Between Researchers and Stakeholders* (pp. 155–197). Rome: Sapienza Università Editrice.

Di Zio, S., Castillo Rosas, J., & Lamelza, L. (2016). Real Time Spatial Delphi: Fast Convergence of Experts' Opinions on the Territory. *Technological Forecasting and Social Change*, on line doi:10.1016/j.techfore.2016.09.029.

Dragicevic, S., & Balram, S. (2004). A Web GIS Collaborative Framework to Structure and Manage Distributed Planning Processes. *Journal of Geographical Systems, 6*(2), 133–153.

Ford, D. A. (1975). Shang Inquiry as an Alternative to Delphi: Some Experimental Findings. *Technological Forecasting and Social Change, 7*(2), 139–164.

Glenn, J. C. (2009). Participatory Methods. In J. C. Glenn & T. J. Gordon (Eds.), *Futures Research Methodology*, *CD-ROM Version 3.0*. Washington: The Millennium Project, American Council for the United Nations University.

Gordon, T. J. (2009a). The Delphi Method. In J. C. Glenn & T. J. Gordon (Eds.), *Futures Research Methodology*, *CD-ROM Version 3.0*. Washington: The Millennium Project, American Council for the United Nations University.

Gordon, T. J. (2009b). The Real-Time Delphi Method. In J. C. Glenn & T. J. Gordon (Eds.), *Futures Research Methodology*, *CD-ROM Version 3.0*. Washington: The Millennium Project, American Council for the United Nations University.

Gordon, T. J., & Helmer, O. (1964). *Report on a Long-Range Forecasting Study. R-2982*. Santa Monica: The Rand Corporation.

Gordon, T. J., & Pease, A. (2006). RT Delphi: An Efficient, "Round-Less" Almost Real Time Delphi Method. *Technological Forecasting and Social Change, 73*(4), 321–333.

Grime, M. M., & Wright, G. (2016). *Delphi Method*. New York: Wiley StatsRef: Statistics Reference Online: John Wiley & Sons.

Gustafson, D. H., Shukla, R. K., Delbecq, A. L., & Walster, W. G. (1973). A Comparative Study of Differences in Subjective Likelihood Estimates Made by Individuals, Interacting Groups, Delphi Groups, and Nominal Groups. *Organizational Behavior and Human Performance, 9*(2), 280–291.

Hassan, G. (2013). Groupthink Principles and Fundamentals in Organizations. *Interdisciplinary Journal of Contemporary Research in Business, 5*(8), 225–240.

Hastings, H. M., & Sugihara, G. (1993). *Fractals: A User's Guide for the Natural Sciences*. Oxford, UK: Oxford University Press.

Hoffman, L. R. (1965). Group Problem Solving. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology*. New York: Academic Press.

Jankowski, P., Swobodzinski, M., & Ligmann-Zielinska, A. (2008). Choice Modeler: A Web-Based Spatial Multiple Criteria Evaluation Tool. *Transactions in GIS, 12*(4), 541–561.

Neill, S. A. (2009). *The Alternate Channel: How Social Media Is Challenging the Spiral of Silence Theory in GLBT Communities of Color*. Washington, DC: American University.

Riggs, W. E. (1983). The Delphi Technique, an Experimental Evaluation. *Technological Forecasting and Social Change, 23*(1), 89–94.

Ripley, B. D. (1976). The Second-Order Analysis of Stationary Point Processes. *Journal of Applied Probability, 13*(2), 255–266.

Rowe, G., & Wright, G. (2001). Expert Opinions in Forecasting: The Role of the Delphi Technique. In J. Scott Armstrong (Ed.), *Principles of Forecasting* (pp. 125–144). Boston: Kluwer Academic.

Sarra, A., Di Zio, S., & Cappucci, M. (2015). A Quantitative Valuation of Tourist Experience in Lisbon. *Annals of Tourism Research, 53*, 1–16.

Scheibe, M., Skutsch, M., & Schofer, J. L. (1975). Experiments in Delphi Methodology. In H. A. Linstone & M. Turoff (Eds.), *The Delphi Method: Techniques and Applications* (pp. 262–287). Reading: Addison-Wesley.

Torrance, P. E. (1957). Group Decision Making and Disagreement. *Social Forces, 35*(4), 314–318.

Turoff, M. (1970). The Design of a Policy Delphi. *Technological Forecasting and Social Change, 2*(2), 149–171.

Van de Ven, A. H. (1974). *Group Decision Making and Effectiveness: An Experimental Study*. Kent: Kent State University Press.

Van de Ven, A. H., & Delbecq, A. L. (1974). The Effectiveness of Nominal, Delphi, and Interacting Group Decision Making Processes. *The Academy of Management Journal, 17*(4), 605–621.

von der Gracht, H. A. (2012). Consensus Measurement in Delphi Studies. Review and Implications for Future Quality Assurance. *Technological Forecasting and Social Change, 79*(8), 1525–1536.

2

Interactive Scenarios

Theodore J. Gordon and Jerome Glenn

Introduction: Some Definitions

A scenario is a rich and detailed portrait of a plausible route to a future world, a portrait sufficiently vivid to illustrate problems, challenges, and opportunities that might have to be faced as the future unfolds. Scenarios resemble short stories that describe chains of causality and key decisions along the way to the future. An *exploratory scenario* is a plausible description of what *might* occur. A *normative scenario* describes a route from the present to a desired future and thus is a plausible description of what *is hoped* will occur.

Some of the material in this chapter is drawn from two papers by the authors that appeared in Futures Research Methodology version 3.0, a CD ROM, The Millennium Project, Washington, DC, and from prior studies conducted by The Millennium Project. The design and results of these studies can be retrieved from the Project's Global Futures Intelligence System (GFIS) web site at www.themp.org.

T.J. Gordon (✉) • J. Glenn
The Millennium Project, Washington, D.C, USA

L. Moutinho, M. Sokele (eds.), *Innovative Research Methodologies in Management*,
DOI 10.1007/978-3-319-64400-4_2

No scenario is ever seen as probable; the probability of any scenario ever being realized is vanishingly small. It is not accuracy that is the measure of a good scenario, it is rather

- plausibility (telling the story about getting from here to there);
- internal self-consistency;
- usefulness in decision-making.

Sets of scenarios are often used in planning; if the sets encompass a broad span of futures, and plans are generated that cope with the issues described in all of the scenarios, then the plans are considered to be robust and the future can be met with some degree of confidence. In other words, we might not be sure of the world that will emerge, but if a policy is judged to work in a broad range of scenarios, then we assume its success is independent of the actual world that will emerge. For this to be so, of course, the set of scenarios must cover all corners of future *scenario space*.

The term “scenario” comes from the dramatic arts. A scenario in the theater refers to an outline of the plot and in movies is a summary or set of directions for the sequence of action. Of all futures research techniques, scenarios are the most frequently and widely used.

Scenarios can be classed as qualitative or quantitative, normative or exploratory, written as snapshots in time or as future histories. They can be designed to be decision aids, educational tools, and political instruments. They can be static or dynamic, produced by single individuals or groups, result totally from imagination, or be based on simple or complex models.

Herman Kahn, considered the father of scenario construction, introduced the term “scenario” into planning in connection with military and strategic studies conducted by the RAND Corporation in the 1950s. He and Anthony Wiener describe the functions of scenarios as follows:

1. They serve to call attention, sometimes dramatically and persuasively, to the larger range of possibilities that must be considered in the analysis of the future.

2. They dramatize and illustrate the possibilities they focus on in a very useful way. (They may do little or nothing for the possibilities they do not focus on.)
3. They force the analyst to deal with details and dynamics that he might easily avoid treating if he restricted himself to abstract considerations.
4. They help to illuminate the interaction of psychological, social, economic, cultural, political, and military factors, including the influence of individual political personalities upon what otherwise might be abstract considerations, and they do so in a form that permits the comprehension of many such interacting elements at once.
5. They can illustrate forcefully, sometimes in oversimplified fashion, certain principles, issues, or questions that might be ignored or lost if one insisted on taking examples only from the complex and controversial real world.
6. They may also be used to consider alternative possible outcomes of certain real past and present events, such as Suez, Lebanon, Laos, or Berlin.
7. They can be used as artificial "case histories" and "historical anecdotes" to make up to some degree for the paucity of actual examples.[1]

Paul Raskin and Eric Kemp-Benedict of the Tellus Institute which developed the PoleStar global model that has been used extensively to examine environmental futures explain the relationship between algorithm-derived numerical estimates of change and scenarios that employ them or on which the models are based, as follows:

> Compelling scenarios need to be constructed with rigor, detail, and creativity, and evaluated against the criteria of plausibility, self-consistency and sustainability, a process that requires thorough and intensive analysis. Scenarios draw on both science – our understanding of historical patterns, current conditions, and physical and social processes – and the imagination to conceive, articulate and evaluate alternative pathways of development and the environment. In so doing, scenarios can illuminate the links between issues, the relationship between global and regional development, and the role of human actions in shaping the future. It is this added insight, leading to more informed and rational action that is the foremost goal of scenarios, rather than prediction of the future.[2]

Now, many decades and thousands of scenarios later, these statements still hold. The early scenarios were usually qualitative and prepared by single authors.

In this chapter, we are most interested in *interactive scenarios*. Such scenarios result from creative processes that involve people (experts, users, broader audiences), and endogenous or exogenous models that are based on historically demonstrated statistical relationships or judgments about these relationships. These models can guide the scenarios' statements of causality and are used to provide quantitative context and help assure internal self-consistency. The interactions among people, models, and all permutations take place during their formation or during their use.

We describe several interactive modes:

- ***Designing Content***: A Delphi questionnaire might be used in the design of a set of scenarios to collect suggestions for scenario axes (major dimensions that define a domain of interest) or technological, economic, political, or other driving elements that will propel the scenario narrative forward.
- ***Cooperative Feedback:*** The classic approach to scenario generation is for a single individual, or small group to simply write a description about how some aspects of the future might evolve from the present. The interactive part of the process comes from feedback and comments from readers and users of the scenarios and may lead to redrafting of the story.
- ***Using Models***: A cross-impact matrix, Futures Wheel, or other modeling system might be used to develop the richness of consequences and relations among possible events and trends. These can also be built into a scenario substrate to show how changing one statement in the scenario might affect other portions of the scenario.
- ***Filling in the Blanks***: A scenario might be written with key decision points open, to be filled in by voting participants. The remainder of the scenario then illustrates the consequences of the selected decision; for example, a movie or play with three alternative endings that depend on a jury decision provided by the audience.
- ***Scenario Clustering***: An approach in which large numbers of projections or scenarios are constructed to avoid the burden imposed by point scenarios.

These approaches are described in more detail in the following subsections of this chapter.

Designing Content

The designer of a single scenario or set of scenarios may solicit suggestions from colleagues and experts about the principal features to be addressed. The first step in constructing a set of useful scenarios is the choice of axes or factors about which the scenarios of the set differ. For example, a set of global scenarios might use

- Demographic change, including changes in family size, migration to cities, and population aging
- Technological change, which affects our society and economy as well as the tools available for education and other part of the knowledge system.

If each of these factors had three levels of expression: high, medium, and low, then nine scenarios could be constructed using the permutations of the axes, as shown in Fig. 2.1. If there had been 3 axes with 3 possible levels, then 27 possible scenarios could have been defined.

Thus, scenario 1 would have an assumption of high demographic growth and high technological development and scenario 7, low demographic growth and high technological development. In using this

Num ber	*Demogr High*	*Demogr Med*	*Demogr Low*	*Tech High*	*Tech Med*	*Tech Low*
1	*X*			*X*		
2	*X*				*X*	
3	*X*					*X*
4		*X*		*X*		
5		*X*			*X*	
6		*X*				*X*
7			*X*	*X*		
8			*X*		*X*	
9			*X*			*X*

Fig. 2.1 Scenario axes define content

kind of layout, certain combinations would be much less plausible than others, and these could be put aside for later exploration. The more appealing combinations more useful for decision-making would (or should?) become the focus of the subsequent work.

The purpose of finding the appropriate axes is to help assure that scenarios in the set surround the "real" future so that policies that seem attractive in all scenarios are apt to be appropriate in the "real" future. The simplest application of this idea is to use three scenarios: two are devoted to "maximum" and "minimum" cases, and the third to a "middle of the road" case. But by selecting axes and drivers, and permuting their characteristics, a much finer grained description can be obtained and most researchers use this approach.[3]

When interactive groups are enlisted to help create the scenario set, they may be asked in interviews or questionnaires to nominate and help choose the axes and the most appropriate combinations. They might also be asked to identify events that should be considered (e.g. changes in fertility, attitudes toward family size, new contraception modes, etc.) and judge the probability of the events in the scenario context (new contraception modes more likely in the high tech world). They might also be asked to help name the scenarios of interest. This is important because the names help later recall of the content, and ingenious names help popularize and de-mystify the process. Some names of past scenarios have been, for example:

- Win-Win
- Seamless Nations
- Big Brother
- Cybertopia
- The Rich Get Richer
- Trading Places
- Systems Breakdown
- Ostrich
- Lame Duck
- Icarus
- Flight of the Flamingos
- The Darkside of Exclusivity

- New Power Politics
- Clash of Modernities

A good example of a questionnaire designed to collect judgments for scenario content can be found in the Millennium Project's study of the **future of work**. The Millennium Project has over 60 "Nodes" around the world consisting of institutions and organizations that have agreed to work together in order to address world issues. The interaction on the future of work scenarios began with a review of recent literature on the topic, and a series of interviews led to the formation of a Real Time Delphi (RTD). The RTD questionnaire posed a number of questions that would be useful in forming later scenarios. The questions were:

1. If socio-political-economic systems stay the same around the world, and technological acceleration, integration, and globalization continue, what percentage of the world do you estimate could be unemployed in the years 2020, 2030, 2040, and 2050?
2. More jobs were created than replaced during both the Industrial and Information Ages. However, many argue that the speed, integration, and globalization of technological changes of the next 35 years (by 2050) will cause massive structural unemployment. What are the technologies or factors that might make this true or false?
3. What questions have to be resolved to answer whether Artificial Intelligence (AI) and other future technologies create more jobs than they eliminate?
4. How likely and effective could these actions be in creating new work and/or income to address technological unemployment by 2050?
5. Will wealth from artificial intelligence and other advanced technologies continue to accumulate income to the very wealthy increasing the income gaps?
6. How necessary or important do you believe some form of guaranteed income will be necessary to end poverty, reduce inequality, and address technological unemployment?
7. Do you expect that the cost of living will be reduced by 2050 due to future forms of AI robotic and nanotech manufacturing, 3D/4D printing, future Internet services, and other future production and distribution systems?

8. What high impact events, developments, surprises, wild cards, or black swans could change the future work-technology relationship?
9. What alternative scenario axes and themes should be written connecting today with 2050 describing cause and effect links and decisions that are important to consider today?

A final question was: "What would make these scenarios worth your time to read them and share with others?" Some of the group's answers were:

- Uncovering new aspects and perspectives of the impact of technology on human life (work and leisure)—in particular potential risky/unwanted situations/conditions … with specific recommendations for actions today
- Challenge basic concepts, such as unemployment, income gaps, technological displacement, classic economics, and so on.
- Include roadmaps and policy recommendations. They should be strategic scenarios, focusing in concrete challenging sectors and fields at global level and regional considerations.
- It would be most important to produce totally new, surprising ideas and approaches and show that we have not yet explored all the possibilities, because we have not. I think we should be ready to take further steps, for example, stop talking about (un) employment and launch new concepts instead: what is the intrinsic value behind work? Why do we need to work? Why not? So, asking "stupid" questions, that is the basic task for researchers.
- Vivid dramatizations, short but informative and conceptual. Good sci-fi authors on commission might be able to achieve this, but it is not easy. Maybe try a short story competition.
- The challenge for humanity is to do away with forced labor. I want to see how humanity can achieve this for all—without generating a species of robots that competes with humans for resources in the process.

Clearly, such comments could be useful to scenario writers. The figure below shows a small portion of the questionnaire[4] (Fig. 2.2)

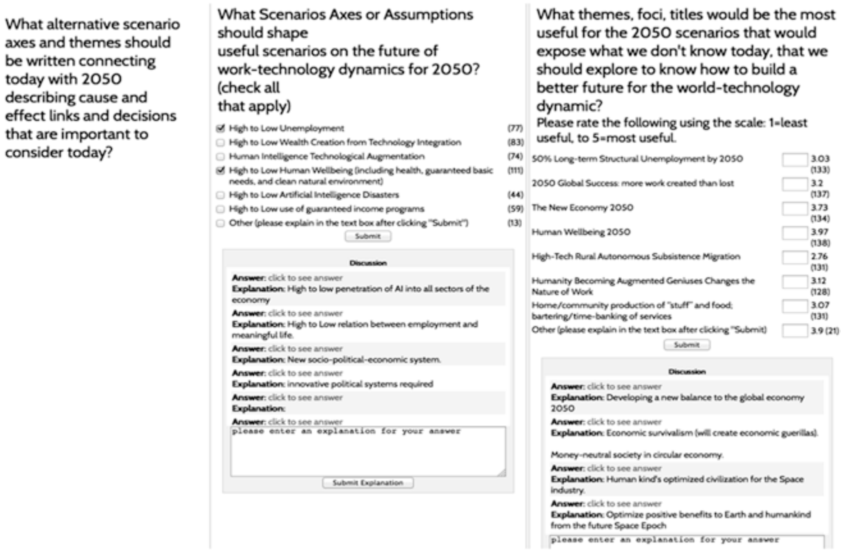

Fig. 2.2 RTD questionnaire on future of work

Note that the RTD process itself is interactive since inputs from prior respondents are displayed for each person to see in real time and to consider in drafting their own answers and suggestions.

Cooperative Feedback; Policy Feed-Forward

This approach to interaction is straightforward: write the scenario and then ask groups of people to critique it or improve it. Then, redraft considering their inputs. This approach is particularly applicable when it is anticipated that the scenario will be controversial, or when consensus from broad slices of society would be needed, as in the Mont Fleur Process. Mont Fleur is a conference center outside Cape Town, where a diverse group of 22 prominent South Africans met in 1991 (decades before the end of Apartheid) with a team of scenario writers from Shell Oil Company to create four scenarios. Funded by a private foundation, the scenarios were intended to "stimulate debate on how to shape the next ten years" for South Africa.[5]

One of the first successes of this project was to bring together the people and ideas from the extremes as well as the center, including the South African government, the African National Congress (ANC), the Inkatta, and the far right-wing extremists. The discussions were facilitated by Adam Kahane, a Shell employee at that time. The key axes involved in the scenario space were political settlement and economic policy. The outputs were series of papers and a very effective video presentation of the scenarios.

Four scenarios were produced as follows:

- "Ostrich," in which a negotiated settlement to the crisis in South Africa is not achieved, and the government continues to be non-representative: a white government sticks its head in the sand.
- "Lame Duck," in which a settlement is achieved, but the transition to a new dispensation is slow and indecisive; in trying to satisfy all, no one is happy.
- "Icarus," in which transition is rapid, but the new government unwisely pursues unsustainable populist economic policies; a constitutionally unconstrained black government engages in unsustainable public spending.
- "Flight of the Flamingos," in which the government's policies are sustainable and country takes a path of inclusive growth and democracy.

The first three of these portrayed dark worlds and in a very simple manner (using cartoons and bird fables) the scenarios highlighted the dangers ahead if a political settlement was not reached between the anti-apartheid movement and the Government. It also indicated the impacts that ill-advised economic policies could have on the future of South Africa. The scenarios were credited with nudging the National Party toward a negotiated settlement and convincing the ANC about the need for a sensible economic policy.

There was give and take in the design and discussion of content, and as external positions helped shape the scenarios, the scenarios helped shape the external positions. One reporter quoted Nelson Mandela as saying:

> "And I got frightened", Mandela recalls. "Before Trevor finished, I said to him, 'Now what does this mean as far as negotiations are concerned?

> Because it appears to me that if we allow the situation to continue … the economy is going to be so destroyed that when a democratic government comes to power, it will not be able to solve it.'" Mandela made a decision—the deadlock [in the then-stalled negotiations on a post-apartheid political order] must be broken.[6]

The scenarios were published in a 14-page insert in *The Weekly Mail* and *The Guardian Weekly*, major South African newspapers. Over the rest of the year, the team presented the scenarios to more than 50 influential groups throughout South Africa. A 30-minute video presenting the scenarios was also released.

The scenarios became widely discussed in South Africa at all levels, including taxi drivers and talk radio shows. The extent of the influence of the scenarios is not measurable, but all these years later we know that South Africa made a peaceful transition to representative government. It could have been much different.

Using Models

When the numbers take the spotlight, the scenarios are quantitative and are usually based on models that use algorithms and economic or social theory to connect cause and effect, for example, a scenario of a country's economic development might use an **econometric model** that examines (as a cause) the consequences of rapidly increasing energy prices or immigration overload primarily using statistical relationships derived from historical data. In 1972, Dennis Meadows and his colleagues produced a report titled "The Limits To Growth" for the Club of Rome using a modeling approach called **systems dynamics** in which stocks and flows of parameters that describe production and consumption occurring in an economy are captured in logical algorithms.[7]

It was not the first model designed to create scenarios as an aid to high-level decision-making. Larry Kline developed the Wharton Econometric Forecasting Model in the mid-1960s; it was used, for example, to examine the consequences of changes in commodity prices, tax rates, and tariffs on GDP, business conditions, and other economic factors.

Another model that has been used extensively as the basis for quantitative scenarios is the **International Futures** (Ifs) model developed at the University of Denver by Professor Barry Hughes. This model provides the backbone for scenarios included in the RAND analyses of education, poverty, and health in their series *Patterns of Potential Human Progress*.[8] This model has also been used by many government agencies, including the Department of Defense and the National Intelligence Council.[9]

The Ifs model is an example of a generalized macroeconomic/demographic model. Hughes says

> International Futures is a large-scale, long term, integrated global modeling system. It represents demographic, economic, energy, agricultural, sociopolitical, and environmental subsystems for 183 countries interacting in the global system. The central purpose of IFs is to facilitate exploration of global futures through alternative scenarios.[10]

Demographic models are also used frequently in forming quantitative pictures of future populations. These models are based on assumptions about fertility, mortality, migration, and other such factors and while complex in execution are often simple in concept. Imagine, for example, that we are interested in estimating the size of the future population of the world. We could reason that next year the population would be the same as this year's, plus the number of new babies born, minus the number of people who died. The scenarios would describe the evolution of factors that could affect the number of births and deaths (including, for example, new attitudes toward family size and suicide, epidemics, genetics, and life extension), and the model would be used to show the consequences of these assumptions. "Standard" assumptions about such factors have been made and are used by census bureaus and international agencies such as the United Nations Population Division to forecast details of the world's population canvas.[11] If our supposed demographic model were to deviate from those assumptions, we could substitute our own assumptions, derived from interactive interviews or an RTD conducted among experts. We could show the results of the model run, using the new judgments and say to the participants: "Taking your inputs about population-controlling factors and running our demographic models, leads to these forecasts. Do they look right to you? If not, what input assumptions would you change?" In this way, the model becomes part of an interactive scenario-generating feedback process.

Filling in the Blanks

Interactive scenarios can be constructed when it is up to the audience or readers of a scenario to complete them. The dramatic arts have used this technique for some time. For example, in *Night of January 16th*, a 1934 play by Ayn Rand, members of the audience act as a jury and based on testimony presented in the play, vote "guilty" or "innocent." Depending on the vote, one or another of the previously prepared endings is presented. More recently, *The Mystery of Edmond Drood*, a 1985 musical, requires that the audience vote on the killer's yet undisclosed identity and, depending on the outcome, different conclusions are presented. Other plays have used the same approach. Books and games have used the same interactive device.

In futures research, readers of a draft scenario can be asked to "fill in the blanks" by supplying words that are missing from the text. Drop-down menus could also be used to present a few alternative choices to the participants. Consider this excerpt derived from a Millennium Project scenario on the future of work.

> During the early 21st century, political leaders were so mired in short-term political conflicts, and me-first, selfish economic thinking that they did not anticipate how fast [*1. PLEASE ENTER THE MOST APPROPRIATE WORD OR PHRASE*], and other technologies would make business after business obsolete beginning dramatically in the late 2020s and early 2030s. Too many economists, lawyers, and [*2. PLEASE ENTER THE MOST APPROPRIATE WORD OR PHRASE*] who knew little of the coming technology-induced unemployment crowded out those with knowledge of what was coming. Corporate lobbyists protected short-term profit decisions. Most of the political/economic systems around the world did not reward long-term strategic planning but rewarded short-term profits and immediate political favors. Hence there were no long-term strategies in

> place to reduce the devastating impacts of the dramatic growth in unemployment around the world, especially in *[3. PLEASE ENTER THE MOST APPROPRIATE WORD OR PHRASE*].

The analysts would simply compile the suggested words for the three blanks, decide if they were appropriate, and insert them in the final draft. The questionnaire might also ask the respondents to supply reasons for their suggestions, and these could be fed back to the participants and lead to insights that would have otherwise remained hidden.

It is often possible in scenarios to identify key decision points and allow readers to choose the decision; the trick is then to amend the scenario to reflect the outcome of the readers' votes. This may be accomplished through the use of a simple program not visible to the readers and an imbedded RTD program that collects judgments and responds to the choices by displaying new text depending on the vote. This approach involves the prior construction of a cross-impact matrix.

Cross-Impact Modeling

This interactive technique focuses on cause/effect relationships among scenario statements; it is a systematic means for creating and expanding scenarios, for testing internal consistency, and for incorporating feedback of judgments in interactive applications. It was invented by Gordon and Helmer in about 1968 and has been used in numerous studies since then. It has also been used in combination with other methods such as Delphi and systems dynamics.

In their initial application of cross-impact principles, Gordon and Helmer developed a game for Kaiser Aluminum and Chemical Company called *Future*. The company produced many thousands of copies of the game and used it for promotional gifts in conjunction with its 100th anniversary. The game, which is long out of print, involved a series of cards, each describing a single future event. The cards were given an a priori probability of occurrence, based largely on Gordon and Helmer's judgment. Then a die was rolled to determine whether or not, in the scenario that was being constructed, the event "occurred." In the game, the die was an icosahedron with numbers written on the faces to correspond to the probability that the face would turn up (e.g. 20% of the faces displayed the number 20; 80% of the faces displayed the number 80, etc.). If the probability shown on the die face was equal to or greater than the event probability, it was said to have "occurred."

If an event occurred, the card was flipped over. On the back face of the card, the "cross impacts" were described: for example, "if this event happens, then the probability of event 12 increases by 10 percent; the probability of event 53 decreases by 15 percent, etc." Brief reasons were given for the stated interactions, and a simple system (plastic sliders on a track, as shown in Fig. 2.3) was provided for keeping track of the evolving probabilities as the game progressed.

At the end of the game, one stack of cards represented events that had happened, and another stack, events that had not. The stacks of cards represented a scenario that was, in fact, determined by chance, the predetermined probabilities, and the cross impacts.[12]

Fig. 2.3 *Future* game

Gordon and Hayward programmed the approach at the University of California Los Angeles (UCLA, USA) in 1968.[13] The conditional probabilities were expressed as impact coefficients and ranged from −10 to +10. The first programs played almost exactly like the *Future* game: events were chosen in random order, decided, and the probabilities of cross-impacted events then determined. One "play" was completed when all events were decided. Then, in Monte Carlo fashion, the process was repeated many times.[14] The computer kept track of the number of scenarios that contained each event. This count of event "occurrences" was used to compute the final probabilities of the events, given their cross impacts.

The game appeared in many classrooms in the 70s, reworked to address the problems under study (e.g. urban crises). Gordon, Rochberg, and Enzer at The Institute for the Future experimented with a form of cross impact that included time series rather than "snapshot" approach.[15] Norman Dalkey used conditional probabilities in the matrix.[16] Helmer applied the approach to gaming,[17] and more recently, Heiko used it as a gaming tool in decision-making.[18]

KSIM, a simulation technique developed by Julius Kane, was based on the expected interactions among time-series variables rather than events.[19] In this approach, Kane treated all of the variables as a percentage of their maximum value, and the cross impacts were used to adjust the variables in each time interval.

Duval, Fontela, and Gabus at the Battelle Institute in Geneva developed EXPLOR-SIM, a cross-impact/scenario approach,[20] and Duperrin and Godet developed SMIC, a cross-impact approach that asks experts to provide initial, conditional occurrences, and conditional non-occurrence probabilities and to form scenarios based on the cross-impact results.[21]

At The Futures Group, probabilistic systems dynamics was a joining of systems dynamics and a time-dependent version of cross impact, an approach first explored by John Stover in stimulating the economy of Uruguay[22] A simulation method, called Interax incorporated cross-impact concepts, was developed by Selwyn Enzer at the University of California (USA).[23] Ducos integrated Delphi and cross impact.[24] Bonnicksen at Texas A&M University (USA), in a process called EZ-IMPACT, used the cross-impact approach in a workshop gaming application to explore policy options among contentious parties.[25] His algorithm was based on Kane's[26] approach.

The cross-impact method has been applied to many research questions on a stand-alone basis or in combination with other techniques. Godet, for example, lists application of SMIC to subjects as diverse as aircraft construction, world geopolitical evolution, the nuclear industry, and corporate activities and jobs to the year 2000.[27] Other[28] examples include Brent Vickers in 1992 studying the European automobile industry and Albert Schuler et al. studying the softwood lumber industry in Canada.[29] Fontela and Rueda-Cantuche showed the application and similarities of cross-impact methods to input-output scenario analysis.[30] Winterscheid used cross-impact analysis in the study of flood risks in a portion of the Rhine with 25 experts completing independent cross-impact matrixes linking flood cause and effects.[31]

Bañuls, Turoff, and Hiltz used a group collaborative scenario that incorporated a cross-impact matrix and Interpretative Structural Modeling (CIA-ISM)[32] in an elegant study directed toward improving emergency planning and preparedness.[33] They imagined a dirty bomb explosion on the top floor of a crowded urban shopping center and built a cross-impact model linking a few key assumptions about pre-conditions (such as decontamination preparedness), events that might occur during the response (dynamic events, such as an information leak to the media and participation of the military), and outcome descriptions (such as loss of income). They collected judgments about the initial probability and conditional probability estimates and used Bayesian equations to compute the group response from the individual inputs. Finally, a sensitivity analysis identified how various courses of action could change the expected outcomes of the model. The method they describe is meticulous and certainly has wide applicability to complex issues.

In the remaining portions of this section, we illustrate a few cross-impact matrixes and show their use in modifying scenarios:

The US Central Intelligence Agency has used cross-impact analysis in in a number of studies; one forecasted the political future of Rhodesia.[34] More recently, a cross-impact matrix approach was recommended as a tool to intelligence analysts.[35] Here, however the cells were filled with qualitative judgments about the influence of one entry on another, rather than quantitative estimates of conditional probabilities.

In more complex applications, the events listed in the rows could be assigned a probability and judgments about conditional probabilities placed in the cells. The matrix could be played out using a Monte Carlo program to obtain a quantitate estimate of the outcome. In this approach, expert judgments are provided for all of the events on the rows and judgments for all of the conditional probabilities in the cells. The play of the matrix results in a revision to the initial marginal probabilities that now recognizes the effects of the interactions among events.[36]

In another example, suppose a cybercrime scenario contained seven statements:

1. Online crime and mischief is doubling every decade
2. Hackers begin inserting false data into big databases as a means of discrediting some institutions (e.g. Internal Revenue Service (IRS) and big banks)
3. Most advanced countries begin tracking a new category for cause of death: cyber-induced death (e.g. shutdown of hospital facilities)
4. Twenty-two major UN databases have to be abandoned as a result of spurious insertions that are too difficult to identify and correct
5. Most big police forces have pre-crime detection units charged with avoiding planned terror attacks
6. Big data forensics is now used extensively to identify future criminals by forming a risk score based on sets of behavioral attributes
7. Most people say the age of "Big Brother" has arrived.

A cross-impact influence matrix was constructed using these seven statements. In Table 2.1, the first row and first column contain numbers that relate to the seven statements above. The second column shows a priori judgments about these statements that might be obtained from a review group, and the remaining columns, the influence of one element on the others. A "0" in a cell means no effect. A "3" means large positive correlation, a "-3" means a large negative correlation, and so on. Thus, an increased likelihood of item 5 (pre-crime detection) was seen to result in diminished likelihood item 1 (increasing crime) that is, if the first occurs, the second is less likely.

The Monte Carlo model works as follows: an item is picked at random. Its probability is compared to a random number that is taken to be the threshold for occurrence. If the probability is more than the random number, it is said to occur and the probabilities of the remaining items

Table 2.1 A crime cross-impact matrix

Number	Question	Probability	1	2	3	4	5	6	7
1	Online crime and mischief is doubling every decade	50	0	0	0	1	3	3	1
2	Hackers begin inserting huge amounts of false data into big databases	20	2	0	1	2	2	−2	0
3	Most advanced countries begin tracking a new category for cause of death: cyber-induced death	25	0	0	0	0	2	0	1
4	Twenty-two major UN databases have to be abandoned as a result of spurious insertions that are too difficult to identify and correct	15	0	0	0	0	1	−2	1
5	Most big police forces have pre-crime detection units charged with detecting planned terror attacks	65	−2	−1	0	−1	0	2	3
6	Big data forensics is now used extensively to identify future criminals by forming a risk score based on sets of behavioral attributes	50	−2	−1	−1	−1	1	0	3
7	Most people say the age of "Big Brother" has arrived	25	0	1	0	1	−1	−1	0

are adjusted according to the cross-impact estimates in the matrix. A second item is then chosen from those that remain, decided, and followed by adjustments to the remaining items. This process is repeated until all items have been decided. The items that "happened" constitute a single scenario. Then by repeating the process many times (e.g. 1000), one can observe the number of times that item appears and thus calculate a new probability. (Incidentally, this particular matrix results in strong increases in the probabilities of items 5 and 7.)

The interactions with an expert group can occur at three levels. First, using a RTD or other such technique, the participants can be asked to extend the list of events addressed in a scenario and to provide judgments about the initial probabilities of the events. Second, after having chosen the events and their probabilities, the participants can be asked to provide the numbers in the cells, that is, coefficients linking the events. Finally when the group recommends a change, its effect can be traced through the model: "increasing the probability of event 1 leads to an increase in the probability of event 9; increasing the probability of event 9 leads to a lower probability of events 2 and 3, etc." This kind of analysis can be done manually, or if the matrix is "played out" using Monte Carlo techniques, the process can be performed automatically.[37] If automated, the result is a new set of event probabilities and key activities that could help achieve an objective such as lowering the probability of the outcome of item 12, "Big Brother."

One of the most tedious aspects of using a cross-impact matrix to help create or modify a scenario is construction of the matrix itself. If 20 events are depicted, then the matrix will involve consideration of (20 × 20) −20 or 380 conditional probabilities for each person participating, and the number goes up with the square of the number of events included. However, the chore of filling in the matrix can be considerably simplified using RTD. RTD provides all of the advantages of RTD: asynchronous participation, no need to do it all at once, seeing the inputs of other participants, and so on. An illustrative RTD/XIA program has been written for this chapter; it uses RTD to collect judgments about initial and conditional probabilities to improve the efficiency of building a cross-impact matrix and can be seen at www.realtimedelphi.org. It is free and demonstrates a questionnaire with seven events. Sign in and, when prompted for a study code, enter "XIA" all caps and no punctuation. You can fill in the matrix if you wish. This software results in a matrix constructed by the participants that can later be used in a Monte Carlo analysis.

Futures Wheel

The Futures Wheel is a method for identifying and tracing primary, secondary, and tertiary consequences of trends, events, issues, and possible decisions and is an excellent means of generating scenario elements through interaction. It is a kind of structured brainstorming, invented by Jerome C. Glenn in 1971.[38] Variations of the Futures Wheel have been called the *Implementation Wheel, Impact Wheel, Mind Mapping,* and *Webbing.* These variations have been used by futurists in a wide variety of situations, but we recognize its use here as a means of generating, extending, and interacting with scenarios.

The Futures Wheel is a simple but powerful technique, requiring only blank paper, a pen, and one or more fertile minds, and it is also an extremely powerful method of exploring the future. The name of a trend or event is written in the middle of a piece of paper; then small spokes are drawn wheel-like from the center. Primary impacts or consequences are written at the end of each spoke. Next, the secondary impacts of each primary impact form a second ring of the wheel. This ripple effect continues until a useful picture of the implications of the central event or trend is clear.

The Futures Wheel is most commonly used to:

- Think through possible impacts of trends or events
- Create forecasts within alternative scenarios
- Show complex interrelationships
- Engage workshop participants into thinking together about the future
- Aid in group brainstorming
- Help avoid surprises

Futurists find it easy to use the wheel to think through the implications of, and organize their thoughts about, possible future events or trends. As the least expensive technique to use, it is also flexible for use in advanced situations as well as in primary school classrooms. After identifying trends or possible future events, some futurists ask their clients, "If this event occurs, then what happens next?" or "What are the impacts or consequences?" These impacts comprise a mental map of the future, acting to stimulate new thinking and help populate scenarios.

To illustrate the interactive use of a Futures Wheel in composing a scenario dealing with, say, increasing speed and diminishing size of computers, the author might convene a brainstorming session and begin the session by drawing an oval around the item on which the scenario will focus, as shown in Fig. 2.4. The author might then ask the group to suggest what necessarily "goes with" this item. As impacts or consequences are

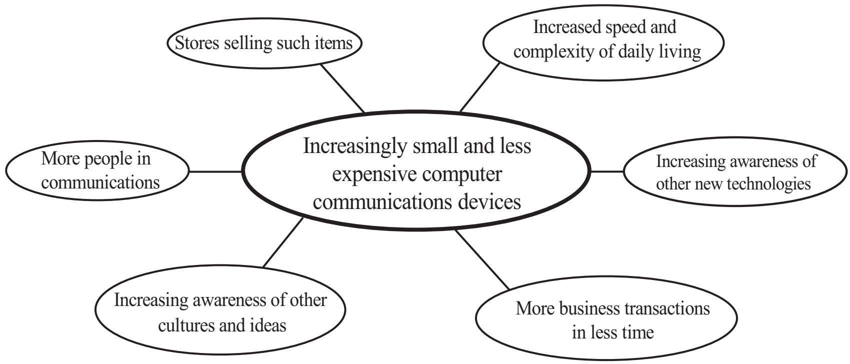

Fig. 2.4 Example of primary impacts of trend

proposed by the group, the leader draws short wheel-like spokes out from the central oval and writes these developments at the end of each spoke.

In the next step, illustrated in Fig. 2.5, ovals are drawn around each of the primary impacts. A ring can be drawn connecting the primary impacts. Next, the leader asks the group to forget about the original item in the middle of the Futures Wheel and to give the most likely impacts for each of the primary impacts of the first ring of primary consequences. As these secondary impacts are offered by the group, the leader draws two or three short spokes out from each of the ovals to form a second ring and writes the names of these secondary impacts at the end of each spoke and draws ovals around them.

At first, this process goes quickly, with participants listing second-, third-, and fourth-order consequences with little or no evaluation. After the group members agree their thinking is represented on the wheel, they can evaluate and edit the wheel to be more compliant with the main themes of the scenario under construction.

Alternatively, the impacts of an event or trend can be processed more slowly and deliberately by inviting criticism prior to entering anything on the wheel. In this approach, the group discusses the plausibility of every impact in the scenario context. If all members judge an impact to be plausible, then it is accepted as an element in the scenario. Peter Wagschal refers to this as the "rule of unanimity." He argues that making sure everyone agrees is one way of ensuring that the impacts are reasonable. He says: "The

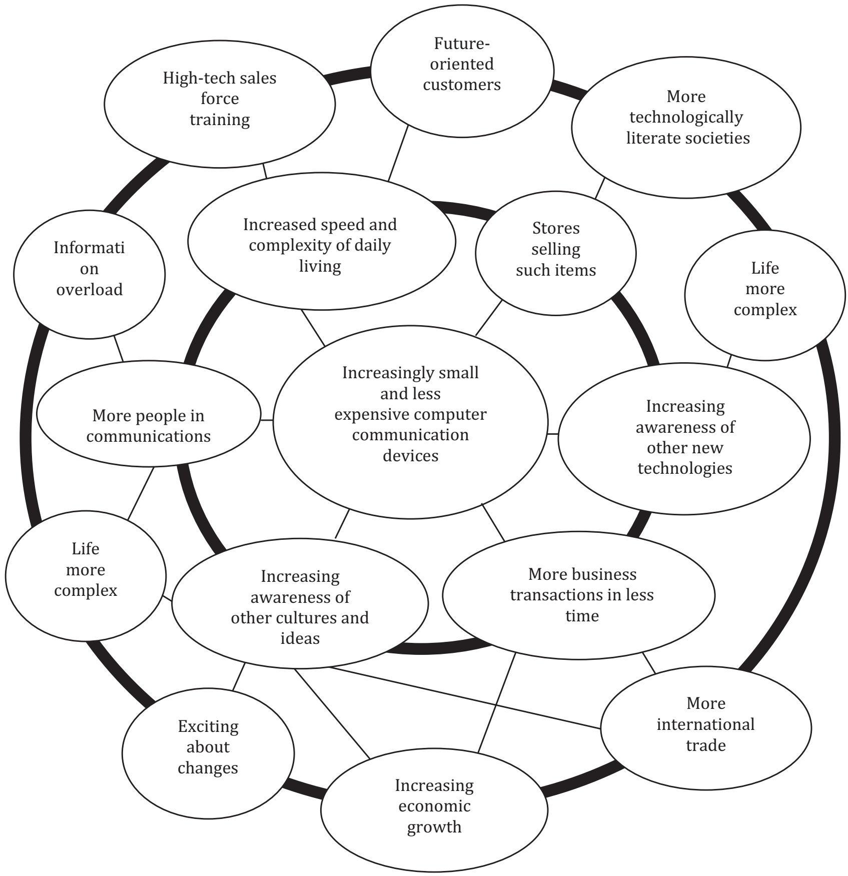

Fig. 2.5 Example of primary and secondary impacts of a trend

Futures Wheel process leads rapidly to unexpected consequences and, thus, requires a restriction on the group to prevent them from arriving at conclusions that are so speculative as to be of little worth in assessing alternative futures.[39]

The Futures Wheel can be drawn to show distinctions between primary, secondary, and tertiary linkages in another way. Instead of rings, one can draw single lines from the central oval to the primary impacts, double lines between the primary and secondary impacts, and triple lines between the secondary and tertiary impacts. Using this approach, the

Futures Wheel shown in Fig. 2.6 illustrates the possibility of cross-linkage of impacts. For example, "increased funds required for software" is a *primary* consequence of the National Security Agency (NSA) experiencing "growing costs for and dependence on acquisition and maintenance of software," a *secondary* consequence of "increased dependency on contractors," and a *tertiary* consequence of "increased costs" in general.[40]

The Futures Wheel is easy and often enjoyable to use: no equipment or software is necessary. It stimulates thinking about causes and effects of scenario elements quickly and easily. One futurist said that whenever he gets stuck in a strategic planning exercise, he does a Futures Wheel with the group and "everything starts flowing again."[41] It does not require

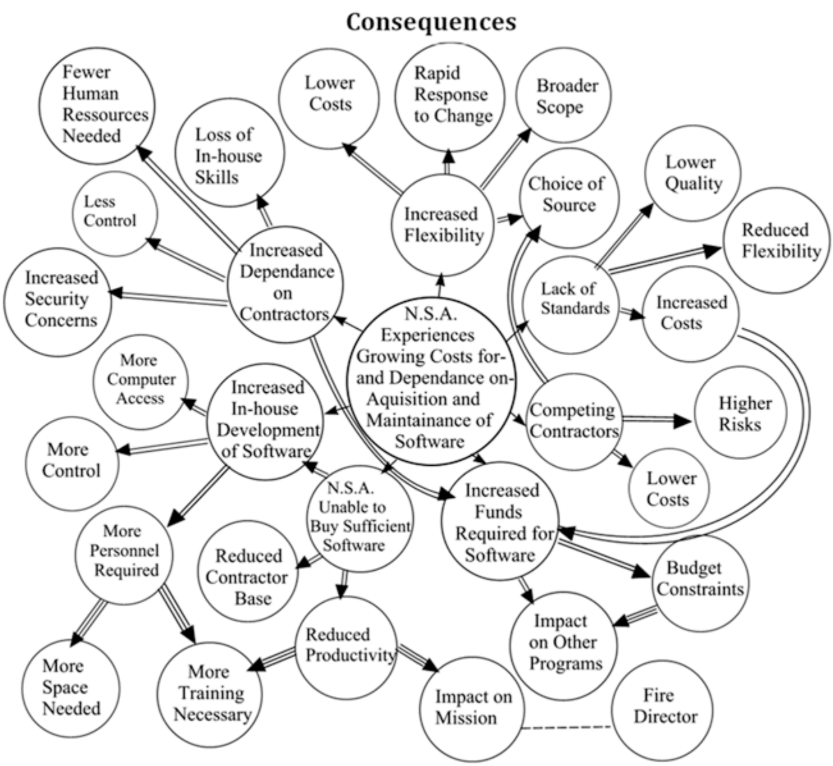

Fig. 2.6 Variation of a Futures Wheel with lines indicting sequence of consequences

advanced education or training and is easily transferred and adapted to a variety of situations. It is an easy means of diagnosing any group's collective thinking about the future.

Scenario Clustering

Researching the Interactive Dynamics Among Future Events

Imagine that 1000 forecasts were downloaded into a cross-impact matrix (see section "Cross-Impact Modeling"). Also imagine that through some tireless AI, conditional probabilities were entered in the 990,000 open cells and the matrix played through. Emerging would be a scenario resulting from the perceptions about the countless cause/effect assumptions. Now instead of placing single numbers in the cells, place a distribution of estimates in the cells. Using random numbers, choose a conditional probability in each cell. Run a solution of the matrix hundreds of thousands of times, each time resulting in a different scenario. Our guess is that out of the multitude of scenarios, there will be clusters of similar scenarios produced by different routes; that there will be stable "saddle points" or preferred outcomes, maybe several, and if that turned out to be the case, policies could be assessed against those "reference" scenarios. But the size of that task is daunting, and it is a long way off.

Robust Decision-Making

Researchers Robert Lempert, Steven Popper, and Steve Bankes of RAND have, in fact, designed and have been using an approach that they call Robust Decision Making (RDM) in analyzing policy problems under conditions of deep uncertainty. They say[42]:

> Robust decisionmaking (RDM) is a decision theoretic framework that aims to make systematic use of a large number of highly imperfect forecasts of the future[43]. [44] Rather than relying on improved point forecasts or probabilistic predictions, RDM explicitly embraces a multiplicity of plausible futures

> [usually created by repeated runs of a computer model, each run differing in its assumptions about initial conditions] as the best representation of the available information about an indeterminate future. RDM then helps analysts and decision-makers identify near-term actions which are robust across a very wide range of futures, that is, actions which promise to do a reasonable job of achieving the decision-makers' goals compared to the alternative options no matter what future comes to pass. RDM transforms the age-old question—what will the future bring?—to the more answerable—what can we do today to better shape the future to our liking?

The RDM methods have been applied to a wide variety of problems under uncertainty, including defense issues,[45] climate change policies,[46] and S&T planning.[47]

While the models used in such analyses may have been developed to predict aspects of the future in utility-based decisions, they are used in a basically different way in RDM; they are used to create a multitude of scenarios, each of which differs in input assumptions that result from the distribution of uncertainties associated with the variables and the assumed policies; but all link cause and effect to expected consequences. The output of the model runs (many thousands) is stored in a database that is later analyzed to identify policies that have moved the set of scenarios toward generally favorable outcomes. Several different models have been used in such analyses; the Ifs model of the University of Denver (see section "Using Models") has been use frequently. The runs of the model differ by systematically changing the values of input parameters over plausible ranges and ideally covering all (or at least the most important) combinations of input assumptions, including assumptions that simulate the effects of alternative policies.

Thus, RDM is a marriage between multi-scenario simulation approaches with exploratory modeling.[48]

Testing Scenarios for Resiliency

It is important to ask whether there have been omissions that, if corrected, would improve the scenarios. One approach, as noted earlier, is to distribute the drafts and ask colleagues and other readers "How can this scenario (these scenarios) be improved?" In addition, a new database containing over 1000 imagined future newspaper headlines has become available that will produce statements about random future developments. These statements are based on assumed completion of projects that are "in the works" or issues that need resolution. This database can be tapped to produce a list of "what if" possibilities that can be assessed as possible stimulants for enriching scenarios or testing them for completeness.[49]

Conclusions

We believe that in the future most scenarios will be built and tested using interactive techniques. These techniques will permit more minds to systematically focus on possible futures. There are some dangers: feedback processes can lead to instability and chaos. Further, no matter how many people are involved or complex the systems that allow them to contribute, scenarios do not (nor should they be expected to) forecast the "real" future. And mistaking a scenario as a forecast can badly mislead policy makers and other readers. Rather the best scenarios will describe plausible means for improvement and means for removing roadblocks to a desired future. The very best scenarios will also contribute to a better understanding of future risks and uncertainties. And the hope is that new techniques of elicitation and description will contribute to that goal. Imagine scenarios that are not presented as dry essays but rather electronic virtual reality worlds that—like war games—embed the reader in situations that require resolution, where the emerging story line is based on behavior in those worlds, where outcomes are probabilistic and measurable as functions of input variations, where outliers are not discarded, and where scores are kept on the costs and returns of actions taken. In short, there is a lot of room for exciting invention in this venerable field.

Notes

1. Herman Kahn and Anthony Wiener, "The Use of Scenarios," Adapted from the Year 2000: A Framework for Speculation on the Next Thirty-Three Years (Macmillan, 1967), © Hudson Institute, January 1, 1967; http://www.hudson.org/research/2214-the-use-of-scenarios; retrieved June 15, 2016.

2. Paul Raskin and Eric Kemp-Benedict, Global Environmental Outlook Scenario Framework, Background Paper for UNEP's Third Global Environmental Outlook Report, Tellus Institute, Boston, 2002
3. Hendrik Carlsen, E. Anders Eriksson, Karl Henrik Dreborg, Bengt Johansson, Örjan Bodin, "Systematic Exploration Of Scenario Spaces", *Foresight*, Vol. 18, No. 1, pp. 59–75, 2016; http://www.emeraldinsight.com/doi/pdfplus/10.1108/FS-02-2015-0011; retrieved June 24, 2016.
4. The full questionnaire can be found at: https://themp.org/rtd/future_worktechnology_2050/
5. "Factors Required for Successful Implementation of Futures Research in Decision Making," The Millennium Project, Washington, DC. See paragraph 1.4; http://www.millennium-project.org/millennium/applic-ch1.html; retrieved June 23, 2016, and quoted from "Short Term Country Scenarios" United Nations Environment Programme; see: http://www.unep.org/ieacp/iea/training/manual/module6/1248.aspx/; retrieved June 23, 2016.
6. "Learning from Experience, The Mont Fleur Scenario Exercise," March 14, 2010. http://reospartners.com/learning-from-experience-the-mont-fleur-scenario-exercise/; retrieved June 23, 2016.
7. Meadows, Donella, Dennis Meadows, Jorgen Randers, and William Behrens, The Limits to Growth, Potomac Associates Book, 1072; http://collections.dartmouth.edu/published-derivatives/meadows/pdf/meadows_ltg-001.pdf; retrieved June 25, 2016.
8. Hughes, Barry, Series Editor, "Patterns Of Potential Human Progress," The Pardee Center, Oxford University Press, 2010.
9. Mapping the Global Future, Report of the National Intelligence Council's 2020 Project, December, 2004.
10. Hughes, Advancing Global Education, Patterns of Potential Human Progress (2010).
11. For a good discussion of demographic modeling, see: Newell, Colin, "Methods and Models in Demography," The Guilford Press, New York, 1988.
12. The game was patented by Gordon, Helmer, and Goldschmidt (of Kaiser Aluminum and Chemical); patent 3,473, 802.
13. Gordon, Theodore, Hayward, "Initial Experiments with the Cross-Impact Matrix Method of Forecasting," *Futures*, Vol. 1, No. 2, pp. 100–116, 1968. This was the original cross impact paper and introduced the method.

14. "Monte Carlo" is the name of a technique often used in operations research when modeling in a closed form is not feasible. Instead specific values of independent variables are selected (usually by random processes), and the value of the dependent variable is computed. This is known as a single "run." Then new values are assigned to the independent variables, and the process is repeated numerous times, perhaps thousands of times, in order to produce a range of results for the independent variable.

15. Gordon, Theodore; Rochberg, Richard; Enzer, Selwyn, "Research on Cross Impact Techniques with Selected Problems in Economics, Political Science and Technology Assessment," Institute for the Future, 1970.
16. Dalkey, Norman, "An Elementary Cross Impact Model," *Technological Forecasting and Social Change,* Vol.3, No.3, pp. 341–351, 1972.
17. Helmer, Olaf, "Cross Impact Gaming," *Futures,* Vol.4, No.2, pp. 149–167, 1972.
18. Heiko, D. (2007). Causal cross-impact analysis as a gaming tool for strategic decision making. Informally published manuscript, BIBA—Bremen Institute of Industrial Technology and Applied Work Science, Retrieved from http://games.biba.uni-bremen.de/info/2007-07.pdf; retrieved June 24, 2016.
19. Kane, J, A Primer for a New Cross Impact Language- KSIM. *Technological Forecasting and Social Change*, Vol. 4, No. 2, pp. 129–142, 1972.
20. Duval, Fontela, and A. Gabus, "Cross Impact: A Handbook of Concepts and Applications," Geneva: Battelle-Geneva, 1974.
21. Duperrin, J.C., Godet, M, "SMIC-74, A method for constructing and ranking scenarios," *Futures*, Vol. 7, No. 4, pp. 302–312, August 1975.
22. Stover, J. (1975). The Use of Probabilistic System Dynamics in Analysis of National Development Policies: A Study of Economic Growth and Income Distribution in Uruguay. Proceedings of the 1975 Summer Computer Simulation Conference, San Francisco, California.
23. Enzer, Selwyn, "Interax: An Interactive Model for Studying Future Business Environment," *Technological Forecasting and Social Change*, Vols.2 and 3, 1980.
24. Ducos, G., "Delphi et Analyse d'Interactions," *Futuribles*, No. 71, 1984.
25. Bonnicksen, Thomas and Robert H. Becker, "Environmental Impact Studies: An Interdisciplinary Approach for Assigning Priorities," *Journal: Environmental Managemen*t, Vol.7, No. 2, March 1983.
26. Kane, Julius, "A Primer for a New Cross Impact Language- KSIM," *Technological Forecasting and Social Change,* Vol. 4, No. 2, pp. 129–142, 1972.

27. Godet, Michel, *From Anticipation to Action: A Handbook of Strategic Prospective*, UNESCO Publishing, 1993.
28. Vickers, Brent, "Using GDSS to Examine the Future European Automobile Industry," *Futures*, October. 1992. See Section V for an abstract of this report.
29. Schuler, Albert, William A Thompson, Ilan Vertinsky, Yishi Ziv, "Cross-Impact Analysis of Technological Innovation and Development in the Softwood Lumber Industry in Canada: A Structural Modeling Approach," *IEEE Transactions on Engineering Management,* August, 1991.
30. Fontela, Emilio and José M. Rueda-Cantuche, "Linking cross-impact probabilistic scenarios to input-output models," June 2004 http://www.uam.es/otroscentros/klein/stone/fiirs/cuadernos/pdf/FIIRS012.PDF; retrieved June 24, 2016
31. Winterscheid, A "Flood risk analysis using Cross-Impact Matrix," Darmstadt University of Technology, Germany, Geophysical Research Abstracts, Vol. 9, 02741, 2007.
32. Warfield, J. N. 1974a. Structuring complex systems, Battelle Monograph No 4, Battelle Memorial Institute, Columbus. Ohio, USA.
33. Victor A. Banus, Murray Turoff, Roxanne Hiltz, "Supporting Collaborative Scenario Analysis Through Cross-Impact," ISCRAM 2012, http://www.iscramlive.org/ISCRAM2012/proceedings/196.pdf; retrieved June 25, 2016.
34. Moritz, Frank, "Cross Impact Analysis and Forecasting the Future of Rhodesia," *Quantitative Approaches to Political Intelligence: The CIA Experience,* ed. Richard J. Heuer, Jr., Westview Press, 1978.
35. Whaley, Barton, "Textbook of Political-Military Counter Deception; Basic Principles and Methods." August 2007; at URL: 20131119_TextbookofPolitical_MilitaryCounterdeception.pdf; retrieved June 21, 2016.
36. See Enzer, op. cit. and MI Mphahlele, OO Olugbara, SO Ojo and DG Kourie, "Cross-Impact Analysis Experimentation Using Two Techniques to Revise Marginal Probabilities of Interdependent Events," *Orion*, Vol. 27, No. 1, pp. 1–15 (2011), http://orion.journals.ac.za/pub/article/viewFile/95/96; retrieved June 24, 2016.
37. For a more complete discussion of this quantitative cross-impact approach, see: T. J. Gordon, Cross-Impact Analysis, a chapter in the CD Futures Research Methods version 3.0, The Millennium Project, Washington, D.C.

38. Full details on this method can be found in the CD: Futures Research Methods version 3.0, The Millennium Project, Washington, D.C.
39. Wagschal, Peter. "Futuring: A Process for Exploring Detailed Alternative Futures," *World Future Society Bulletin* (now the *Futures Research Quarterly*), September/October 1981, pp. 25–31.
40. This Futures Wheel, which was developed by Futurist David Snyder during consulting with the US National Security Agency, illustrates the use of single, double, and triple lines to represent primary, secondary, and tertiary impacts (reprinted with permission of the author).
41. Interview by Glenn with futurist David Snyder.
42. Lempert, R. J., S. W. Popper, and S. C. Bankes, *Futures Research Methodology, Robust Decision Making v3*, CD ROM , Millennium Project, Washington, DC.
43. Lempert, R. J., S. W. Popper, and S. C. Bankes, *Shaping the Next One Hundred Years: New Methods for Quantitative, Long-Term Policy Analysis*, Santa Monica, Calif.: RAND, MR-1626-CR. 2003.
44. Lempert, Robert J. and Myles T. Collins, "Managing the Risk of Uncertain Threshold Response: Comparison of Robust, Optimum, and Precautionary Approaches," *Risk Analysis*, Vol. 27, No. 4, 1009–1026, 2007.
45. Brooks, Arthur, Steve Bankes, Bart Bennett, "An Application of Exploratory Analysis: The Weapon Mix Problem," *Military Operations Research*, Vol. 4, No. 1, pp. 67–80, 1999.
46. Lempert, Robert J., Nebojsa Nakicenovic, Daniel Sarewitz, Michael Schlesinger, "Characterizing Climate-Change Uncertainties for Decision-makers," *Climatic Change*, Vol. 65, pp. 1–9, 2004.
47. Lempert, Robert and Jim Bonomo, 1998: *New Tools for Robust Science and Technology Planning*, RAND DB-238-DARPA.
48. Steven C. Bankes, Robert J. Lempert, and Steven W. Popper, "Computer-Assisted Reasoning," *Computing in Science and Industry*, Vol. 3, No. 2, pp. 71–77 March/April 2001.
49. Theodore Gordon, "1,000 Futures: Testing Resiliency Using Plausible Future Headlines," *World Future Review*, Vol. 8, No. 2, pp. 75–86, 2016. Free access to the database is via www.changesignals.com. The access code is "whatif."

3

Virtual Reality for Marketing Research

Raymond R. Burke

Academic and commercial interest in understanding shopper behavior and the influence of the store environment has increased dramatically over the last 30 years due to a number of factors. Retailers have become more powerful through consolidation and the aggressive development of private label programs; traditional brand advertising has lost effectiveness due to mass media fragmentation; increased competition has led to product proliferation and shortened product lifecycles; and shoppers have become more connected and informed than ever before (Desforges and Anthony 2013).

Within this fast-changing environment, conventional marketing approaches are not meeting expectations. Consumer packaged goods manufacturers have introduced a continuous stream of new products in the hopes of engaging shoppers and growing brand and category sales, but the results have been disappointing. According to Nielsen (2016), the

R.R. Burke (✉)
Indiana University, Kelley School of Business, Bloomington, IN, USA

L. Moutinho, M. Sokele (eds.), *Innovative Research Methodologies in Management*,
DOI 10.1007/978-3-319-64400-4_3

top 25 largest food and beverage companies generated 45% of category sales in the USA in 2015, but drove only 3% of the total category growth from 2011 to 2015 (roughly $1 billion in sales out of $35 billion in category growth) and grew at a compound annual growth rate of 0.1%. In contrast, private label drove 23% of the growth ($8 billion) and grew at a 2.6% compound annual growth rate. Investments in trade promotion have also been relatively unproductive. Of the estimated $280 billion that manufacturers spend on in-store promotions, 70% (or $196 billion) lost money for the manufacturer, with an average return on investment (ROI) of 30 cents on a dollar (Desforges and Anthony 2013).

To effectively manage their brands in this complex environment, marketers need research tools to identify the points of engagement and friction in the shopping process and discover what drives shopper behavior at the point of purchase. They need to understand the "First Moment of Truth," a phrase coined by Alan Lafley, former president and CEO of Procter & Gamble. Virtual reality simulations are one such tool, offering several advantages over conventional research techniques for shopper research. They can be used to create photorealistic simulations of the retail environment, conduct marketing experiments, and collect detailed behavioral data without requiring a retailer's cooperation. Multiple concepts can be tested in a controlled and confidential setting, at a lower cost and faster than an in-market test. Practitioners report that virtual store simulations offer a variety of business benefits, including the facilitation of internal planning and collaboration, and the fostering of stronger relationships with key retail accounts (Breen 2009).

The next section outlines the challenges to developing innovative marketing concepts, the limitations of existing research techniques, and the benefits of virtual reality simulations. The following sections describe the steps involved in creating and running a simulated shopping study, discuss the validity of the simulation technique, provide examples of several commercial and academic research applications, and summarize the future prospects for using the virtual store for marketing research and other business applications.

The Innovation Challenge

Innovation is critical for many business decisions, including the design of new products, product-line planning, pricing, promotion, package design, and merchandising. The importance of innovation is perhaps easiest to see in the context of new product introductions (Jelinek and Schoonhoven 1993). Of the tens of thousands of new consumer packaged goods that are introduced each year, only a small fraction deliver genuine innovations in formulation, positioning, packaging, or technology (Erickson 1995; Nielsen 2016). Without innovation, products fail to generate consumer interest and excitement. More money must therefore be spent on advertising and promotion to communicate a brand's marginal benefits and stimulate trial. Consumers switch between brands because they perceive them as commodities. High marketing expenses and low sales lead to the eventual withdrawal of these products from the market. Less than 3% of new consumer packaged goods exceed first-year sales of $50 million, considered the benchmark of a successful product launch (Schneider and Hall 2011).

So why is there so little innovation? A critical problem is that the research methods and criteria used to screen new marketing ideas are inherently biased toward maintaining the status quo. These methods include informal and formal reviews by individuals or groups of managers, discussions with salespeople, and personal interviews and focus groups with customers.

To illustrate the problem, let us consider the new product example. One can think of the new product development process as a funnel (Wheelwright and Clark 1992). Firms often start out with hundreds, and in some cases thousands, of new product ideas. As the project evolves, the company eliminates most of these ideas and commits to specific concepts, product designs, and manufacturing processes; that is, the funnel narrows. Because of the high costs of producing and launching a new product, only a few of the original concepts make it through the entire process. Unfortunately, the criteria and methods used at the early stages to screen new product ideas tend to underestimate the value of innovation.

- *Hard data drive out soft data*. It is easier for a firm to assess a new product's technical feasibility and fit with the company's resources than to predict potential customer demand. Really new products are often more difficult and expensive to manufacture, and require the greatest change in the organization's structure and activities (Tushman 1994). With no hard evidence to the contrary, managers will often refuse to consider developing products outside of their original, successful product lines.
- *Really new concepts are hard to imagine*. At the early stages of the development cycle, consumers are often asked to react to written or verbal descriptions of concepts, perhaps with an artist's rendering of the product. People are comfortable evaluating concepts that are similar to existing products, but they have trouble envisioning the benefits of really new products (Fusco 1994).
- *Concepts are not tested in context*. Research has shown that the most important new product success factor is having a unique, superior benefit in the eyes of the customer (Cooper and Kleinschmidt 1987). People judge new products relative to existing offerings. It is therefore surprising that most concept-screening tasks ask customers to evaluate new product ideas in the abstract. The benefits of a truly unique product are therefore underestimated.
- *Concepts are not tested under dynamic market conditions*. One of the most common reasons for new product failure is unexpected competitive retaliation (Wind and Mahajan 1988). Yet, most research methods do not evaluate the performance of new concepts in alternative competitive scenarios. Marginally improved products may be preferred over existing products under current conditions, but fail under likely, future scenarios of lower prices and increased competition.
- *Consumer acceptance depends on execution*. The execution of a new marketing idea is often not tested until late in the development process. Yet, an innovative execution can be a core part of the product concept. For example, a new fat-free frozen entree could be positioned as a diet product for weight watchers or as a meal for active, health-conscious adults. This positioning can have a dramatic impact on the potential market size and level of demand for the product.

In summary, the research techniques that are routinely used to screen new marketing ideas do not have sufficient realism in terms of the concept presentation, the competitive context, and the measurement of customer behavior to adequately assess the value of innovation.

Why aren't more realistic and valid research methods employed? The answer is time, cost, control, and confidentiality. Traditional test markets have high external validity, but they are slow and expensive, often taking six months to a year or more to complete, and can cost millions of dollars. Controlled field experiments—where the manufacturer handles the distribution and display of the new product in test stores, and maintains conditions in a matched set of control sets—are faster and less expensive, but can still take several months to conduct and cost over $100,000 for a single test. Three critical limitations of both techniques are the necessity of obtaining the retailer's cooperation (which is becoming increasingly difficult to obtain); the difficulty in calculating a baseline level of sales against which to assess the effects of the marketing manipulation; and the possibility that the test will alert competitors to a manufacturer's activities (Burke 1996).

Virtual Shopping

Virtual reality simulations allow the marketer to recreate the look and feel of a retail store on a computer monitor, projected display wall, VR CAVE, or head-mounted display. In a typical desktop VR simulation, the shopper navigates through the store using a mouse, joystick, or keyboard commands, and can pick up a package from the shelf by clicking on or touching its image on the monitor. The product then flies to the center of the screen. The consumer can rotate the product to examine package information and then touch an image of a shopping cart to "purchase" the item. During the shopping process, the computer unobtrusively records all aspects of the interaction, including the time spent shopping in the category, the sequence and duration of each product interaction, and the quantity of items purchased.

The sophistication and realism of VR simulations has significantly improved during the last 30 years; evolving from simple two-dimensional (2D) simulations developed in the late 1980s, to interactive simulations with three-dimensional (3D) graphics and product interaction in the 1990s, to today's fully immersive simulations using life-size projected images and stereo glasses, as illustrated in Fig. 3.1. The most recent innovations include integrating eye tracking into the simulation to measure patterns of visual attention (Burke and Leykin 2014), and using biometric, EEG, and facial coding measures to capture emotional response.

A. Ad testing simulation presented at the 1988 AMA Doctoral Consortium, UC Berkeley (Burke 1988)

B. "Realistic Laboratory Simulation" from Burke, Harlam, Kahn and Lodish (1992)

C. Planogram-based 3D shopping simulation from Burke (1996)

D. Contemporary VR simulations at Indiana University's Customer Interface Laboratory and Advanced Visualization Laboratory (Bloomington, IN)

Fig. 3.1 Evolution of Virtual Shopping Simulations

The computer-simulated environment offers several advantages over traditional marketing research tools. It allows firms to test new marketing concepts in a realistic, competitive context without incurring the costs of physical production and distribution. A broad range of concepts can be tested early in the development cycle, without fear of disclosing new ideas to competitors. The performance of new concepts can be evaluated under a variety of competitive scenarios. Once the virtual store is constructed, individual tests can be set up and run in days rather than months or years. Data are coded automatically by the computer, so the results are available quickly. The simulation provides quantitative estimates of sales, market share, sources of volume, and profitability that relate directly to the firms' objectives.

Creating and Running a Virtual Shopping Simulation

Most of the digital assets necessary to build a photorealistic model of a shelf fixture or store aisle are readily available. Commercial shelf-space management systems record product dimensions, shelf locations, and images of the product packages; store floor plans contain the dimensions and locations of shelf fixtures and display cases; and point-of-sale systems track product inventory and prices. Research, consulting, and service firms have assembled this information into comprehensive databases for the grocery, drug, and general merchandise trades, which can be used to construct 2D or 3D models of the products, shelf displays, and store environments for marketing research. Digital images of new product concepts, package designs, merchandising, and promotional materials can be incorporated directly into a shopping simulation, allowing marketers to test alternative scenarios.

There are some business contexts where the photographic and dimensional information necessary to build a simulation are not available. In these cases, one can photograph the environment and texture-map the images onto the corresponding architectural models for 3D walkthroughs, or use panoramic video as a backdrop for computer-generated products and displays. The simulated shopping experience can be augmented with mock salesperson interactions, and exposure to television programs, magazines, web pages, product brochures, and so on (see, e.g., Urban et al. 1996).

Conducting a simulation study typically involves the following steps: (1) pre-recruit respondents to a central research facility or a web portal; (2) present a cover story (e.g., "you will be testing a new type of home shopping system…"), and instruct respondents to shop as they normally would; (3) train respondents on the VR simulation's navigation and product selection process; (4) take respondents on a series of "shopping trips" where they can interact with and purchase products in several categories; and (5) administer a post-simulation survey, and debrief and pay the respondents. During the course of the study, the researcher can manipulate the appearance and organization of the shelf display (e.g., product assortment, adjacencies, signage) and measure the effects on a variety of dependent variables, including product category dwell time, product pickup/return order, package examination time, SKU sales, brand switching, and category sales volume.

The sequence of trips plays an important role in creating a realistic grocery shopping experience. The "practice trials" allow shoppers to become comfortable with the computer interface and familiar with the layout of the virtual store environment and shelf displays. By the third or fourth shopping trip, respondents transition into a more typical, low-involvement mindset that is much closer to what is observed in field studies of supermarket customers (Dickson and Sawyer 1990; Sorensen 2009). It is at this stage that the test concepts are introduced into the virtual store. To better approximate the actual retail environment, the simulation can be tailored to the shopping habits of individual consumers. For example, consumers who normally purchase health and beauty products from a drug store can be taken through a drug store simulation, while individuals who buy these products in a grocery store can shop in a grocery simulation.

Validation

Validation is a critical issue if the virtual shopping simulation will be used to screen new product concepts and forecast consumer acceptance. The results from several academic and commercial validation studies suggest that simulations can accurately predict shoppers' in-store behavior (Bressoud 2013; Burke 1996; Burke et al. 1992; Campo et al. 1999;

Desmet et al. 2013; Tjøstheim and Saether-Larsen 2005; van Herpen et al. 2016).

In an early test of a computer graphic simulation, Burke et al. (1992) recruited a small panel of consumers; tracked their grocery shopping activities over a seven-month period in five categories (soft drinks, paper towels, canned tuna, orange juice, and toilet paper); and then asked them to take a series of trips through "realistic" and "rudimentary" shopping simulations. For the realistic simulation, the computer screen showed high-resolution color images of the shelf displays, with the actual prices and signage that had appeared in the physical stores (see Fig. 3.1, panel B). The findings demonstrated that computer simulations are generally accurate at predicting brand market shares and consumer price sensitivity observed in the supermarket, especially when the simulation reproduces the visual cues that consumers use to make their purchase decisions. However, there are a few notable differences in behavior: shoppers tend to purchase larger quantities in the simulation than in the physical store, exhibit less brand switching across shopping trips, and (in the rudimentary simulation) are more responsive to price features.

A follow-up study tested new package designs for products in the cleaning, health and beauty care (HBC), and snack food categories using both traditional field experiments and virtual shopping simulations (Burke 1996). The measured effects of the package changes were similar across the two methods: 27% and 31% lifts in HBC, 26% and 23% drops in snack foods, and no change for cleaning products, for the laboratory and field experiments, respectively. The computer simulation demonstrated strong predictive validity, with estimated market shares closely matching the results from syndicated scanner data. The correlations ranged from 0.90 for HBC products to 0.94 for household cleaning products.

Campo et al. (1999) conducted a between-subjects validation study in a grocery context using carefully matched samples and store conditions, and explored a number of potential biases that might affect the validity of simulation findings. They compared shoppers' patterns of purchasing jam, paper towels, and margarine in a computer simulation over a 12-week period, with shopper behavior captured by UPC scanners and customer loyalty cards over 16 weeks. The authors found that simulation shoppers tend to buy in a similar, realistic manner across shopping occasions, but they purchase larger quantities in the simulation than in the physical store, favor national brands over generics, and respond more to promotions (especially for national brands).

Tjøstheim and Saether-Larsen (2005) conducted a validation study of a 3D grocery shopping simulation based on a video game engine, comparing market share estimates from the virtual shopping application with in-store purchases. A sample of 603 customers were intercepted as they entered a grocery store and asked to (1) take a virtual shopping trip, purchasing what they intended to buy, and then (2) complete their regular shopping in the physical store. For the 36 beverage brands tested, the authors found no significant differences in market share for two-thirds of the products.

Desmet et al. (2013) compared how customers shopped for ground arabica coffee in a 3D shopping simulation versus a physical laboratory store, where the research protocols and product displays were matched to be as similar as possible. They report that the article purchase rates observed in the two simulated shopping environments were highly correlated (Pearson $r = 0.81$), and the average difference between purchase rates was small (mean = 0.04%; standard deviation = 0.028). However, shoppers in the 3D simulation took twice as long to buy, were more likely to handle products, explored less, and paid less attention to price. (The findings from the attitudinal measures were similar across the two methodologies.) Similarly, Bressoud (2013) found that shoppers took longer to buy a new cereal and purchased more in a virtual environment than in an experimental real store, but these differences were not statistically significant.

A recent study by van Herpen et al. (2016) compared shopper behavior in a real supermarket with a 3D virtual shopping simulation and a 2D pictorial representation, using matched product assortments and shelf arrangements. Shoppers were pre-recruited for the study, qualified on category usage, and asked to shop for three types of products (milk, produce, and biscuits) in either the 3D simulation, pictorial simulation, or physical store. The authors found that shoppers in the simulation conditions tended to buy more products, more varied products, more national brands, and more promoted products than in the real store, echoing the findings of several prior studies. The 3D simulation was a better predictor of milk purchases than the 2D simulation, but the simulations were comparable for produce and biscuits. The authors did not compare estimated market shares of products across conditions.

Applying the Virtual Store

During the last 20 years, virtual shopping simulations have grown in popularity for both academic and commercial research. According to a report by the In-Store Marketing Institute (Breen 2009), consumer packaged goods manufacturers and retailers use simulations to test the impact of new product packaging, pricing strategies, product assortments, and shelf organization on SKU- and category-level purchases. Some companies have also tested the effectiveness of in-store marketing and merchandising concepts, examined category adjacencies, and evaluated overall store layout. Academic applications include experiments on product package design (e.g., Garber et al. 2000; Underwood et al. 2001), shelf organization, and store layout (Burke and Leykin 2014).

Beginning in 1993, the present author collaborated with an Atlanta-based research firm (Simulation Research) on a series of studies using the virtual shopping software. During a 9-year period, 238 commercial studies were completed worldwide, covering a wide range of marketing applications. The most popular application was testing new product concepts (30%), followed by planogram assortment and shelf organization studies (23%), product packaging tests (21%), pricing experiments (16%), promotion tests (6%), and restaging/substitution studies (4%). Most studies were conducted with grocery/food products, but there were also many tests with household items, over-the-counter drug products, and personal care items.

The following examples illustrate some of the issues that were investigated (Feder 1997). General Mills benchmarked the virtual shopping simulation against traditional concept testing and found that the shelf simulation provided unique insights into the shelf impact and consumer appeal of new product concepts including Honey Nut Chex and Chocolate Cheerios cereals, which were successfully launched by the company. Using a simulation of the snack food aisle in Spanish markets, Frito-Lay discovered that end-of-aisle displays produce more lift than advertising directly above the shelf. The company also adapted the simulation to create a virtual vending machine and measure how likely buyers of salty snacks would be to substitute candy bars, cookies, crackers, or fruit, revealing category structure. Johnson & Johnson's advanced care products division used simulations to determine how to price Uristat, a pain killer used for urinary diseases. Burke (1996) and Needel (1998) discuss a number of other commercial examples, including measuring the brand equity of Goodyear tires, optimizing the shelf organization of frozen foods, simplifying fast food menus, managing product assortments, and SKU rationalization.

Shaking Up Consumers at the Point of Purchase

For a more detailed look at the application of virtual shopping, let's return to the problem of marketing innovation. As noted earlier, manufacturers face the difficult task of designing marketing programs to break through the high levels of in-store clutter and attract consumers' attention. But which marketing tools are best able to grab and hold the consumer at the point of purchase? Three simulation studies were designed to address this question in the context of consumer packaged goods. The first study measured the effects of three alternatives: (1) introducing a new product, (2) changing an existing product's package, or (3) running a price promotion (a 30% reduction with a "Sale" sign), on consumer attention, brand and category sales, and brand switching.

One hundred grocery shoppers, aged 18–65, were recruited to participate in this study. Respondents were asked to take a series of seven shopping trips through the simulated store. On each trip, they purchased products from four categories: soft drinks, paper towels, canned tuna, and orange juice. These categories were selected because they had high penetration and high purchase frequency among the sampled consumers. On the fifth and subsequent shopping trips, a change was introduced into three of the four product categories, with the fourth category serving as a control condition. For example, for some respondents, a new product was introduced into soft drinks, a new package was shown in paper

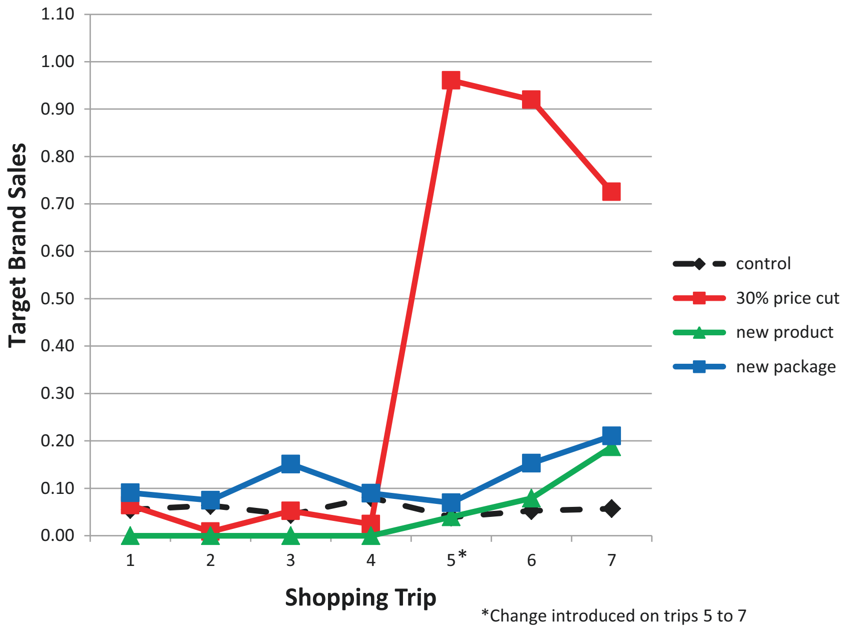

Fig. 3.2 Shaking up consumers at the point of purchase

towels, a 30% price reduction was taken in canned tuna, and no change was made in the orange juice category. A single respondent saw only one experimental condition in each product category, but the pairing of conditions and categories was balanced across respondents using a confounded block experimental design (Kirk 1968, pp. 327–339; Winer 1971, pp. 639–650). On each trip, shoppers were told to assume that their household inventory of the product was depleted.

The study revealed that the price promotion had a dramatic impact on choice, increasing brand sales up to 20 times the level observed during nonpromoted periods (Fig. 3.2). This effect was greatest in the paper towel category, where brands were perceived to be of similar quality, and smallest in the orange juice category, where the perishable nature of the product limited the quantity purchased. An examination of the total category sales and brand switching data revealed that most of the sales increase during the price promotion was due to stockpiling by the brand's current customers, rather than brand switching by new customers. Thus, the price promotion would reduce brand sales in future nonpromoted periods as consumers worked down their at-home inventory. The effects of the new product introductions and the new packaging were more gradual, with sales rising during the fifth, sixth, and seventh shopping trips. Surprisingly, the package change produced about the same level of brand switching as the price promotion, but without the undesirable stockpiling effects and consequent loss of margin. On average, consumers' attention to the target brands was about 12.5 seconds in the experimental conditions, and only 7.5 seconds in the control condition.

One might argue that the large price effect observed in the first study was due to its novelty in the laboratory context. When shopping at a supermarket, consumers see a variety of products on sale. If the same promotion is seen repeatedly, shoppers may "habituate" to this stimulus, responding less over time. Therefore, a second study was designed to examine how consumers respond to repeated price promotions by competing brands. As in the previous study, participants were asked to take seven shopping trips through each of four product categories. There were four experimental conditions, where:

1. The target brand was promoted on the second, fourth, and sixth shopping trips,
2. A competitor's brand was promoted on the second and fourth trips, while the target brand was promoted during the sixth trip,
3. The target brand was promoted only once, during the sixth trip, or
4. None of the brands was promoted (control).

As in the first study, each respondent saw only one condition in each product category, but the pairing of conditions and categories was balanced across respondents.

The research revealed that, contrary to the habituation hypothesis, consumer response to the target brand's price promotion increased with repetition (condition #1). However, consumers gradually stopped buying this brand during nonpromoted periods (Fig. 3.3). It appears that they learned to wait for the brand to go on sale. Consumers had an even greater response to the target brand's promotion when the leading competitor had been frequently promoted in the past (condition #2). The promotion of the

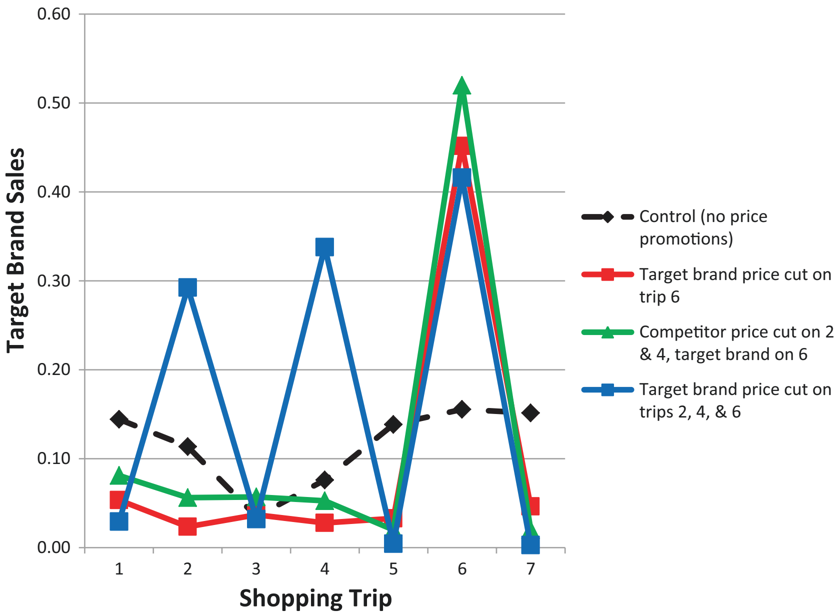

Fig. 3.3 Measuring the impact of repeated price promotions

competing brand seemed to increase the distinctiveness of the target brand's promotion, thereby increasing its impact. In all cases, consumers were able to quickly respond to the price promotions, taking only about 1.5 seconds longer to make a decision than in the no-promotion conditions.

While consumers eagerly respond to price promotions, they are slow to react to new product introductions. Therefore, a third study was conducted to see if we can accelerate consumer trial of new products by pairing them with attention-getting merchandising and promotions. The research was conducted with a sample of 96 people who were the primary grocery shoppers in their households. Participants were asked to take seven shopping trips through the simulated store, making purchases in four different product categories. A new product was introduced into each category on the fifth and following shopping trips. This product was accompanied (on just the fifth trip) with either no special merchandising, a "NEW!" sign, a 30% price promotion (with a "SALE" sign), or twice the amount of shelf space. To increase the realism of the task, half of the respondents also had the opportunity to purchase a leading competitor's product at a 30% price reduction. This is a common, competitive practice, designed to disrupt a manufacturer's new product introductions and help the competitor to retain its existing customers. While prior research has explored the judgmental effects of competitive products and actions (e.g., Huber et al. 1982; Huber and Puto 1983; Simonson 1989; Simonson and Tversky 1992), the perceptual and attentional influences have been largely neglected.

The promotional activities of competitors had a significant influence on consumer response. In terms of brand attention, consumers spent about the same amount of time (11 seconds) examining the new product in each of the four merchandising/promotion conditions when there was no competitive price reduction. However, when the competitor cut its price, consumer attention to the unmerchandised new product dropped to 5 seconds. With additional shelf space or a "NEW!" sign, brand viewing time was 10 seconds. When the new product was put on promotion, attention peaked at 16 seconds. Trial of the new product was highest in the price promotion condition and occurred immediately after the product was introduced. In other conditions, the purchase rate was lower, with the "NEW!" sign accelerating the purchase rate relative to the control and high-shelf-space conditions. The results suggest that new products can get lost in the clutter at the point of purchase. Merchandising and promotions may enhance consumer attention, especially in the context of competitive promotions.

Limitations of Virtual Shopping

The virtual shopping methodology has several weaknesses that are important to keep in mind. When making a purchase decision, consumers use certain cues to select between products. In some cases, the brand name and packaging are the critical stimuli. In others, price, shelf space, signage, smell, weight, taste, and/or texture might be important (e.g., Madzharov et al. 2015; Peck and Childers 2003). For the simulation to produce a valid measure of consumer behavior, it must accurately recreate these cues in the virtual store. It is relatively easy to generate a credible visual and auditory simulated environment, but the senses of touch, taste, and smell are harder to simulate. The absence of tactile feedback is likely to curtail the use of this approach for products where weight, firmness, and texture are important (e.g., produce, soft goods), and for complex products, where trial is critical to product evaluation and purchase. Researchers may be able to address this concern by using a hybrid approach where the actual product is demonstrated and/or sampled as part of the virtual shopping exercise.

A second limitation of most simulations is that they do not include a social component. The virtual stores are deserted, with no visible salespeople or shoppers. Social factors have been found to influence how shoppers navigate a store environment, and what they touch and purchase (Zhang et al. 2014). This is a greater concern for store-level simulations, where factors such as crowding, customer service, and checkout line length are more likely to play a role in the shopping process. As technology evolves, simulations may incorporate photorealistic, dynamic avatars who populate the store environments and are programmed to interact with shoppers. In the meantime, video interactions can be incorporated as an element of the simulation (see, e.g., Urban et al. 1996).

Conclusion

The ongoing development and refinement of virtual reality simulations will continue to enhance their performance and expand the range of applications. These include identifying which new marketing concepts stand out from the competitive clutter, estimating the relative contributions of product design, branding, packaging, pricing, and promotion to product sales, and measuring the impact of possible competitive retaliation. By using the simulation at the early stages of the new product development process, firms can test a broad selection of marketing ideas in a realistic setting, and thus improve their chances of identifying breakthrough concepts that will expand their businesses.

Perhaps the most exciting future use of the simulation is in the area of ecommerce. With improvements in computer graphics, immersive and augmented display technologies, mobile computing, network connectivity, and product fulfillment, the virtual store is evolving from a research tool to a viable channel of distribution. Virtual stores can combine the convenience and transparency of today's online shopping environments with the emotional engagement and reward of a video game or movie. They can carry an unlimited variety of products, styles, flavors, and sizes, shown in entertaining and relevant contexts. New products can be stocked which do not yet physically exist, but are produced in response to the desires of shoppers. The store can be tailored to the preferences and purchasing habits of individual customers, displaying just those products with the features, price, and performance they require. Simulations can also bring these benefits into physical store environments, creating engaging and responsive hybrid experiences that effectively connect supply and demand.

References

Breen, P. (2009). *Shaping Retail: The Use of Virtual Store Simulations in Marketing Research and Beyond*. In-Store Marketing Institute White Paper, available online at http://kelley.iu.edu/cerr/files/09ismi_virtualretailing.pdf

Bressoud, E. (2013). Testing FMCG Innovations: Experimental Real Store Versus Virtual. *Journal of Product & Brand Management, 22*(4), 286–292.

Burke, R. R. (1988). *Computer Graphic Simulations for Marketing Research*. Faculty presentation at the American Marketing Association Doctoral Consortium, University of California Berkeley, Berkeley.

Burke, R. R. (1996). Virtual Shopping: Breakthrough in Marketing Research. *Harvard Business Review, 74*(March–April), 120–131.

Burke, R. R., & Leykin, A. (2014). Identifying the Drivers of Shopper Attention, Engagement, and Purchase. In D. Grewal, A. L. Roggeveen, & J. Nordfält (Eds.), *Review of Marketing Research: Shopper Marketing and the Role of In-store Marketing* (Vol. 11, pp. 147–187). Bingley: Emerald Publishing.

Burke, R. R., Harlam, B., Kahn, B., & Lodish, L. (1992). Comparing Dynamic Consumer Choice in Real and Computer-Simulated Environments. *Journal of Consumer Research, 19*(1), 71–82.

Campo, K., Gijsbrechts, E., & Guerra, F. (1999). Computer Simulated Shopping Experiments for Analyzing Dynamic Purchasing Patterns: Validation and Guidelines. *Journal of Empirical Generalisations in Marketing Science, 4*, 22–61.

Cooper, R. G., & Kleinschmidt, E. J. (1987). New Products: What Separates Winners from Losers? *Journal of Product Innovation Management, 4*, 169–184.

Desforges, T., & Anthony, M. (2013). *The Shopper Marketing Revolution*. Highland Park: RTC Publishing.

Desmet, P., Bordenave, R., & Traynor, J. (2013). Differences in Purchasing Behaviour Between Physical and Virtual Laboratory Stores. *Recherche et Applications en Marketing, 28*(2), 70–85.

Dickson, P. R., & Sawyer, A. G. (1990). The Price Knowledge and Search of Supermarket Shoppers. *Journal of Marketing, 54*(July), 42–53.

Erickson, G. (1995). New Package Makes a New Product Complete. *Marketing News, 29*(10), 10.

Feder, B. J. (1997, December 22). Test Marketers Use Virtual Shopping to Gauge Potential of Real Products. *The New York Times*.

Fusco, C. (1994, December). New Product Introduction: Challenges in Researching Customer Acceptance. In M. Adams & J. LaCugna (Eds.), *And Now for Something Completely Different: "Really" New Products* (pp. 29–31). Conference Summary, Report Number 94-124, Marketing Science Institute.

Garber, L. L. Jr., Burke, R. R., & Morgan Jones, J. (2000). The Role of Package Color in Consumer Purchase Consideration and Choice. *MSI Technical Report* (pp. 1–46), Report No. 00-104, Boston.

Huber, J., Payne, J. W., & Puto, C. (1982). Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis. *Journal of Consumer Research, 9*(June), 90–98.

Huber, J., & Puto, C. (1983). Market Boundaries and Product Choice: Illustrating Attraction and Substitution Effects. *Journal of Consumer Research, 10*(June), 31–44.

Jelinek, M., & Schoonhoven, C. B. (1993). *The Innovation Marathon*. San Francisco: Jossey-Bass Publishers.

Kirk, R. E. (1968). *Experimental Design: Procedures for the Behavioral Sciences*. Belmont: Brooks/Cole.

Madzharov, A. V., Block, L. G., & Morrin, M. (2015). The Cool Scent of Power: Effects of Ambient Scent on Consumer Preferences and Choice Behavior. *Journal of Marketing, 79*(1), 83–96.

Needel, S. P. (1998, July/August). Understanding Consumer Response to Category Management Through Virtual Reality. *Journal of Advertising Research, 38*, 61–67.

Nielsen Company. (2016, June). *Nielsen Breakthrough Innovation Report*. Available online at http://www.nielsen.com/breakthrough

Peck, J., & Childers, T. L. (2003). To Have and to Hold: The Influence of Haptic Information on Product Judgments. *Journal of Marketing, 67*(2), 35–48.

Schneider, J., & Hall, J. (2011, April). Why Most Product Launches Fail. *Harvard Business Review, 89*, 21–23.
Simonson, I. (1989). Choice Based on Reasons: The Case of Attraction and Compromise Effects. *Journal of Consumer Research, 16*(September), 158–174.
Simonson, I., & Tversky, A. (1992). Choice in Context: Tradeoff Contrast and Extremeness Aversion. *Journal of Marketing Research, 29*(August), 281–295.
Sorensen, H. (2009). *Inside the Mind of the Shopper: The Science of Retailing*. Philadelphia: Wharton School Publishing, ISBN-10: 0137126859.
Tjøstheim, I., & Saether-Larsen, H. (2005, February). *How to Validate a New MR Tool? A Case Study in FMCG*. ESOMAR, Innovate! Conference, Paris.
Tushman, M. (1994, December). Radical Innovation, Dominant Designs, and Organizational Evolution. In M. Adams & J. LaCugna (Eds.), *And Now for Something Completely Different: "Really" New Products* (pp. 22–26). Conference Summary, Report Number 94-124, Marketing Science Institute.
Underwood, R. L., Klein, N. M., & Burke, R. R. (2001). Packaging Communication: Attentional Effects of Product Imagery. *Journal of Product and Brand Management, 10*(7), 403–422.
Urban, G. L., Weinberg, B. D., & Hauser, J. R. (1996). Premarket Forecasting of Really-New Products. *Journal of Marketing, 60*(January), 47–60.
van Herpen, E., van den Broek, E., van Trijp, H. C. M., & Yu, T. (2016). Can a Virtual Supermarket Bring Realism into the Lab? Comparing Shopping Behavior Using Virtual and Pictorial Store Representations to Behavior in a Physical Store. *Appetite*. doi:https://doi.org/10.1016/j.appet.2016.07.033.
Wheelwright, S. C., & Clark, K. B. (1992). *Revolutionizing Product Development: Quantum Leaps in Speed, Efficiency, and Quality*. New York: The Free Press.
Wind, Y., & Mahajan, V. (1988). New Product Development Process: A Perspective for Reexamination. *Journal of Product Innovation Management, 5*, 304–310.
Winer, B. J. (1971). *Statistical Principles in Experimental Design*. New York: McGraw-Hill.
Zhang, X., Li, S., Burke, R. R., & Leykin, A. (2014). An Examination of Social Influence on Shopper Behavior Using Video Tracking Data. *Journal of Marketing, 78*(5), 24–41.

4

The Knowledge Domain of Affective Computing: A Scientometric Review

Maria Helena Pestana, Wan-Chen Wang, and Luiz Moutinho

Introduction

Emotions play an important role not only in successful a effective human-human communication, but also in human rational learning (Cambria 2016). Affective Computing recognizes this inextricable link between emotions and cognition, and works to narrow the communication gap between the highly emotional human and the emotionally challenged computer, by developing computational systems that respond to the affective states of the user (Calvo and D'Mello 2010). According to Rukavina et al. (2016), Affective Computing aims to detect users' mental states, revealing which feature customers enjoy and excluding those that

M.H. Pestana (✉)
ISCTE_IUL, Lisbon, Portugal

W.-C. Wang
Feng Chia University, Taiwan, Republic of China

L. Moutinho
University of Suffolk, Suffolk, England, UK

The University of the South Pacific, Suva, Fiji

L. Moutinho, M. Sokele (eds.), *Innovative Research Methodologies in Management*,
DOI 10.1007/978-3-319-64400-4_4

receive negative feedback. It, therefore, shows great potential to enhance companies' capabilities to manage their customer relationships, by improving their marketing strategies, and constantly gathering and predicting the attitudes of the general public toward their products and brands. The basic principle behind most Affective Computing systems is that they automatically recognize and respond to users' affective states during their interactions with a computer and thus provide data which can be used to enhance the quality of the interaction. Essentially, this is achieved by measuring multimodal signals, namely, speech, facial expressions, and/or psychobiology, and from these measurements, designing a computer interface which is more usable and effective. Affective Computing focuses on extracting a set of emotion labels (Picard 1997; Zeng et al. 2009; Calvo and D'Mello 2010; Schuller et al. 2011; Gunes and Schuller 2012) and polarity detection, usually a binary classification task with output such as positive versus negative, or like versus dislike (Pang and Lee 2008; Liu 2012; Cambria et al. 2013).

According to Frijda (2007), emotions have a short life in the field of consciousness, motivating behavior that requires immediate attention. Lang (2010) confirms the measurement of emotions as being critical in the advertising process, pointing to the fact that emotions are often conveyed by an advertising slogan, which arouses an appeal within consumers that positively predisposes them toward the message being communicated and thereby helping to deliver the desired image of brand position, which could generate enormous profit (Teixeira et al. 2012). Several marketing researchers (e.g., Bagozzi et al. 1999; Lee et al. 2007; Wang et al. 2015) state that Affective Computing supersedes self-report measures as a vehicle for evaluating emotions; and Cambria (2016) refers to the design of automatic Web mining tools for use in real time, as one of the most active research and development areas in Affective Computing.

Notable fallouts in marketing and financial market prediction have raised the interest of the scientific community and the business world in Affective Computing, which allows for the leverage of human-computer interaction, information retrieval, and multimodal signal processing.

This study provides information about Affective Computing, by measuring and visualizing the retrieved bibliographic references from the Web-of-Science Core Collection, between 1991 and 2016. It organizes these references into homogeneous clusters, identifying the most relevant sub-areas of research, and those references which are the most innovative, with more citations, with more connections between sub-areas, and which are responsible for recent advances in Affective Computing.

Marketing intelligence gathering, and strategic planning based on this body of knowledge on Affective Computing, provides opportunities for companies to enhance their customer relationship capabilities, to capture general public emotions in respect of their products and services, and to react accordingly.

The remainder of this chapter consists of four sections. The next section deals with the methodology, explaining the data collection approach and the quality of CiteSpace results. The third section presents the results, identifying in clusters the most efficient information on Affective Computing in metrical and graphical terms. Finally, concluding thoughts are offered in section four.

Methodology

Bibliographical records concerning Affective Computing, and published since 1991, were collected from the Web-of-Science of Thomson Reuters in the field of marketing, including the related areas of management and psychology. The resultant dataset contained 5078 records relating to 782 core articles and 4296 references that cite those articles at least once. This dataset was exported to CiteSpace for the scientometric review. CiteSpace is a free Java application for visualizing and analyzing emerging trends and changes in scientific literature. It allows for a multiple-perspective structural and temporal approach, and semantic patterns of references which help in interpreting the nature of clusters (Chen 1999, 2013). Structural metrics include centrality, modularity, and silhouette. High centrality identifies references that connect different clusters, and are responsible for the expansion of knowledge; high modularity shows that the references are distributed in non-overlapping clusters with clear boundaries; and high silhouette means great homogeneity of clusters, facilitating the uniformity of their labels.

Temporal metrics include citation burstness and novelty, burstness being used as a term to identify emergent references whose citation counts increase abruptly in a short period of time, and high novelty indicating references that represent creative ideas.

All the records were grouped into 40 non-overlapping and homogeneous clusters (#) based on their interconnectivity, applying the pathfinder pruning algorithm and a g-index with a scaling factor k = 5. Seven major clusters were found in the references and labeled #0 to #6. These account for 71.3% of the references (Fig. 4.1).

Semantic metrics allows for clusters to be labeled according to three algorithms (weight term frequency, TF*Inverse Document Frequency (IDF); log-likelihood ratio, LLR; mutual information, MI), identifying sub-areas of research. The clusters sorted by size and labeled by TF*IDF are: "multichannel physiology" (#0), "physiological signal" (#1), "emergence" (#2), "naturalistic interaction" (#3), "speech" (#4), "delivery" (#5), and "agent" (#6). The first two clusters (#0 and #1) have been the more active sub-areas of research up to this point in time (2016).

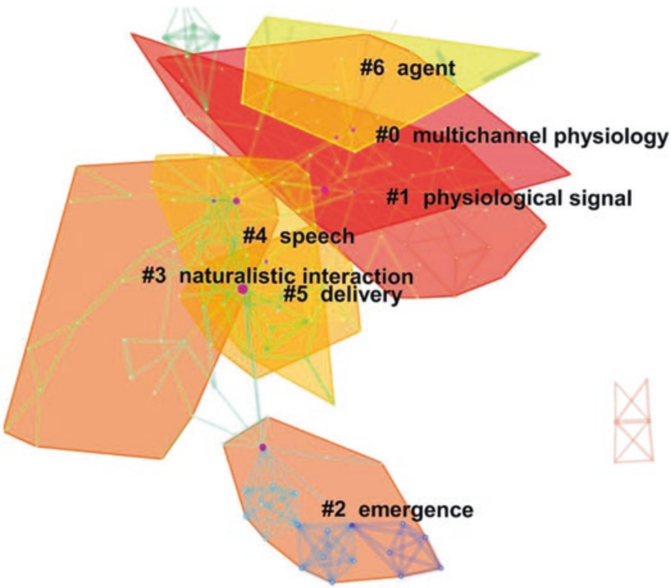

Fig. 4.1 Seven major clusters in AC, labeled by TF*IDF

Results from CiteSpace

The advantage of CiteSpace is its ability to provide the most useful information about Affective Computing in metrical and graphical terms, distributing the 5078 references by clusters, identifying those with more links (centrality), that are cited, more creative (sigma), and more explosive (burstness). CiteSpace represents a fountain of opportunities for managers to capture general public emotion in respect of products and services, thereby enhancing the capabilities of companies to improve their customer relationships.

Characterization of the Major Clusters

The seven major clusters or areas of research are labeled by these algorithms (TF*IDF, LLR, MI), according to the index terms coming from those references that cite members of each cluster, presented in Table 4.1. The oldest cluster in Affective Computing is "emergence" (#2), with an average year of publications being 1997, and the youngest cluster is "physiological signal" (#1), with an average year of publications being 2009. On average, the mean year difference between publication and citation varies between a minimum of 3 years, corresponding to the major clusters #0 (multichannel physiology) and #1, and a maximum of 7 years, corresponding to the "speech", cluster #4.

The largest cluster (#0) has 35 members and a silhouette value of 0.731, being labeled as "multichannel physiology" by both LLR and TF*IDF, and as "modality" by MI. This cluster has 491 citations, being most quoted (0.17) in "Affect detection: an interdisciplinary review of models, methods, and their applications" (Calvo and D'Mello 2010).

The second largest cluster (#1) has 27 members and a silhouette value of 0.802, being labeled as "response" by LLR, "physiological signal" by TF*IDF, and "music" by MI. This cluster has 263 citations, being most quoted (0.11) in "Consistent but modest: a meta-analysis on unimodal and multimodal affect detection accuracies from 30 studies" (D'Mello and Kory 2012).

Table 4.1 Major clusters, average year of publications, and citations

#	Cluster size	Cited size	Silhouette	TF*IDF	LLR	MI	Mean year publication	Mean year citation	Mean year difference
0	35	491	0.731	Multichannel physiology	Multichannel physiology	Modality	2008	2011	3
1	27	263	0.802	Physiological signal	Response	Music	2009	2012	3
2	23	92	0.965	Emergence	Computer	Human-computer interaction	1997	2001	4
3	21	180	0.857	Naturalistic interaction	Naturalistic interaction	Order crossing	2002	2008	6
4	21	229	0.724	Speech	Strategies	Audiovisual emotion	2002	2009	7
5	18	157	0.731	Delivery	Content delivery	Modality	2002	2008	6
6	12	78	0.95	Agent	Agent	Multi-score learning	2004	2009	5

The third largest cluster (#2) has 23 members and a silhouette value of 0.965, being labeled as "computer" by LLR, "emergence" by TF*IDF, and "human-computer interaction" by MI. This cluster has 92 citations, being most quoted (0.3) in "On the role of embodiment in the emergence of cognition and emotion" (Pfeifer & Scheier 1999).

The fourth largest cluster (#3) has 21 members and a silhouette value of 0.857, being labeled as "naturalistic interaction" by both LLR and TF*IDF, and as "order crossing" by MI. This cluster has 180 citations, being most quoted (0.38) in "Affect detection: an interdisciplinary review of models, methods, and their applications" (Calvo and D'Mello 2010).

The fifth largest cluster (#4) has 21 members and a silhouette value of 0.724, being labeled as "strategies" by LLR, "speech" by TF*IDF, and "audio-visual spontaneous emotion recognition" by MI. This cluster has 229 citations, being most quoted (0.24) in "Cross-corpus acoustic emotion recognition: variances and strategies" (Schuller et al. 2010).

The sixth largest cluster (#5) has 18 members and a silhouette value of 0.731, being labeled as "delivery" by TF*IDF, "content delivery" by LLR, and "modality" by MI. This cluster has 157 citations, being most quoted (0.17) in "Affect detection: an interdisciplinary review of models, methods, and their applications" (Calvo and D'Mello 2010).

The seventh largest cluster (#6) has 12 members and a silhouette value of 0.95, being labeled as "agent" by TF*IDF and LLR and "multi-score learning" by MI. This cluster has 78 citations, being most quoted (0.42) in "Multimodal semi-automated affect detection from conversational cues, gross body language, and facial features" (D'Mello and Graesser 2010).

More Innovative, Central, and Cited References by Cluster

Intellectual collaboration between references is fundamental to the overall understanding of a knowledge domain (Hu and Racherla 2008). References in the literature with more connections between different clusters or sub-areas of research (centrality), more citations, and more innovative (sigma) are revolutionary scientific publications.

"*Affective computing*" (Picard 1997) and "*Toward machine emotional intelligence: Analysis of affective physiological state*" (Picard et al. 2001) are the two most creative references ever in this area, according to their high sigma values, and, together with "Toward an affect-sensitive multimodal human-computer interaction" (Pantic and Rothkrantz 2003), are central references for the expansion of knowledge due to their connection between different clusters, summarized below.

The book *Affective computing* (Picard 1997), from cluster "emergence" (#2), states that the future "ubiquitous computing" environments will need to have human-centered designs instead of computer-centered designs, a fundamental component of human-human communication. Computing will move to the background, weaving itself into the fabric of our everyday living spaces and projecting the human user into the foreground. This reference prompted a wave of interest among computer scientists and engineers looking for ways to improve human-computer interfaces by coordinating emotion and cognition with task constraints and demands. Picard described three types of Affective Computing applications: first, systems that detect the emotions of the user; second, systems that express what a human would perceive as an emotion (e.g., an avatar, robot, and animated conversational agent); and third, systems that actually "feel" an emotion.

The paper "Toward machine emotional intelligence: Analysis of affective physiological state" (Picard et al. 2001), from cluster "delivery" (#5), proposed that machine intelligence needed to include emotional intelligence and demonstrated results suggesting the potential for developing a machine's ability to recognize human affective states given physiological signals.

The paper "Toward an affect-sensitive multimodal human-computer interaction" (Pantic and Rothkrantz 2003), from cluster "naturalistic interaction" (#3), reviewed the efforts toward the single modal analysis of artificial affective expressions and discussed how to integrate into computers a number of components of human behavior in the context-constrained analysis of multimodal behavioral signals.

The more innovative, central, and cited references are shown in Table 4.2, by title, authors, year, source, and cluster.

Table 4.2 The more innovative, central, and cited references by clusters in Affective Computing

Title	Authors	Year	Citations	Sigma	Centrality	Source	#
A survey of affect recognition methods: audio, visual, and spontaneous expressions	Zeng, Z.H., Roisman, G.I; & Huang, T.S.	2009	79	5.09	0.24	IEE PATTERN ANL M I	0
Affect detection: an interdisciplinary review of models, methods, and their applications	Calvo, R.A.; & D'Mello, S.	2010	61	4.66	0.15	IEEE T AFFECT COMPUT	0
Toward machine emotional intelligence: analysis of affective physiological state	Picard, RW; Vyzas, E; & Healey, J.	2001	38	334.24	0.50	IEE PATTERN ANL M I	5
Emotion recognition in human-computer interaction	Cowie, R., Douglas-Cowie, E.; Tsapatsoulis, N.; & et al.	2001	38	7.42	0.14	IEEE SIGNAL PROC MAG	4
Affective computing	Picard, R.W.	1997	35	688.39	0.35	MIT Press	2
Toward an affect-sensitive multimodal human-computer interaction	Pantic, M; & Rothkrantz, L.J.M.	2003	34	15.22	0.40	P IEEE	3
Emotion recognition based on physiological changes in music listening	Kim, J., & Elisabeth, A.	2008	31	1.22	0.04	P IEEE	0
Automatic prediction of frustration	Kapoor, A.; Burleson,W.; & Picard, R.	2007	30	1.56	0.12	Int.J.Human-Computer Studies	0
Affective computing: challenges	Picard, R.W.	2003	28	1.24	0.03	Int.J.Human-Computer Studies	3
DEAP: A database for emotion analysis using physiological signals	Koelstra, S.	2012	26	1.18	0.02	IEEE T AFFECT COMPUT	1

"A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions" (Zeng et al. 2009) and "Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications" (Calvo and D'Mello 2010) are the references with more citations, both from the major cluster "multichannel physiology" (#0), summarized below.

The paper "A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions" (Zeng et al. 2009) developed algorithms to detect subtleties and changes in the user's affective behavior, in order to initiate interactions based on this implicit information rather than on explicit messages usually involved in the tradition interface devices, such as the keyboard and mouse. These algorithms are intended to process naturally occurring human affective behavior, with a view to multimodal fusion for human affect analysis, including audiovisual, linguistic, paralinguistic, and multi-cue visual fusion based on facial expressions, head movements, and body gestures.

The paper "Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications" (Calvo and D'Mello 2010) stresses the need to include within Affective Computing practice emotion theories that have emerged and that rely on cross-disciplinary collaboration and active sharing of knowledge. It reviews models that emphasize emotions as expressions, embodiments, outcomes of cognitive appraisals, social constructs, and products of neural circuitry, aiming to be a useful tool for new researchers by providing taxonomy of resources for further exploration and discussing different theoretical viewpoints and applications.

Trends in Affective Computing

A timeline view shows references through time and clusters (Fig. 4.2). CiteSpace identifies all references by the first name of the author(s). The size of the nodes corresponds to the number of citations of the reference, that is, a large node indicates many citations. Citations with an unexpected increase over a short period of time (burst strength) are marked as red rings around a node.

The position of "Automatic prediction of frustration" (Kapoor et al. 2007) in cluster #0 and "DEAP: A Database for Emotion Analysis Using

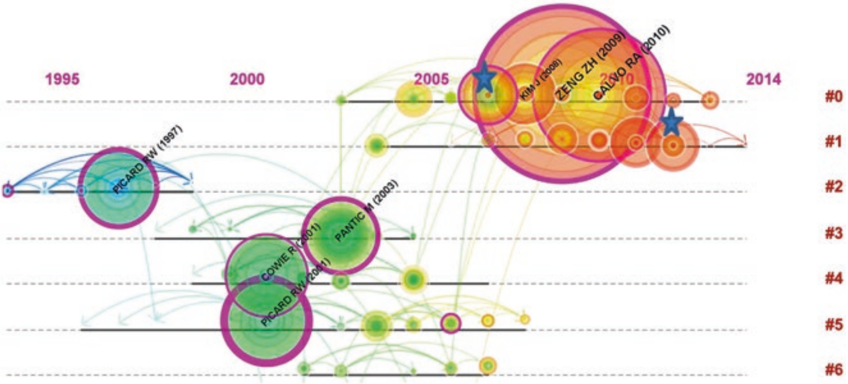

Fig. 4.2 Timeline view by author's first name and cluster

Physiological Signals" (Koelstra et al. 2012) in cluster #1 are both marked with a star. "Toward an affect-sensitive multimodal human-computer interaction" (Pantic and Rothkrantz 2003) is superimposed on "Affective computing: challenges" (Picard 2003) in cluster #3.

The majority of references were published after 2000, and the more active clusters are #0 and #1, being essential to the literature of Affective Computing, the main papers which have already been discussed.

Burstness provides a temporal perspective, indicating where the frequency of a reference increases abruptly in relation to its peers during a short period of time (Lee et al. 2016). A citation burst has two attributes: the intensity (strength) and the length of time the status lasts.

Table 4.3 lists the references with the strongest citation bursts across the entire dataset, according to the clusters to which they belong. It can be seen that most of the references started to burst in year 2000 and have continued until 2016.

"Affective Computing" (Picard 1997), from cluster #2, is the reference with the highest burst citation, having a significant statistical fluctuation over 2000–2005. "Emotion recognition in human-computer interaction" (Cowie et al. 2001), from cluster #4, over 2005–2009, and "Toward machine emotional intelligence: Analysis of affective physiological state" (Picard et al. 2001), are respectively the second and third references with strong intensity, the last also showing the highest duration of six years, over 2003–2009, as

Table 4.3 Trends on Affective Computing by periods of time

Authors	Source	Year	Strength	Begin	End	1999 – 2016	#
Picard	AFFECTIVE COMPUTING	1997	21.680	**2000**	2005		2
Cowie, Douglas-Cowie, Tsapatsoulis, et al.	IEEE SIGNAL PROC MAG	2001	15.039	**2005**	2009		4
Picard, Vyzas and Healey	IEEE T PATTERN ANAL	2001	14.436	**2003**	2009		5
Calvo and D'Mello	IEEE T AFFECT COMPUT	2010	10.694	**2013**	2016		0
Chang and Lin	ACM T INTEL SYST TEC	2011	8.262	**2013**	2016		1
Koelstra	IEEE T AFFECT COMPUT	2012	8.257	**2013**	2016		1
Pantic and Rothkrantz	P IEEE	2003	8.175	**2005**	2011		3
Zeng, Roisman and Huang	IEEE T PATTERN ANAL	2009	7.701	**2013**	2016		0
Picard	INT J HUM-COMPUT ST	2003	7.442	**2006**	2008		3
Kim	IEEE T PATTERN ANAL	2008	5.698	**2012**	2016		0

well as "Toward an affect-sensitive multimodal human-computer interaction" (Pantic and Rothkrantz 2003), over the period 2005–2011.

"Emotion recognition in human-computer interaction" (Cowie et al. 2001) has also been one of the most comprehensive and widely cited reference in reviewing the efforts to reach a single modal analysis of artificial affective expressions, and in providing a comprehensive summary of qualitative acoustic correlations for prototypical emotions. The writers of that paper discuss the recognition of seven different human negative and neutral emotions (bored, disengaged, frustrated, helpless, over-strained, angry, and impatient) by technical systems, focusing on problems of data gathering and modeling, in an attempt to create a "Companion Technology" for Human Computer Interaction that allows the computer to react to human emotional signals.

"Affective computing: challenges" (Picard 2003), from cluster #3, shows a statistically significant fluctuation over the period 2006–2008. It raises and responds to several criticisms of Affective Computing, emphasizing the need for a balance and articulated state-of-the-art research challenges, especially with respect to affect in human-computer interaction. This paper suggested that designers of future computing can continue with the development of computers that ignore emotions, or they can take the risk of making machines that recognize emotions, communicate them, and perhaps even "have" them, at least in the ways in which emotions aid in intelligent interaction and decision making.

References with greater burst impact by 2016 are "Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications" (Calvo and D'Mello 2010), from cluster #0, over 2013–2016; "LIBSVM: A Library for Support Vector Machines" (Chang and Lin 2011), over 2013–2016; "DEAP: A Database for Emotion Analysis Using Physiological Signals" (Koelstra et al. 2012), from cluster #1, over 2013–2016; "A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions" (Zeng et al. 2009), from cluster #0, over 2013–2016; and "Emotion recognition based on physiological changes in music listening" (Kim and Elizabeth 2008), from cluster #0, over 2012–2016.

With the exception of Zeng et al. (2009) and Calvo and D'Mello (2010), previously noted, below is a summary of the references that have an abrupt increase of citations by 2016.

"LIBSVM: A Library for Support Vector Machines" (Chang and Lin 2011) presents details of how to implement a library for support vector machines (SVMs) and the package LIBSVM. It discusses in detail issues concerning how to solve SVMs' optimization problems, theoretical convergence, multiclass classification, probability estimates, and parameter selection.

"DEAP: A Database for Emotion Analysis Using Physiological Signals" (Koelstra et al. 2012), considered vital for multimedia information retrieval, characterizes multimedia content with relevant, reliable, and discriminating tags. This reference presents a multimodal dataset for the analysis of spontaneous emotions, where implicit tagging of videos using affective information helps recommendation and retrieval systems to improve their performance. The dataset was made publicly available, and other researchers were encouraged to use this data for testing their own affective state estimation methods.

"Emotion recognition based on physiological changes in music listening" (Kim and Elizabeth 2008) investigates the potential of physiological signals for emotion recognition as opposed to audiovisual emotion channels such as facial expression or speech. This paper develops a novel scheme of emotion-specific multilevel dichotomous classification and shows an improvement in its performance compared with direct multiclass classification.

Computer systems are now attempting to interact more naturally with the users as human beings. Graphical user interfaces are becoming more flexible and intelligent to adapt to different human interests. The adaptive systems' applications learn the user interactions and then make user-friendly platforms. The application of the recognized affect in creating a diversified and context-specific response would be more domain specific, forming another level of work in Affective Computing, one with importance because of its wide application in different areas.

The accuracy of human emotion recognition has been improved by utilizing advanced analysis methods and techniques including image processing, voice recognition, natural language processing, and electroencephalography devices, which are time and resource-consuming, and which have certain prerequisites such as the availability of a webcam, microphone, and highly technical equipment. This contrasts with the common situation whereby computers and portable digital devices, such as handhelds, tablets, and mobile phones, process and analyze keyboard keystroke dynamics, mouse movements, and touch screen interactions (Bakhtiyari and Husain 2014). Since the mouse is available in all computers, and users rely very much on this, it is possible to capture the human physiological signals for determining the human affective state via the mouse in a non-intrusive manner, without the user being consciously being aware of a physiological data-capturing device. Hence, a solution is found through the use of low-cost equipment, without imposing any burdens on the user (Fu et al. 2014). The importance of touch in communicating emotions and intensifying interpersonal communication has been analyzed in Affective Computing to detect and display emotions (Eid and Osman 2016).

The development of systems capable of mining emotions and sentiments over the Web in real time to track public viewpoints on a large scale represents one of the most active research and development areas, being important not only for commercial purposes but also for monitoring hostile communications or model cyber-issue diffusion (Cambria 2016). The Web is evolving to become an area in which consumers are defining future products and services, as it has made users more enthusiastic about sharing their emotions and opinions through several online collaboration media, like social networks, online communities, blogs, and wikis, in all fields related to everyday life, such as commerce and tourism. With the increasing number of webcams installed in end-user devices such as smart phones, touchpads, and netbooks, a greater volume of affective information is being posted to social online services in an audio or audiovisual format rather than on a purely textual basis. Public opinion is destined to gain increased prominence and so is Affective Computing with its capacity for recognizing and classifying emotions.

Table 4.4 provides examples of some recent research studies that have investigated references included in the two major clusters (#0 and #1) just discussed.

Affective Computing is mainly interpreted in terms of emotions and sentiments, with an emphasis on the classification and recognition of emotions via human-computer interaction, and the provision of implicit information about the changes in the affective states. This is depicted in Fig. 4.3, which displays the keywords assigned to each reference in the dataset.

Table 4.4 More relevant citers of the most active clusters

Title	Authors	Year	Source
Fuzzy model of dominance emotions in affective computing	Bakhtiyari, K.; & Husain, H.	2014	EURAL COMPUTING & APPLICATIONS
Physiological mouse: towards an emotion-aware mouse	Fu, Y.; Leong, H.V.; Ngai, G.; Huang, M.X.; & Chan, S.C.F.	2014	IEEE 38th Annual International COMPSAC
Affective computing and sentiment analysis	Cambria, E.	2016	IEEE Intelligent Systems
Affective haptics: current research and future directions	Eid, M A., & Osman, H.A.	2016	IEEE Access
Time-delay neural network for continuous emotional dimension prediction from facial expression sequences	Meng, H.; Deng, J.; Chen, J., & Cosmas, J.	2016	IEEE TRANSACTIONS ON CYBERNETICS

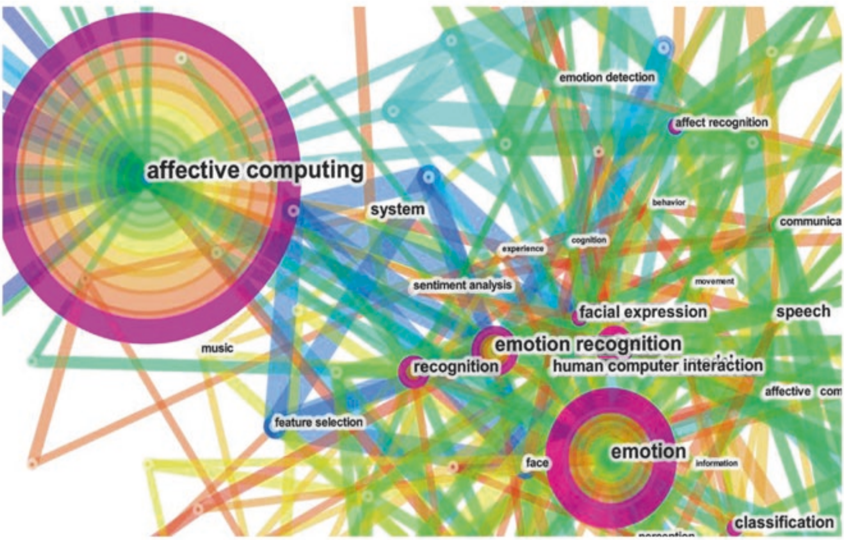

Fig. 4.3 Network with the relevant keywords

Concluding Thoughts

This chapter has traced the advancement of Affective Computing through the analysis of expert references in the literature. It has done this by using computational techniques to discern patterns and trends at various levels of abstraction: cited, central, innovative, and burstness references; sources of publications; and keywords. The major clusters are #0, "multichannel physiology", and the newest cluster #1 "physiological signal", which continue to be the most cited and to demonstrate the largest burst citations. The most creative and central references are those of Picard (1997) and Picard et al. (2001); and those with recent burstness are seen to come from Kim and Elizabeth (2008), Calvo and D'Mello (2010), Chang and Lin (2011), Koelstra et al. (2012), and Zeng et al. (2009), the latter also being the most cited reference ever in this area.

Through its descriptive findings about Affective Computing, obtained efficiently via the use of CiteSpace, this chapter provides opportunities for companies to enhance their capabilities in respect of customer relationships. It is up to managers to choose the most useful tools to capture and respond to the emotions of their customers about the products or services offered by their companies.

References

Bagozzi, R. P., Gopinoth, M., & Nyer, P. U. (1999). The Role of Emotions in Marketing. *Journal of the Academy of Marketing Science, 27*(2), 184–206.

Bakhtiyari, K., & Husain, H. (2014). Fuzzy Model of Dominance Emotions in Affective Computing. *Neural Computing & Applications, 25* (6th ed.), 1467–1477.

Calvo, R. A., & D'Mello, S. K. (2010). Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications. *IEEE Transactions on Affective Computing, 1*(1), 18–37.

Cambria, E. (2016). Affective Computing and Sentiment Analysis. *IEEE Intelligent Systems, 31*(2), 1541–1672.

Cambria, E., et al. (2013). New Avenues in Opinion Mining and Sentiment Analysis. *IEEE Intelligent Systems, 28*(2), 15–21.

Chang, C. C., & Lin, C. J. (2011). LIBSVM: A Library for Support Vector Machines. *ACM Transactions on Intelligent Systems and Technology*, *2*(3), Article 27.

Chen, C. (1999). Visualizing Semantic Spaces and Author Cocitation Net-Works in Digital Libraries. *Information Processing & Management, 35*(3), 401–420.

Chen, C. (2013). The Structure and Dynamics of Scientific Knowledge. In *Mapping Scientific Frontiers* (pp. 163–199). London: Springer.

Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., & Taylor, J. G. (2001, January). Emotion Recognition in Human-Computer Interaction. *IEEE Signal Processing Magazine, 18,* 32–80.

D'Mello, S., & Graesser, A. (2010). Multimodal Semi-automated Affect Detection from Conversational Cues, Gross Body Language, and Facial Features. *User Modelling and User-Adapted Interaction, 20*(2), 147–187.

D'Mello, S., & Kory, J. (2012, October 22–26). *Consistent but Modest: A Meta-Analysis on Unimodal and Multimodal Affect Detection Accuracies from 30 Studies*. ICMI '12, Santa Monica, California.

Eid, M. A., & Osman, H. A. (2016). Affective Haptics: Current Research and Future Directions. *IEEE Access*. https://doi.org/10.1109/ACCESS.2015.2497316.

Frijda, N. (2007). *The Laws of Emotion*. London: Routledge.

Fu, Y., Leong, H. V., Ngai, G., Huang, M. X., & Chan, S. C. F. (2014, July 21–25). Physiological Mouse: Towards an Emotion-Aware Mouse. In C. K. Chang, Y. Gao, A. Hurson, et al. (Eds.), *38th Annual IEEE International Computer Software and Applications Conference (COMPSAC)* (pp. 258–263). Local: Vasteras, SWEDEN.

Gunes, H., & Schuller, B. (2012). Categorical and Dimensional Affect Analysis in Continuous Input: Current Trends and Future Directions. *Image and Vision Computing, 31*(2), 120–136.

Hu, C., & Racherla, P. (2008). Visual Representation of Knowledge Networks: A Social Network Analysis of Hospitality Research Domain. *International Journal of Hospitality Management, 27*, 302–312.

Kapoor, A., Burlesonc, W., & Picard, R. W. (2007). Automatic Prediction of Frustration. *International Journal of Human-Computer Studies, 65*, 724–736.

Kim, J., & Elizabeth, A. (2008). Emotion Recognition Based on Physiological Changes in Music Listening. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 30*(2), 2067–2083.

Koelstra, S., Soleymani, M., Yardani, A., & Nijholt, A. (2012). DEAP: A Database for Emotion Analysis Using Physiological Signals. *IEEE Transactions on Affective Computing, 3*(1), 18–31.

Lang, P. J. (2010). Emotion and Motivation: Toward Consensus Definitions and a Common Research Purpose. *Emotion Review, 2*(3), 229–233.

Lee, N., Broderick, A. J., & Chamberlain, L. (2007). What is 'Neuromarketing'? A Discussion and Agenda for Future Research. *International Journal of Psychophysiology, 63*, 199–204.

Lee, Y. C., Chen, C., & Tsai, X. T. (2016). Visualizing the Knowledge Domain of Nanoparticle Drug Delivery Technologies: A Scientometric Review. *Applied Sciences, 6*(1), 11. https://doi.org/10.3390/app6010011.

Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. San Rafael: Morgan and Claypool.

Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval, 2*(1–2), 1–135.

Pantic, M., & Rothkrantz, L. J. M. (2003). Towards Emotion Recognition in Human Computer Interaction. *IEEE, Proceedings of the IEEE, 91*(9), 1370–1390.

Pfeifer, R., & Scheier, C. (1999). *Understanding Intelligence*. Cambridge, MA: MIT Press.

Picard, R. (1997). *Affective Computing*. Cambridge: The MIT Press.

Picard, R. W. (2003). Affective Computing: Challenges. *International Journal of Human-Computer Studies, 59*, 55–64.

Picard, R. W., Vyzas, E., & Healey, J. (2001). Toward Machine Emotional Intelligence: Analysis of Affective Physiological State. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 23*(10), 1175–1191.

Rukavina, S., Sascha, G., Holder, H., Jun-Weng, T., Walter, S., & Traue, H. (2016). Affective Computing and the Impact of Gender and Age. *PLoS One, 11*, e0150584.

Schuller, B., Vlasenko, R., Eyben, F., et al. (2010). Cross-Corpus Acoustic Emotion Recognition: Variances and Strategies. *IEEE Transactions on Affective Computing, 1*(2), 119–131. https://doi.org/10.1109/T-AFFC.2010.8.

Schuller, B., Batliner, A., Steidl, S., & Seppi, D. (2011). Recognising Realistic Emotions and Affect in Speech: State of the Art and Lessons Learnt from the First Challenge. *Speech Communication, 53*(9), 1062–1087.

Teixeira, T., Wedel, M., & Pieters, R. (2012). Emotion-Induced Engagement in Internet Video Advertisements. *Journal of Marketing Research, 49*(2), 144–159.

Wang, W. C., Chien, C. S., & Moutinho, L. (2015). Do You Really Feel Happy? Some Implications of Voice Emotion Response in Mandarin Chinese. *Marketing Letters, 26*(3), 391–409.

Zeng, Z. H., Riisman, G. I., & Huang, T. S. (2009). A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 31*(1), 39–58.

5

The Effect of Emotions on Brand Recall by Gender Using Voice Emotion Response with Optimal Data Analysis

Wan-Chen Wang, Maria Helena Pestana, and Luiz Moutinho

Introduction

Emotions have an important role in directing responses to stimuli and have been extensively investigated in the field of psychology, receiving unprecedented recognition in the field of marketing (Consoli 2010). Bagozzi et al. (1999) described emotions as producing a mental state of readiness arising from cognitive appraisals of events or thoughts; Frijda (2007) says that emotions motivate behaviour, but as they are short-lived

W.-C. Wang (✉)
Feng Chia University, Taiwan, Republic of China

M.H. Pestana
ISCTE_IUL, Lisbon, Portugal

L. Moutinho
University of Suffolk, Suffolk, England, UK

The University of the South Pacific, Suva, Fiji

L. Moutinho, M. Sokele (eds.), *Innovative Research Methodologies in Management*,
DOI 10.1007/978-3-319-64400-4_5

in the field of consciousness, they require immediate attention; Bagozzi and Dholakia (2006) state that emotions are important to accomplish collective goals; Romani et al. (2012) state that negative emotions drive consumers away from brands; Chowdhry et al. (2015) state that not all negative emotions lead to concrete construal; Dubé and Morgan (1998) affirm that higher states of positive emotions magnified increasing trends in satisfaction; Orth and Holancova (2004) show that emotions evoked by advertisements vary according to gender; Pham et al. (2013) show the ad-evoked feelings exert a positive influence on brand attitudes; Wierenga (2011) says that sophisticated behavioural laboratories and brain imaging methods are the next research frontier in managerial decision-making in marketing.

Recall is an important physiological factor of the human learning process, where information is stored in the mind as collaborated nodes, creating a semantic network. Recalled data with stronger emotional background will ultimately lead to greater attention to stimuli and increase communication effects (Vakratsas and Ambler 1999). Precise measurement of emotions is important given the significance of emotion in the advertising process (e.g., Martensen et al. 2007). An appealing advertising campaign arouses consumer's positive emotion towards communicating message that is often themed with advertising slogan. According to Fiske and Taylor (1991), priming exists when current ideas come to consumer's mind with greater ease than ideas that are not currently activated. Slogan, as a themed emotion of advertising campaign, can be applied to prime various attributes of brand perceptions (Bouch 1993).

Psychophysiological measures are more connected to brand recall (Hazlett and Hazlett 1999; Wang et al. 2015). According to Wang et al. (2015), brand recall in Mandarin Chinese is better captured by voice emotion response than by self-reported measures. Their study in Taiwan in 2015 involved a sample of 142 participants, from 18 to 55 years old. Emotions were measured in five nominal categories (happy, bored, neutral, sad, and angry), in order to classify eight advertising slogans of long-established brands, familiar to consumers in the Greater Chinese market: KFC (Just want KFC), Burger King (Roast is delicious), Pepsi (Enjoy delicacies, drink Pepsi!), Coca-Cola (Enjoy cold, My Coca Cola), 7-Eleven (Oh! Thanks heaven! 7-Eleven), Family Mart (Family Mart is your home!), Suzuki (Firepower is to win), and SYM (Burn my hot blood chopper dream). To develop this previous study, the authors intend to select the voice emotion response and analyse its effect on brand recall by gender, applying optimal data analysis (ODA) which is robust with small samples (Yarnold and Soltysik 2005). A mixed method of both qualitative and quantitative designs was used in the study, which, according to Fodness (1994), results in a comprehensive measurement in understanding tourist motivations.

In order to statistically measure the scientific contribution of the study, a scientometric review of the bibliographical references, which are growing rapidly in this digital information era, is undertaken. Such a review has theoretical and practical implications since the detection of potentially valuable ideas is essential to safeguard the integrity of scientific knowledge (White and McCain 1998; Morris and Van 2008; Tabah 1999).

The perceived contribution of this chapter is its illumination of when, how, and why the effect of emotions on brand recall by gender, using voice emotion response with ODA, can be useful in the scientific domain. Hence, it is necessary to obtain an up-to-date understanding of the scientific field's intellectual structure, and to identify exactly how the current issue connects with previously disparate patches of knowledge, creating a network of cognitively demanding ideas. The insights into the structure and dynamics of this issue are gained in computational terms, using publication records from Web-of-Science and exporting them to CiteSpace, a scientometric branch of informatics enabling the analysis of bibliographical records, articles actively cited from the domain issue, and emerging trends and changes in scientific literature over time. It is a tool that applies multiple temporal, structural, and semantic metrics and allows the visualisation of patterns from both citing and cited items (White and Griffith 1982; Chen 2013; Chen et al. 2014). The temporal interval is sliced into equal mutually exclusive length segments of one year, and an individual co-citation network is derived from each slice. The merged networks and the major changes between adjacent periods can be highlighted in a panoramic visualisation. The following sections present literature review, methodology, results, and conclusion of this study.

Literature Review

The literature review about the subject that entitles this research was done using structural, temporal, and semantic metrics, applying CiteSpace, which allows knowing the distribution and contribution for knowledge of those references by areas of investigation. This purpose is explained in five steps: first, what is a scientometric review; second, the quality of the obtained analysis; third, the identification of the most cited, central, burst, and novelty references; fourth, the identification of emerging trends; and fifth, the identification of ODA as an isolated area in the network of all references.

Scientometric Review of the Literature with CiteSpace

The bibliographical records published since 1900 appropriate to the subject title "the effect of emotions on brand recall by gender using voice emotion response with optimal data analysis" were collected from the Web-of-Science of Thomson Reuters, using the broadest possible terms to ensure that subsequent analysis via CiteSpace covered all major components of a knowledge domain (Chen 2016). The terms chosen with the logical operator "AND" were emotions, advertising recall, and affective computing; and with the logical operator "OR" were gender recall and ODA. Two datasets were obtained: one corresponding directly to the core or topic search, producing 226 papers, and the other indirectly including 3826 citations appearing at least once in any paper of the original core set, because they can be thematically relevant to the subject matter (Chen 2013; Chen et al. 2014). References are a general term for any written scientific work, that is, articles, books, and conferences. The resultant expanded dataset containing 4052 references was then merged and exported to CiteSpace for scientometric review.

The Scope and Quality of the Network

CiteSpace filters the analysis, narrowing the period of time to 1991–2016. The corresponding network has 294 unique nodes, each one representing a cited reference, and 686 links connecting them, showing various topics, and it has always been possible to find some irrelevant ones. Each node cites a number of related references, where the connectivity between cited and citing papers captures the underlying intellectual structure of each cluster.

All the references published in a given year create a network called a time slice. The entire time period is divided into equal length mutually exclusive segments of one year. The network configuration for each time slice is based on one of seven criteria: modify g-index, top N, top N%, threshold, by citations, usage 180, usage 2013. The modify g-index with a positive scaling factor k gives the average number of an author's most important publications. In the current chapter, the network configuration of citations and co-citations uses the g-index with a threshold ($k = 9$).

Top N represents the number of most cited references or occurrences chosen. Top N% represents the percentage of most cited references or occurrences. Threshold interpolation combines nodes and links stipulating minimum values for citation counts (c), co-citation counts (cc), and co-citation coefficients (ccv). Citations give the simple and cumulative frequency distribution of the number of cited articles. Usage 180 represents the highest number of the most download papers in the last 180 days, and usage 2013 the highest number of the most download papers since 2013.

The sequence of time-sliced networks is merged into one containing all nodes appearing between 1991 and 2016, giving an overview of how the scientific field has been evolving over time. Each link from individual networks is merged based on the earliest time stamp (default option), and subsequent links connecting the same pair of nodes are dropped in order to detect the earliest moment when a connection was first made in the literature (Chen 2008).

The network can change over time with the addition or elimination of references and links. The objective is to simplify a dense network through effective pruning, allowing its visualisation to be clarified (Chen 2008). CiteSpace supports two pruning algorithms: the minimum spanning tree and Pathfinder. Pathfinder is the default option and was used in this study because it keeps the relevant links at a minimum and preserves the chronological growth patterns (Chen et al. 2001; Chen 1998).

The new improved network includes individual components of bibliographic records (nodes), and their relationships and changes over time, established via the method of Document Co-Citation (DCC), which partitioned the network into 50 non-overlapping clusters, measured by cosine coefficients (Small 1980). The application of structural metrics (betweenness centrality, modularity, silhouette) and temporal metrics (burstness, sigma) allowed the network to be filtered and reduced to seven (regarded as optimal) relevant clusters, using the spectral clustering method (Luxburg et al. 2009; Luxburg 2006) in which strongly connected nodes were assigned to the same cluster, and non-connected nodes were assigned to different clusters. These seven major clusters correspond to 59.5% (=175) of the 249 references, with silhouettes in the interval [0.850; 0.967], containing at least 10 references. Clusters of smaller size, despite having a high silhouette, probably indicate that the same author provides all the references, thus being of no interest for analysis (Schneider 2006).

The scientometric analysis via CiteSpace includes modularity and silhouette as structural metrics that measure the quality of the network. Modularity (Q) ranges from zero to one, measuring the extent to which a network can be divided into a number of independent groups, named clusters, such that nodes within the same group are more tightly connected than nodes between different ones. A low value of Q suggests a network that cannot be reduced to clusters with clear boundaries, whereas a high value, like the one in the current chapter (0.827), suggests a well-constructed network. According to Chen (2006), networks with modularity scores of 1 or very close to 1 may turn out to be trivial special cases in which individual components are simply isolated from one another. The silhouette ranges from −1 to 1 and is useful in estimating the uncertainty involved in identifying or interpreting the nature of a cluster (Rousseeuw 1987). A value of 1 represents a perfect separation from other clusters, while a negative value suggests its diversity or heterogeneity. The mean silhouette for the whole period defined by the merged network (0.544) is higher than 0.5, thereby fulfilling the conditions required by Chen (1999) for further analysis.

The topics involved in the field of research can be delineated in terms of keywords assigned to each reference in the dataset, as shown in Fig. 5.1. Adjacent ones are often attributed to the same reference, with characters' size proportional to their frequency. The keywords of ODA, affective computing, human-computer interaction, and emotion recognition are marked with arrows due their connection with the subject title.

Fig. 5.1 Network of keywords assigned to papers in the field of research

Table 5.1 Network of 359 keywords' citations

Citation		
Counts	%	Keywords
134	37.33	Affective computing
78	21.73	Human-computer interaction
41	11.42	Emotion
31	8.64	Emotion recognition
18	5.01	Recognition
16	4.46	System
15	4.18	Facial expression
13	3.62	Model
13	3.62	Classification

The frequency of the more relevant keywords is shown in Table 5.1. Emotions and Affective Computing were used in Web-of-Science for the current domain.

The panoramic bibliographical landscape shows the spatial distribution of clusters throughout the period 1991–2016. The collection of articles representing the state of the art in the field are identified by the first author and have their characteristic dimensions proportional to their size (Yu and Sample 1965). One isolated node, identified by an arrow at the top of Fig. 5.2, refers to Yarnold and Soltysik (2005), and is related to ODA, which appears in the subject title. According to Yarnold and Soltysik (2016), ODA accommodates all metrics, requires no distributional assumptions, allows for analytic weighting of individual observations, explicitly maximises predictive accuracy, and supports multiple methods of assessing validity.

The citation history shown in Table 5.2 shows that ODA was cited in seven papers during the period 2008–2013, none of them connected with the domain of the effect of emotions on brand recall by gender using voice emotion response.

Relevant References in the Literature

The most important references in the literature for the subject that entitles this research are those: more cited, with extraordinary attribute values, visualised by a large radius; more central, shared by different clusters or being the only link between two different clusters, being

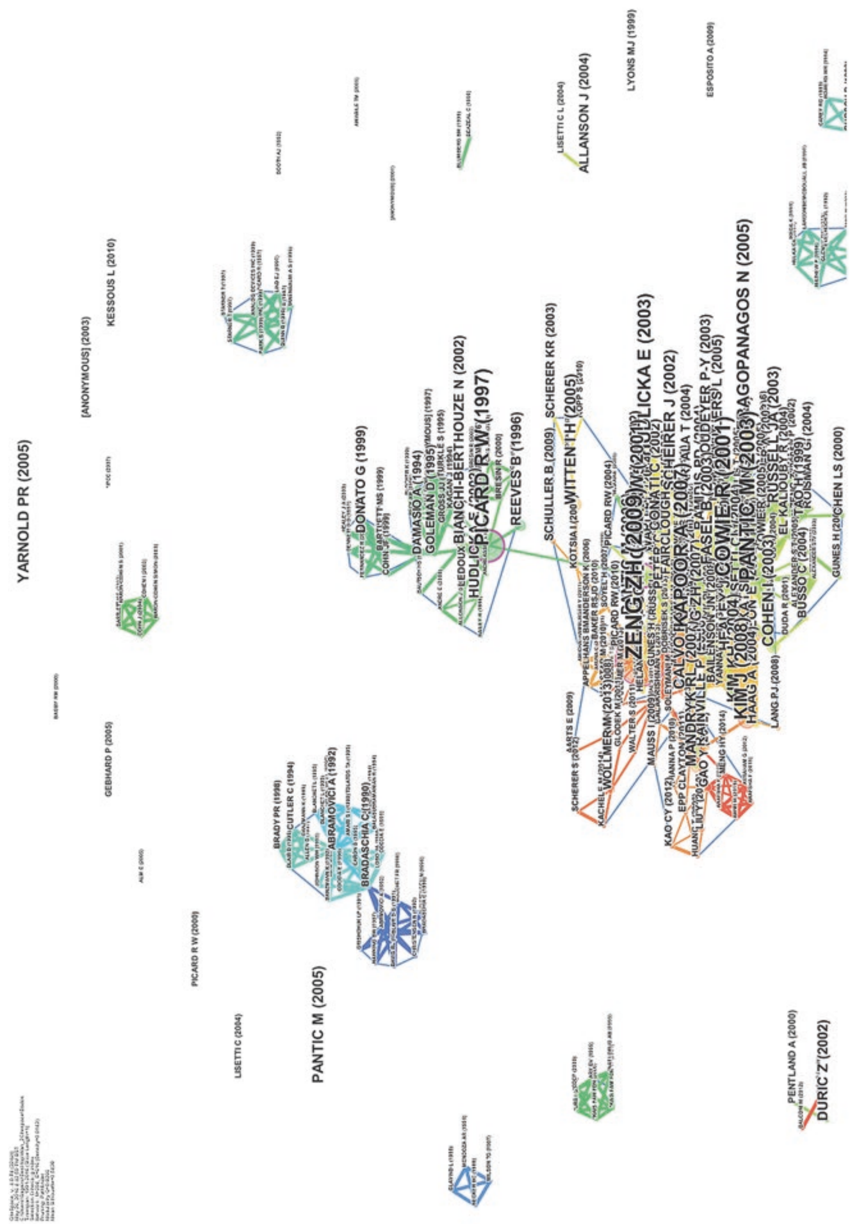

Fig. 5.2 Panoramic bibliographical landscape between 1991 and 2016

Table 5.2 Citation history of Yarnold P R (2005), optimal data analysis

Records	Citations	Citing article of Yarnold PR 2005
1.	7	HAN SD, 2009, J HEAD TRAUMA REHAB, V24, P57
2.	1	HAN SD, 2010, J INT NEUROPSYCH SOC, V16, P721, DOI https://doi.org/10.1017/S1355617710000512
3.	13	JONES A, 2011, CLIN NEUROPSYCHOL, V25, P1207, DOI https://doi.org/10.1080/13854046.2011.600726
4.	20	LAVIGNE JV, 2010, BEHAV THER,V41, P198
5.	2	PAPE TLB, 2013, REHABIL PSYCHOL, V58, P253, DOI https://doi.org/10.1037/a0032610
6.	1	RUPERT PA, 2012, PROF PSYCHOL-RES PR, V43, P495, DOI https://doi.org/10.1037/a0029420
7.	11	SMART CM, 2008, J INT NEUROPSYCH SOC, V14, P842, DOI https://doi.org/10.1017/S1355617708081034

candidates for intellectual turning points, identifying progressive knowledge domain; with novelty; and having a high frequency, or bursts, in a short period of time. Those references are identified by area of research they belong. For that in the network of all literature references, it is necessary labelling those clusters (semantic metrics), and after that identifying the references with high betweenness centrality (structural metrics), and those with citation bursts and novelty (temporal metrics).

Areas of Research

To identify the areas of research included in the network, we apply semantic metrics where clusters are labelled from extracted titles according to three algorithms [weight terms frequency (TF (Term Frequency)*IDF (Inverse Document Frequency)); likelihood-ratio (LLR); mutual information (MI)], providing a set of cues that facilitate their interpretation and serving as symbols for scientific ideas and methods (Schneider 2009). The algorithm TF*IDF tends to represent the most salient aspects of a cluster (Salton et al. 1975); and MI and LLR reflect its unique aspect (Dunning 1993). According to Chen (2008), LLR usually gives the best result in terms of uniqueness and coverage of each cluster.

The most common area of research, with an average year of publications being 2003, corresponds to cluster (#0), has the major number of references (39), and a silhouette 0.85, indicating high homogeneity between a large number of articles in the domain, linked to each other with a very widely accepted concept, labelled as *spontaneous expression* by TFI*IDF, *audio* by LLR, and *speech* by MI. The oldest area of research, with an average year of publications being 1994, corresponds to cluster #5 with 22 homogeneous members (silhouette 0.967), being labelled as *gravitational wave* by both TFI*IDF and LLR, and as *sensitivity* by MI. The newest emerging area, with an average year of publications being 2011, corresponds to cluster #6, containing 17 homogeneous members (silhouette 0.891), and is labelled as *mouse* by TFI*IDF, *future direction by* LLR, and as *arousal classification* by MI (Table 5.3).

The most relevant 15 clusters, labelled via the TFI*IDF algorithm, are represented in circle packing graph, the size and distance of each circle being proportional to their intercognitivity and relevance (Chen 1999). Cluster #0 (spontaneous expression) with 39 references is identified by the largest circle, while cluster #6 (mouse), with 17 references, has the seventh smallest size, marked with an arrow. Other small clusters with a small number of references are gravitating around the first seven, as shown in Fig. 5.3.

The representation of the network with the most salient connections between nodes is in Fig. 5.4, where clusters have been labelled by the LLR algorithm. The larger characters and circles are proportional to the dimension of clusters.

Cluster #5, *gravitational wave,* has no connections with other clusters, and consequently has no betweenness centrality references, not being useful as a theoretical support for the subject title.

The Most Relevant References of the Literature by Area of Research

The most relevant references, analysed by citations, centrality, burst, and novelty (sigma), are identified by the first author in Table 5.4. Centrality assumes values in the interval [0;1] and measures the extent to which a reference is in the middle of a path that connects different clusters (Brandes 2001). High centrality identifies potentially revolutionary scientific publications as well as gatekeepers in social network (Freeman 1977). Bursts aim to investigate

Table 5.3 Semantic metrics, mean year of publications, and citations

Cluster	Size	Silhouette	Label (TFI*IDF)	Label (LLR)	Label (MI)	Publication mean (year)	Citation mean (year)	Year difference
0	39	0.85	(17.8) spontaneous expression	Audio (348.86, 1.0E-4)	Speech	2003	2009	6
1	27	0.897	(11.99) human-computer interaction	Detection system (136.3, 1.0E-4)	Emotion representation	2007	2012	5
2	24	0.954	(15.06) user emotion	Computer (402.11, 1.0E-4)	Computer	1998	2004	6
3	23	0.893	(10.27) empathic technologies	Data fusion (97.52, 1.0E-4)	Speech	2005	2010	5
4	23	0.782	(12.76) continuous emotion	Co-present creative collaboration (119.22)	Arousal classification	2007	2013	6
5	22	0.967	(19.55) gravitational wave	Gravitational wave (373.89, 1.0E-4)	Sensitivity	1994	1998	4
6	17	0.891	(16.16) mouse	Future direction (26.43, 1.0E-4)	Arousal classification	2011	2014	3
7	10	1	(16.16) Georgia tech	Georgia tech (70.39, 1.0E-4)	…	1997	–	–
8	9	1	(18.36) gravitational-wave data	(13.36) gravitational-wave data analysis	Gravitational-wave data analysis	2002	–	–

(*continued*)

Table 5.3 (continued)

Cluster	Size	Silhouette	Label (TFI*IDF)	Label (LLR)	Label (MI)	Publication mean (year)	Citation mean (year)	Year difference
9	9	1	Learning environment	(19.79) video user interface	Video user interface	2002	–	–
10	9	0.979	Building hal	(13.36) building hal	Building hal	1998	–	–
11	9	1	Method	(15.64) deuterium NMR tissue perfusion	Deuterium NMR tissue perfusion	1992	–	–
12	9	0.977	Stochastic background	(13.36) stochastic background	Stochastic background	1990	–	–
13	6	1	…	(13.66) root; (13.66) psychiatric care	Root	1999	–	–
14	5	1	…	(12.76) automated inference	Automated inference	2002	–	–

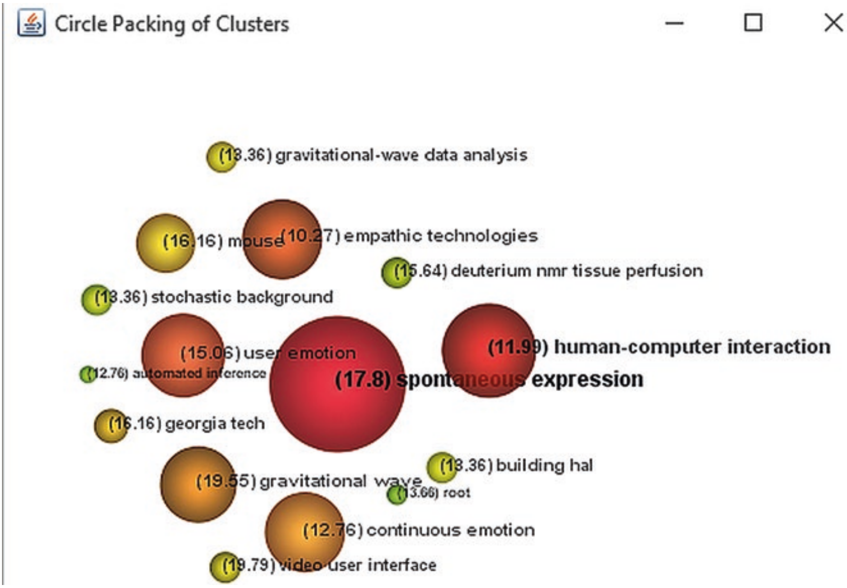

Fig. 5.3 Circle packing graph with the 15 most cited clusters labelled by TF*IDF

where the citations of a reference increase abruptly, using the algorithm introduced by Kleinberg (2002). Burst detects when surge the statistically significant fluctuations on the citation count of a particular reference during a short time interval within the overall time period (Chen et al. 2001), regardless of how many times their host articles are cited. Scientific novelty is measured through sigma, which is a combination of burstness and centrality, identifying publications that represent creative ideas, with a role more prominent than the rate of its recognition by peers.

The timeline view (Fig. 5.5) of most of the documents cited shows they were published after 1995, while the clusters with more recent publications after 2010, marked with an arrow, came from clusters #1 and #6.

The majority of burst references (52.2%) come from cluster #3, followed by the newest cluster #6 which contains 35%.

To understand the links between clusters, in Table 5.5, the most relevant papers that cite or mention members for the top seven clusters can be seen. It can be seen that paper by Balasubramanian et al. (1996),

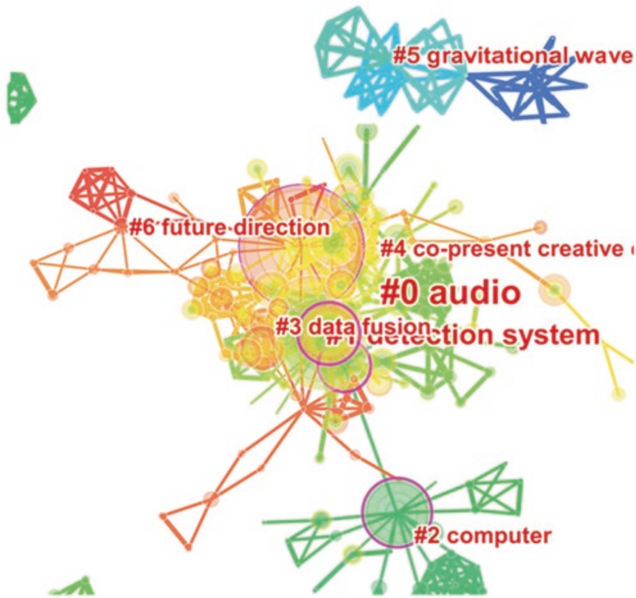

Fig. 5.4 Links between clusters labelled by LLR

"Gravitational waves from coalescing binaries: detection strategies and Monte Carlo estimation of parameters" is connected with 41% (coverage) of the field of research included in the oldest cluster #5. The paper by Zeng, Z H (2007) "A survey of affect recognition methods: audio, visual and spontaneous expressions" is connected with 33% of the field of research included in the major cluster #0. The paper by Eid, M A (2016) "Affective Haptics: Current Research and Future Directions" is connected with 29% of the field of research included in the newest cluster #6.

A citation bursts has two attributes: the intensity and the length of time the status lasts. Table 5.6 lists the ten references with the strongest citation bursts across the entire dataset, according to the clusters to which they belong. Picard (1997) from cluster #2 is the reference with the highest burst citation, having a significant statistical fluctuation over the period 2001–2005. Zeng ZH (2009) from cluster #4 follows with an abrupt increase in relation to its peers over the period 2013–2016.

Table 5.4 Most relevant references by cluster, citations, burst, centrality, and sigma

References	Year	Source	Volume	Pages	Cluster	Citations	Burst	Centrality	Sigma
Zeng ZH	2009	IEEE T PATTERN ANAL	V31	P39	4	29	8.71	0.18	4.32
Pantic M	2003	P IEEE	V91	P1370	0	22		0.06	1.00
Cowie R	2001	IEEE SIGNAL PROC MAG	V18	P32	3	22	7.40	0.06	1.53
Picard RW	1997	AFFECTIVE COMPUTING	V	P	2	16	14.99	0.21	17.22
Kapoor A	2007	INT J HUM-COMPUT ST	V65	P724	1	15		0.30	1.00
Picard RW	2001	IEEE T PATTERN ANAL	V23	P1175	3	12	4.12	0.22	2.26
Hudlicka E	2003	INT J HUM-COMPUT ST	V59	P1	1	11	2.46	0.03	1.08
Kim J	2008	IEEE T PATTERN ANAL	V30	P2067	3	11	6.41	0.01	1.05
Fragopanagos N	2005	NEURAL NETWORKS	V18	P389	4	10		0.01	1.00
Kim KH	2004	MED BIOL ENG COMPUT	V42	P419	3	9	2.56	0.04	1.11
Calvo RA	2010	IEEE T AFFECT COMPUT	V1	P18	3	8	3.78	0.07	1.27
Scherer KR	2005	SOC SCI INFORM	V44	P695	1	7	4.06	0.00	1.00
Healey JA	2005	IEEE T INTELL TRANSP	V6	P156	3	7	2.52	0.03	1.08
Devillers L	2005	NEURAL NETWORKS	V18	P407	0	6		0.06	1.00
Reeves B	1996	MEDIA EQUATION PEOPL	V	P	2	6	6.47	0.00	1.00
Rainville P	2006	INT J PSYCHOPHYSIOL	V61	P5	6	6	3.53	0.00	1.02
Cohn J F	2004	INT J WAVELETS MULTI	V2	P1	0	5		0.05	1.00
Goleman D	1995	EMOTIONAL INTELLIGEN	V	P	2	4	4.15	0.04	1.16
Gao Y	2012	ACM T COMPUT-HUM INT	V19	P	6	4	3.56	0.03	1.11

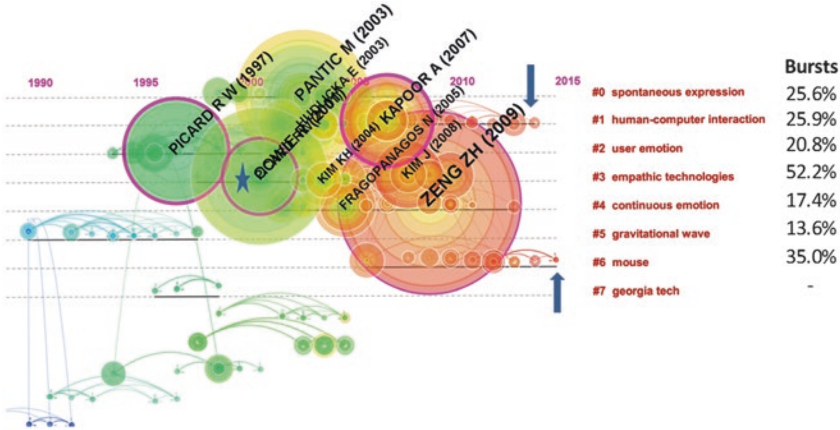

Fig. 5.5 Timeline view of some of the more cited references, where clusters are labelled by TF*IDF

Table 5.5 The most relevant citers of each cluster, identified by the first author

Coverage (%)	Bibliographic details	Cluster
	Most relevant citers of each cluster	
33	Zeng, ZH, 2007, A survey of affect recognition methods: audio, visual and spontaneous expressions	0
22	DMello, SK, 2015, A review and meta-analysis of multimodal affect detection systems	1
33	Picard, RW, 2002, computers that recognise and respond to user emotion: theoretical and practical implications	2
17	Frantzidis, CA, 2010, Toward emotion aware computing: an integrated approach using multichannel neurophysiological recordings and affective visual stimuli	3
17	DMello, SK, 2010, Multimodal semi-automated affect detection from conversational cues, gross body language, and facial features	4
41	Balasubramanian, R, 1996, Gravitational waves from coalescing binaries: detection strategies and Monte Carlo estimation of parameters	5
29	Eid, MA, 2016, Affective haptics: current research and future directions	6

Table 5.6 Top 10 periods of burst references, identified by the first author

References	Year	Strength	Begin	End	1991 – 2016	Cluster
PICARD R W, 1997, AFFECTIVE COMPUTING, V, P	1997	**149.921**	2001	2005		2
ZENG ZH, 2009, IEEE T PATTERN ANAL, V31, P39, DOI	2009	**87.067**	2013	2016		4
COWIE R, 2001, IEEE SIGNAL PROC MAG, V18, P32, DOI	2001	**73.981**	2007	2008		3
KIM J, 2008, IEEE T PATTERN ANAL, V30, P2067, DOI	2008	**64.141**	2012	2016		3
GOLEMAN D, 1995, EMOTIONAL INTELLIGEN, V, P	1995	**41.479**	2001	2003		2
PICARD RW, 2001, IEEE T PATTERN ANAL, V23, P1175, DOI	2001	**41.156**	2008	2009		3
SCHERER KR, 2005, SOC SCI INFORM, V44, P695, DOI	2005	**40.567**	2011	2013		1
CALVO RA, 2010, IEEE T AFFECT COMPUT, V1, P18, DOI	2010	**37.816**	2012	2016		3
GAO Y, 2012, ACM T COMPUT-HUM INT, V19, P, DOI	2012	**35.56**	2014	2016		6
REEVES B, 1996, MEDIA EQUATION PEOPL, V, P	1996	**6.475**	2002	2003		2

In the first seven clusters, the most cited references inspiring interest in their fields, identified by the first author, are shown in a timeline view through the algorithm weight terms frequency (TFI*IDF): Zeng (2009), Pantic et al. (2003), Cowie (2001), Picard (1997), Kapoor et al. (2007), Kim (2008), Fragopanagos and Taylor (2005), and Kim (2008). The landmarks Picard (2001), and Hudlicka (2003), appear superimposed on Fig. 5.4, which is why they are marked with a star.

The most relevant references, analysed by citations, burst, centrality, and sigma, are shown in Table 5.4. Each line of the table provides information about the first author, year of publication, source, the cluster to which it belongs, number of citations received, burst, centrality, and novelty (sigma).

Zeng (2009) is a landmark from cluster #4, labelled as *continuous emotion* (TFI*IDF), *co-presentative collaboration* (LLR), or *arousal classification* (MI). It has strong centrality and burst, being a pioneer reference. This paper develops algorithms that can process spontaneously occurring human affective behaviour. It examines available approaches to solving the problem of machine understanding of human affective behaviour, outlining some of the scientific and engineering challenges to advancing human affect-sensing technology.

The study of Pantic (2003) is the highest landmark of the major cluster #0, labelled *spontaneous expression* (TFI*IDF), *audio* (LLR), or *speech* (MI). This paper discusses how to integrate into computers a number of components of human behaviour in the context-constrained analysis of multimodal behavioural signals.

Cowie (2001) is a landmark of cluster #3, labelled as *empathic technologies* (TFI*IDF), *data fusion* (LLR), or *speech* (MI). It is a pioneer reference, with high burst, responsible for connections between different fields of knowledge (centrality), and a novelty reference (sigma). This paper discusses the recognition of seven different human negative and neutral emotions (bored, disengaged, frustrated, helpless, over-strained, angry, and impatient) by technical systems, focusing on problems of data gathering and modelling, in an attempt to create a "Companion Technology" for human-computer Interaction that allows the computer to react to human emotional signals.

Picard (1997) is a landmark from cluster #2, labelled as user *emotion* (TFI*IDF) or *computer* (LLR, MI). It has strong centrality and burst, and is a structurally essential and inspirational pioneering reference. It advances a compelling argument in favour of the need for affective computers, suggesting that a truly intelligent system, artificial or otherwise, cannot be implemented without emotional mechanisms, drawing upon data and examples taken from a wide spectrum of disciplines, from neurobiology to folk psychology, showing the potential positive applications of affective computing.

Kapoor et al. (2007) is a landmark of cluster #1, labelled as *human-computer interaction* (TFI*IDF), *detection system* (LLR), or *emotion representation* (MI). It has the highest centrality, being a pivotal reference. This paper presents the first automated method that assesses, using multiple channels of affect-related information, whether a learner is frustrated. The new assessment method is based on Gaussian process classification and Bayesian inference, and its performance suggests that non-verbal channels carrying affective cues can help to provide important information for a system to allow it to formulate a more intelligent response.

Those important landmarks and pivot nodes support the relevance of the subject title, concerned with applying a human computer interface to analyse how particular emotions (sad, bored, angry, neutral, and happy) affect brand recall. Table 5.4 shows the most relevant references ranked by citations.

In order to detect emergent terms or understand the significance of a reference within a short period of time, regardless of how many times it was cited, it is important to know the top burst references, which could show some tendency over time to support the subject title.

Goleman (1995), from cluster #2, has a citation burst between 2001 and 2003, describing a model of emotional intelligence based on competences that enable a person to demonstrate intelligent use of their emotions in managing themselves and working with others to be effective at work. This reference seeks to understand the characteristics that predict better performance and more fulfilled lives.

Scherer (2005), from cluster #1, with a burst of citations between 2011 and 2013, attempts to sensitise researchers to the importance of the definition of emotions, in order to guide research and make it comparable across disciplines, which is central for the development of instruments and measurement operations, as well as for the communication of results, and their discussion between scientists. This paper distinguishes emotions from other affective states or traits, and discusses how to measure them in a comprehensive and meaningful way.

Kim (2008), from cluster #3, with a burst of citations between 2012 and 2016, investigates the potential of physiological signals as reliable channels for emotion recognition. This paper designs a musical induction method to acquire a naturalistic dataset for evoking certain emotions based on the voluntary participation of subjects. The emotion recognition problem is decomposed into several refining processes using additional modalities, valence recognition, and the resolution of subtle uncertainties between adjacent emotion classes.

Calvo and D'Mello (2010), from cluster #3, with a recent burst of citations between 2012 and 2016, describes the progress in the field of Affective Computing, with a focus on affect detection. In order to achieve a truly effective real-world system of affective computing, the need is stressed for an integrated examination of emotion theories from multiple areas. This paper provides meta-analyses on existing reviews of affected detection systems that focus on modalities like physiology, face, and voice, and also reviews emerging research on more novel channels such as text, body language, and complex multimodal systems.

Gao et al. (2012), from cluster #6, with a recent burst of citations between 2014 and 2016, analyses whether the touch behaviours when people are playing games on touch-screen mobile phones reflect players' emotional states. The use of touch as an affective communicative channel would be an interesting modality when facial expression and body-movement recognition, or bio-signal detection, may not be feasible.

The investigation of references with strong citations bursts reveals that they can be grouped essentially into two branches: one focusing on the theoretical concept of emotions/affection and the other on emotion recognition through the interface with computers. The exploration of the expanded dataset suggests that designing and executing novel approaches to address the recognition of emotions through computers are significant and widely accepted concerns in the domain knowledge.

Emerging Trends

The networks are intellectual structures of associated co-cited references representing the knowledge of a scientific field (Chen 2013), evolving over time during which newly published articles may introduce profound structural variation or can have little or no impact on the structure (Chen et al. 2014). Changes in modularity are represented by bars in Fig. 5.6, each network being based on a two-year slicing period.

The number of new publications per year is represented by an increased line. The significant decrease in the modularity above 0.5 in three time periods, 1995–1996, 2001–2002, and 2005–2006, is expected to be explained by the appearance of citation burst references, playing an important role in changing the overall intellectual structure. Table 5.4 shows that the emergent trends are explained mainly by five references: Goleman (1995); Reeves and Nasse (1996); Cowie (2001); Picard (2001); and Sherer (2005).

Note that 2015–2016 has modularity less than 0.5 but has no burst references, implying that those references do not contribute to emerging trends.

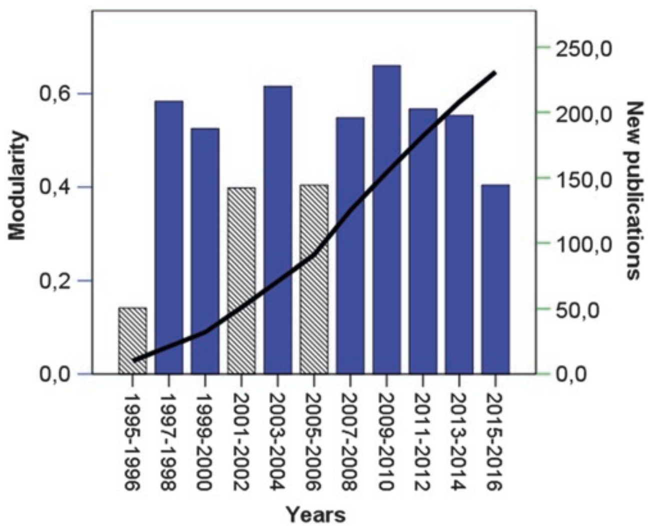

Fig. 5.6 Modularity and new publications between 1991 and 2016

Optimal Data Analysis

ODA is a method developed by Yarnold and Soltysik (2005) which offers maximum predictive accuracy to data, even when the assumptions of the alternative statistical models are not applied. This method is used to identify patterns in the data that distinguish the effect of voice emotion response on brand recall by gender.

The accuracy of ODA is obtained by calculating the following measures: Sensitivity is the proportion of actual females who are correctly predicted by the model; specificity is the proportion of actual males who are correctly predicted by the model; and Effect Size Sensitivity (ESS) is an index of predictive accuracy relative to chance, where values less than 25% indicate a relative weak effect; 25–50% indicate a moderate effect, 50–75% indicate a relatively strong effect, and 75% or greater indicate a strong effect over chance.

To assess generalisability, ODA first estimates using the entire sample (training set), calculating accuracy measures as described previously. Next, the model is cross-validated, and the accuracy measures are recalculated. If the accuracy measures remain consistent with those of the original model using the entire sample, as in the present study, then it can be said that the model is generalisable. The current study applies the approach of "leave-one-out" (LOO) cross-validation, which is simply an n-fold cross-validation, where $n = 141$ observations in the dataset. Each observation in turn is left out, and the model is estimated for all remaining observations. The predicted value is then calculated for the hold-out observation, and the accuracy is determined as female or male in predicting the outcome for that observation. The results of all predictions are used to calculate the final accuracy estimates. Model accuracy measures are calculated using the average values across all hold-out models. All variables included in the ODA model were constrained to achieve identical classification accuracy in training (total sample) and LOO validity analysis. To ensure adequate statistical power, inhibit over-fitting, and increase the likelihood of cross-validation when the model is applied to classify a smaller independent sample, model endpoints were constrained to have N >= 10% of the total sample (Yarnold and Soltysik 2016)

Results

The 141 participants divided into 80 females and 61 males were presented with eight slogans to be classified into five categories of emotion, registered by Voice Emotion Response, in order to determine the effect of the emotion on brand recall (Wang et al. 2015). This was done by gender and by applying ODA. With the exceptions of Family Mart and 7-Eleven, the other slogans have two patterns regarding recall by gender: males feel happier in cars (Suzuki and SYM), showing greater recall than females. In the remaining four slogans, Coca-Cola, Pepsi, KFC, and Burger King, the opposite occurs, with females being happier and showing better recall than males. The results are in line with those found by other researchers (Teixeira et al. 2012; Martensen et al. 2007; Faseur and Geuens 2006; Janssens and De Pelsmacker 2005; Vakratsas and Ambler 1999) who state that a significant relationship exists between advertising effectiveness and positive emotions. The brand recall is higher when associated with positive emotions, as shown in Table 5.7, which reveals the performance indices of ODA to be better than chance for all brands: 8.38% for Burger King (exact $p = 0.02$); 8.77% for KFC (exact $p = 0.003$); 13.31% for Coca-Cola ($p < 0.00001$); 15.1% for Pepsi ($p < 0.00001$); 15.96% for Suzuki ($p < 0.00001$); and 26.57% for SYM ($p < 0.00001$).

The observed values of the relationship between brand recall by gender are shown in Table 5.8. The modal gender for female recall is in Coca-Cola (98.75% > 72.13%); KFC (92.94% > 75.4%); Pepsi (92.50% > 62.3%),

Table 5.7 ODA performance indices for all slogans

Indexes/brands	Coca (%)	KFC (%)	Pepsi (%)	Burger King (%)	SYM (%)	Suzuki (%)
Overall accuracy	68.09	64.38	68.79	60.28	75.18	66.67
Female recall accuracy	98.75	92.94	92.50	72.50	13.11	13.11
Male recall accuracy	27.87	24.59	37.70	44.26	66.25	71.25
Effect size sensitivity (ESS)	26.62	17.53	30.20	16.76	53.14	31.91
Above chance (%)	13.31	8.77	15.1	8.38	26.57	15.96

and Burger King (72.5% > 55.73%). The modal gender for male recall is in SYM (86.89% > 33.75%) and Suzuki (60.66% > 28.75%).

Except for Family Mart and 7-Eleven, where there is a high 95.74% of recall, almost equal between genders, all other brands have a statistically significant relationship with gender. All other brands show two different patterns of recall: males are more likely to recall than female with respect to cars, while in the other four brands, Coca-Cola, Pepsi, KFC, and Burger King, females are more likely to recall than males. For Suzuki, males have at least 1.886 (=1/0.530) times more chance of recall than females, while for Coca-Cola, females have at least 3.929 times more chance of recalling than males.

The magnitudes of recall by similar brands are as follows: SYM (56.74%) higher than Suzuki (42.55%); Coca-Cola (87.23%) higher than Pepsi (79.43%); KFC (85.11%) higher than Burger King (65.25%); Family Mart and 7-Eleven are equal (both 95.74%). This can be explained by the fact that in the Taiwanese market, SYM motorcycles are more popular than Suzuki motorcycles, and Coca-Cola is still the leading soft drink brand. KFC came to Taiwan in 1985 and Burger King in 1990, so KFC is much more well-known and loved by Taiwanese consumers. Finally, Family Mart and 7-Eleven are the top two popular brands of convenience store.

Table 5.8 Voice emotion recall by gender for all slogans

Voice emotion recall	Coca-Cola		Pepsi		Family Mart		7-Eleven	
	Female N (%)	Male N (%)	Female	Male	Female	Male	Female	Male
Yes	79 (98.75%)	44 (72.1%)	74 (92.5%)	38 (62.3%)	74 (92.5%)	61 (100%)	74 (92.5%)	61 –100%
No	1 (1.3%)	17 (27.9%)	6 (7.5%)	23 (37.7%)	6 (7.5%)	0 0%	6 (7.5%)	0 0%

Voice emotion recall	Burger King		KFC		Suzuki		SYM	
	Female	Male	Female	Male	Female	Male	Female	Male
Yes	58 (72.5%)	34 (55.7%)	74 (92.5%)	46 (75.4%)	23 (28.7%)	37 (60.7%)	27 (33.8%)	53 (86.9%)
No	22 (27.5%)	27 (44.3%)	6 (7.5%)	15 (24.6%)	57 (71.3%)	24 (39.3%)	53 (66.3%)	8 (13.1%)

Conclusion

The scientometric review supports the relevance of research on the subject title. The analysis of the field and the citation-based expansion has outlined the evolutionary trajectory of the collective knowledge over 1991–2016 and highlighted the areas of active pursuit. Emerging trends are identified from computational properties from CiteSpace, which is designed to facilitate sense-making tasks in relation to scientific frontiers based on relevant domain literature.

This chapter tracks the advancement of the collective knowledge of a dynamic scientific community through the analysis of expert references in the literature domain, using computational techniques to discern patterns and trends at various levels of abstraction, as cited and co-cited references. The research on the subject title makes the following contributions to science: connecting two isolated areas (ODA and marketing communication); integrating them into six of seven major clusters in this knowledge domain; and differentiating brand recall by gender. Human-computer interaction (cluster #1) and future direction (cluster #6) are the most recent issues that continue to be referenced in 2016, with which the subject title has great probability to be connected.

It was found that men had better recall scores than women when related to cars, whereas women scored higher in recall when dealing with soft drinks and fast food. There has always been a clear association between the memorisation process of advertisements and the triggering of an emotional state towards brands. It has been confirmed that a positive relationship between consumer recall and brand stimuli does exist.

Owing to the preliminary nature of this study, our subject sample was limited in size and scope, and the findings of this investigation may not generalise to other samples. The Voice Emotion Response can only recognise five basic emotions, and this critically constrains this study. Researchers in the Tatung University are trying to develop further techniques to recognise more emotions that better suit marketing and advertising research. Further research will benefit greatly if the technology improves. Additionally, methods' diversity improves the robustness of marketing research (Davis et al. 2011). Future research should, therefore, try to use other psychophysiological measures to obtain more valid and reliable results with the aim of generating a deeper understanding of the construct of emotions.

References

Bagozzi, R. P., & Dholakia, U. M. (2006). Antecedents and Purchase Consequences of Customer Participation in Small Group Brand Communities. *International Journal of Research in Marketing, 23*, 45–61.

Bagozzi, R. P., Gopinoth, M., & Nyer, P. U. (1999). The Role of Emotions in Marketing. *Journal of the Academy of Marketing Science, 27*(2), 184–206.

Balasubramanian, R., Sathyaprakash, B. S., & Dhurandhar, S. V. (1996). Gravitational Waves from Coalescing Binaries: Detection Strategies and Monte Carlo Estimation of Parameters. *Physics Review, D53*, 3033–3055. https://doi.org/10.1103/PhysRevD.53.3033.

Boush, D. M. (1993). How Advertising Slogans Can Prime Evaluations of Brand Extensions. *Psychology and Marketing, 10*(1), 67–78.

Brandes, U. (2001). A Faster Algorithm for Betweenness Centrality. *Journal of Mathematical Sociology, 25*(2), 163–177.

Calvo, R. A., & D'Mello, S. (2010). Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications. *IEEE Transactions on Affective Computing, 1*(1), 18–37.

Chen, C. (1998). Generalized Similarity Analysis and Pathfinder Network Scaling. *Interacting with Computers, 10*, 107–128.

Chen, C. (1999). Visualising Semantic Spaces and Author Cocitation Net-Works in Digital Libraries. *Information Processing and Management, 35*(3), 401–420.

Chen, C. (2006). CitespaceII: Detecting and Visualizing Emerging Trends and Transient Patterns in Scientific Literature. *Journal of the American Society for Information Science and Technology, 57*, 359–377.

Chen, C. (2008). An Information-Theoretic View of Visual Analytics. *IEEE Computer Graphics and Applications, 28*(1), 18–23.

Chen, C. (2013). The Structure and Dynamics of Scientific Knowledge. In *Mapping Scientific Frontiers* (pp. 163–199). London: Springer.

Chen, C. (2016). *The Citespace Manual*. http://cluster.ischool.drexel.edu/~cchen/citespace/Cite SpaceManual.pdf

Chen, C., Kuljis, J., & Paul, R. J. (2001). Visualizing Latent Domain Knowledge. *IEEE Transtions on Systems, Man, and Cybernetics- Parte C: Applications and Reviews, 31*(4), 518–529.

Chen, C., Dubin, R., & Kim, M. C. (2014). Emerging Trends and New Developments in Regenerative Medicine: A Scientometric Update (2000–2014). *Expert Opinion on Biological Therapy, 14*, 1295–1317.

Chowdhry, N., Winterich, K. P., Mittal, V., & Morales, A. C. (2015). Not All Negative Emotions Lead to Concrete Construal. *International Journal of Research in Marketing, 32*, 428–430.

Consoli, D. (2010). A New Concept of Marketing: The Emotional Marketing. *Brand Broad Research in Accounting, Negotiation, and Distribution, 1*(1), 1–8.

Cowie, R. (2001). Machine Understanding of Human Behavior. *IEEE Signal Processing Magazine, 18*, 32.

Davis, D. F., Golicic, S. L., & Boerstler, C. N. (2011). Benefits and Challenges of Multiple Methods Research in Marketing. *Journal of the Academy of Marketing Science, 39*(3), 467–479.

Dube, L., & Morgan, M. S. (1998). Capturing the Dynamics of In-process Consumption Emotions and Satisfaction in Extended Service Transactions. *International Journal of Research in Marketing, 15*, 309–320.

Dunning, T. (1993). Accurate Methods for the Statistics of Surprise and Coincidence. *Computational Linguistics, 19*(1), 61–74.

Faseur, T., & Geuens, M. (2006). Different Positive Feelings Leading to Different Ad Evaluations. *Journal of Advertising, 35*(4), 129–142.

Fiske, S. T., & Taylor, S. E. (1991). *Social Cognition* (2nd ed.). New York: McGraw-Hill.

Fodness, D. (1994). Measuring Tourist Motivation. *Annals of Tourism Research, 21*(3), 555–581.

Fragopanagos, F., & Taylor, J. G. (2005). Emotion Recognition in Human-Computer Interaction. *Neural Networks, 18*, 389–405.

Freeman, L. C. (1977). A Set of Measuring Centrality Based on Betweenness. *Sociometry, 40*, 35–41.

Frijda, N. (2007). *The Laws of Emotıon*. London: Routledge.

Gao, Y., Bianchi-Berthouze, N., & Meng, H. (2012). What Does Touch Tell Us About Emotions in Touchscreen-Based Gameplay? *ACM Transactions on Computer-Human Interaction (TOCHI), 19*(4), article 31. https://doi.org/10.1145/2395131.2395138.

Goleman, D. (1995). *Emotional Intelligence*. New York: Bantam Books.

Goleman, D., Boyatzis, R., & Rhee, K. S. (1995). *Clustering Competence in Emotional Intelligence: Insights from the Emotional Competence Inventory (ECI).*

Hazlett, R. L., & Hazlett, S. Y. (1999). Emotional Response to Television Commercials: Facial EMG vs. Self-Report. *Journal of Advertising Research, 39*(2), 7–23.

Hudlicka, E. (2003). To Feel or Not to Feel: The Role of Affect in Human-Computer Interaction. *International Journal of Human-Computer Studies, 59*(1–2), 1–32.

Janssens, W., & De Pelsmacker, P. (2005). Advertising for New and Existing Brands: The Impact of Media Context and Type of Advertisement. *Journal of Marketing Communications, 11*(2), 113–128.

Kapoor, A., Burlesonc, W., & Picard, R. W. (2007). Automatic Prediction of Frustration. *International Journal of Human-Computer Studies, 65*, 724–736.

Kim, J. (2008). Emotion Recognition Based on Physiological Changes in Music Listening. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 30*(12), 2067–2083.

Kleinberg, J. (2002). Bursty and Hierarchical Structure in Streams. In *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 91–101). New York: ACM Press.

Luxburg, U. v. (2006). *A Tutorial on Spectral Clustering*. Retrieved October 10, 2008, from http://www.kyb.mpg.de/publications/attachments/Luxburg06_TR_%5B0%5D.pdf

Luxburg, U. v., Bousquet, O., & Belkin, M. (2009). *Limits of Spectral Clustering*. Retrieved October 10, 2008, from http://kyb.mpg.de/publications/pdfs/pdf2775.pdf

Martensen, A., Gronholdt, L., Bendtsen, L., & Jensen, M. J. (2007). Application of a Model for the Effectiveness of Event Marketing. *Journal of Advertising Research, 47*(3), 283–301.

Morris, S. A., & Van der Veer Martens, B. (2008). Mapping Research Special-Ties. *Annual Review of Information Science and Technology, 42*, 213–295.

Orth, U. R., & Holancova, D. (2004). Men's and Women's Responses to Sex Role Portrayals in Advertisements. *International Journal of Research in Marketing, 21*, 77–88.

Pantic, M., & Rothkrantz, L. J. M. (2003). Towards Emotion Recognition in Human Computer Interaction. *Proceedings of the IEEE, 91*(9), 1370–1390.

Pham, M. T., Geuens, M., & Pelsmacker, P. D. (2013). The Influence of Ad-evoked Feelings on Brand Evaluations: Empirical Generalizations from Consumer Responses to More than 1000 TV Commercials. *International Journal of Research in Marketing, 30*, 383–394.

Picard, R. (1997). *Affective Computing*. Cambridge: The MIT Press.

Picard, R. W. (2001). Computers that Recognise and Respond to User Emotion: Theoretical and Practical Implications. *IEEE Transactions on Pattern Analysis and Machine İntelligence, 23*, 1175.

Reeves, B., & Nass, C. (1996). *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Chicago/New York: Center for the Study of Language and Information/Cambridge University Press, 305 p.

Romani, S., Grappi, S., & Dalli, D. (2012). Emotions that Drive Consumers Away from Brands: Measuring Negative Emotions Toward Brands and Their Behavioral Effects. *International Journal of Research in Marketing, 29*, 55–67.

Rousseeuw, P. J. (1987). Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis. *Journal of Computational and Applied Mathematics, 20*, 53–65.

Salton, G., Yang, C. S., & Wong, A. (1975). A Vector Space Model for Information Retrieval. *Communications of the ACM, 18*(11), 613–620.

Schneider, J. W. (2006). Concept Symbols Revisited: Naming Clusters by Parsing and Filtering of Noun Phrases from Citation Contexts of Concept Symbols. *Scientometrics, 68*(3), 573–593.

Schneider, J. W. (2009). Mapping of Cross-Reference Activity Between Journals by Use of Multidimensional Unfolding: Implications for Mapping Studies. In B. Larsen & J. Leta (Eds.), *Proceedings of 12th Inter-national Conference on Scientometrics and Informetrics (ISSI 2009)* (pp. 443–454). Rio de Janiero: BIREME/PAHO/WHO and Federal University of Rio de Janeiro.

Sherer, K. R. (2005). *What Are Emotions? And How Can They Be Measured? Social Science Information*. London/Thousand Oaks/New Delhi: SAGE Publications, 0539-0184, *44*(4), 695–729. doi:https://doi.org/10.1177/0539018405058216.

Small, H. (1980). Co-citation Context Analysis and the Structure of Paradigms. *Journal of Documentation, 36*(3), 183–196.

Tabah, A. N. (1999). Literature Dynamics: Studies on Growth, Diffusion, and Epidemics. *Annual Review of Information Science and Technology, 34*, 249–286.

Teixeira, T., Wedel, M., & Pieters, R. (2012). Emotion-Induced Engagement in Internet Video Advertisements. *Journal of Marketing Research, 49*(2), 144–159.

Vakratsas, D., & Ambler, T. (1999). How Advertising Works: What Do We Really Know? *Journal of Marketing, 63*(1), 26–43.

Wang, W. C., Chien, C. S., & Moutinho, L. (2015). Do You Really Feel Happy? Some Implications of Voice Emotion Response in Mandarin Chinese. *Marketing Letters, 26*(3), 391–409.

White, H. D., & Griffith, B. C. (1982). Authors as Markers of Intellectual Space—Co-citation in Studies of Science, Technology and Society. *Journal of Documentation, 38*(4), 255–272.

White, H. D., & McCain, K. W. (1998). Visualizing a Discipline: An Author Co-citation Analysis of Information Science, 1972–1995. *Journal of the American Society for Information Science, 49*(4), 327–355.

Wierenga, B. (2011). Managerial Decision Making in Marketing: The Next Research Frontier. *International Journal of Research in Marketing, 28*, 89–101.

Yarnold, P. R., & Soltysik, R. C. (2005). *Optimal Data Analysis. A Guidebook with Software for Windows*. Washington, DC: American Psychological Association.

Yarnold, P. R., & Soltysik, R. C. (2016). *Maximizing Predictive Accuracy*. Chicago: ODA Books. 10.13140/RG.2.1.1368.3286.

Yu, P., & Sample, V. (1965). Networks of Scientific Papers. *Science, 169*, 510–515.

Zeng, Z. H. (2009). A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions. *IEEE Transactions on Pattern Analysis and Machine İntelligence, 31*, 39.

6

The Neuroscience Research Methods in Management

Jyrki Suomala

Introduction

The first human body scan by magnetic resonance imaging (MRI) was performed by Professor Raymond Damadian in 1977 in New York. It took almost five hours to produce this one image, and the image was very rudimentary (Mattson and Simon 1996). Despite the fact that the basic idea of MRI used nowadays is the same as Damadian invention, the abundant development of technology, mathematical modelling, and computers has made MRI very effective and credible. In addition, by using functional MRI (fMRI), it is nowadays possible to scan tens of pictures from human brain in one second.

Comparable development has been happened relating to other neuro-scientific tools. These methods currently provide unprecedented access to

J. Suomala (✉)
Laurea University of Applied Sciences, NeuroLab, Finland

L. Moutinho, M. Sokele (eds.), *Innovative Research Methodologies in Management*,
DOI 10.1007/978-3-319-64400-4_6

brain functions, and they have become more accessible to the research community of human behaviour (Ruff and Huettel 2014). These new neuroscientific tools have revolutionized the scientific study of the mind. This development has made neuroscientific methods available for behavioural scientists from psychology and economics to marketing and leadership. In addition, there are now emerging fields of social neuroscience, developmental neuroscience, neuroscience and neuromarketing, all largely because of new neuroscientific tools. Experiments that would have been impossible a decade ago are now available even for undergraduate student projects (e.g. Suomala et al. 2012).

In this chapter, the most essential neuroscientific methods, their benefits and drawbacks as well as their applicability in management context will be described. It has been argued that management is decision-making based on relevant data, information, knowledge and wisdom. Thus, it is necessary to understand about the tools and methods that are key to effective and efficient management. In this vein, the neuroscientific methods will give new and promising avenue to expand current methods in management.

It is paradox that at the same time when the tremendous development of neuroscientific methods has made their use technically easy, this development has made it easier for researcher to make mistakes (Ruff and Huettel 2014). Thus, it is important to understand that different neuroscientific techniques address different aspects of neural function. In addition, the interpretation of given data may strongly depend on what has been measured: brain metabolism, neuronal firing, neurotransmitter levels or some other brain property (Ruff and Huettel 2014).

Human behaviour is guided by two main systems (Bargh et al. 2012; Ackerman et al. 2010). The explicit system is based on conscious and deliberate thinking processes. The implicit system in turn is based on subconscious mind processes. The most important meaning of the neuroscientific methods is that they can uncover subconscious mind processes. This is important for behavioural sciences, because the studies have shown that subconscious mind processes have often more impact to human behaviour than conscious processes (Bargh 2013).

Objectives

The most common neuroscientific methods can be divided into two main categories. The most common biometric methods include eye tracking, face reading, skin conductance and heart rate. This chapter concentrates on these methods; however, pupillometry, voice analysis and respiration measurements will be presented shortly too. Despite it has been expected that in future, hormone-level testing, neurotransmitter testing, optogenetics and even gene testing will be commonplace in the neuromarketing and neuromanagement, these methods have not been included in this chapter. These biometric methods measure changes in the implicit mind system by measuring reactions of autonomic nervous system while a research participant engages in some decision or psychological activity. Neuroimaging methods in turn include electroencephalography (EEG), fMRI, position emission tomography (PET) and magnetoencephalography (MEG). The chapter concentrates on EEG and fMRI, because currently they both are most commonly used methods in neuromarketing/neuromanagement. Neuroimaging methods measure changes in brain function while a subject participates in some decision or psychological activity. Neuroimaging methods can uncover both explicit and implicit mental processes.

A comprehensive and deep description to biometric and neuroimaging methods would go beyond the scope of this chapter. An in-depth understanding of any particular neuroscientific methods would require thousands of hours of training and background knowledge of the neurophysiological structures and processes (Ruff and Huettel 2014). The goal of this chapter is to introduce these methods at an elementary level to the management community. Thus, the main focus of this chapter is to give basic concepts and ideas about how to apply neuroscience in order to solve scientific and practical management problems. Thus, it is primarily intended for those who seek guidance on how to evaluate the strengths and limitations of these methods from management point of view. Despite the fact that the application of neuroscientific methods in management is still rare, the goal is that each method is introduced in conjunction with examples from management. When it is not possible, the examples come in from the nearby areas from neuromarketing and neuroeconomics.

Biometric Methods

People often make decisions without having given them much thought. We have implicit behavioural guidance systems (Bargh et al. 2012), which have often more effects to our decisions than conscious behavioural guidance system. There are two experimental designs in order to uncover mind's subconscious processes. In the priming designs, the presentation time of a stimulus has been restricted in order to limit deliberative thinking of participants. An illustrative example to demonstrate the power of the subconscious priming concentrated on the process of deciding whether a candidate was fit to hold public office (Bargh 2013). In the study, a group of mock voters were given photographs of unknown senator candidates. They saw each photograph only once very quickly and their fleeting glimpses of each portrait served as an accurate proxy for the later choices of actual voters in those states (Bargh 2013). In the distractive designs, participants concentrate to demanding task, in which they have no mental resources to respond to other secondary stimuli in the experiment. An illustrative example to demonstrate the effect of secondary stimuli for human decision is the study, in which participants needed to read text (Shapiro et al. 1997) when at the same time the advertisement of product is placed outside of their field of vision. Unexpectedly, this subconscious advertisement increased participants' preference of the products presented on the outside of their vision field. The biometric methods are suitable for studies of both types of subconscious mental processes.

Eye Tracking

Contrary to our intuition, our visual image is not akin to photograph taken by the camera of our eyes. On the contrary, it is construction of many different images based on a constant and almost non-conscious dynamic visual search (Bridger 2015). The most sensitive and exact area in our eyes is the fovea. However, this part in the retina is very small. Surrounding of the fovea is area called peripheral vision. The further from a person's foveal position something is the weaker the visual perception of a person is (Bridger 2015). From retina, the part of the visual information goes to the visual cortex. Most of the saccades and fixations of the eye are guided by automatic processes of the nervous system; however, an individual can control the eye movements consciously, too.

Thus, when we look at a scene or read, we continually make eye movements called saccades (Rayner 1998). A saccade is the switch of the eyes from one pause to another. Between the saccades, our eyes remain still during fixations for 200–500 ms (Zurawicki 2010). Eye trackers are useful tools for the measuring eye's saccades and fixations when a participant sees videos or TV, reads or looks for still pictures. In addition, it is possible to measure the change of pupil's size (pupillometry) during experiments with eye trackers. Pupillometry is based on the idea that the pupil can become larger or smaller depending on light conditions, but also on the person's emotional interest and cognitive load (Bridger 2015).

Currently, there are several manufacturers of eye trackers. Traditionally, eye-tracking data has been collected in laboratory; however, currently it is possible to measure eye tracking online and in real-like environments with mobile device. Most eye-tracking software can perform automatic analyses of researcher's data; thus, basic analysis is easy to do. The most frequently used output that can be generated from eye-tracking data in advertising, marketing and management is the heat map (Bridger 2015). Heat maps can be computed in three ways. It can be based on the number of fixations, the total gaze duration on an area or on the number of people who looked at the area. The second way to present the results of eye-tracking study is the gaze plot (Bridger 2015). It shows the typical patterns of the order of fixations on the image or video.

Resume Review Example

The Ladders (Inc.) company applied eye tracking in order to clarify how professional recruiters (n = 30) review job seekers' resumes and personal profiles (Evans 2012). After reviews, they made initial decision about candidate's potential for the job (fit/no fit decision). All information was on the screen, thus it was possible to measure with eye tracking where and how long a participant focuses when she/he digested information from resume. The eye-tracking data showed that almost 80% of the gaze duration (total fixation time) of the recruiters resume reviews hit on the following six data points: name, current title/company, previous title/

company, previous position start and end dates, current position start and end dates and finally, education. Thus, the decisions were based mostly on the six pieces of data listed above. Other pieces of information of candidates became filler and had little to no impact on the initial decision making (Evans 2012). The researcher of the study asked before eye-tracking phase that how much time they spend for reviewing an individual resume. Based on this self-reported data, they spend four to five minutes for this review. However, the eye-tracking study found unexpectedly that recruiters spend only six seconds reviewing an individual resume on their initial fit/no fit decision (Evans 2012).

Benefits and Drawbacks of Eye Tracking

Eye tracking is probably the oldest and most common method used in consumer science and marketing. It is comparably cheap and requires no complex knowledge to produce results. By online and mobile eye tracking, it is possible to make eye-tracking studies in real marketing, workplace and management contexts. The eye-tracking data shows where and how long someone looked; however, based on this method, it is not possible to explain why someone looks. The most viewed products do not necessarily sell so much than less viewed products (Iyengar and Lepper 2000). An individuals' gaze may have stayed on a particular area because they found hard to understand or simply because they were resting their eyes there whilst thinking what to do next (Bridger 2015). Despite the fact that it can be useful to see where and how long a person's eyes are fixated, the most critical aspect is how effective our message is in order to change the behaviour of people under considerations. By using eye tracking, it is not possible to expect behavioural change of human subjects (Venkatraman et al. 2012); thus, it is important to use eye tracking with other neuromarketing (neuromanagement) tools.

Face Reading

In the 1970s, the psychologist Paul Ekman and his colleagues discovered six primary emotional expressions of humans' faces (Ekman and

Friesen 1971). These emotions are anger, disgust, fear, happiness, sadness and surprise. Later, the contempt has been added to the basic emotional expressions (Genco et al. 2013). These primary emotional expressions seem to be universal across different cultures, and it is possible to measure these expressions by face reading. The first face reading technologies (facial electromyography) were based on direct measurements with sensors on the person's face, measuring the electrical activity generated when muscles become active. Today, face reading technology has been developed to allow automated facial coding via a camera and software. Thus, face reading is an easy and a relatively cheap tool to install in customers' and consumers' home laptops in order to test their basic emotional responses to different materials in real and online environments. In addition to basic emotions by using the face reading technology, it is possible to measure the movements of faces' micromuscles.

Internet Page-Loading Experience Example

Web users are often impatient and fickle, and they want pages to load as soon as possible. Websites are becoming increasingly image- and information-rich. Thus, it is critical for companies to know how much the users emotionally engage to the website (Bridger 2015). The Radware company has developed specific image-loading technique called "Progressive Image". They needed to know which were the users' emotional engagements when they used Radware's loading methods compared with other methods. In order to understand how emotional engagement moved over time in response to the web pages loading using the different image-loading techniques, Darren Bridger with his colleagues conducted face reading—study for Radware (Bridger 2015). They measured second-by-second happiness of the participants and cute and funny images by using face reading. The pictures and videos came from three different webpages. First, YouTube, listing funny cat videos, second, a clothing site with images of cute toddlers, and finally, a greetings card website with cards depicting funny animals. They used different image-loading methods and found that Radware's "Progressive Image" techniques made both men and women smile sooner in the page-loading experience and for longer period (Bridger 2015).

Benefits and Drawbacks of Face Reading

By using face reading, it is possible to measure basic emotional states of participants. This information reflects subconscious emotional reactions to materials showed to the participants. Bridger (2015) argues that the heartland of the use of face reading is emotive videos. In this way it is possible to give a moment-by-moment trace of emotional responses in order to test TV ads or movie trailers. It is also possible to measure evolving emotional responses over time, for example, the journey through websites. In addition, the face reading technology is currently cheap, easy and fast to use. Thus, the main use of face reading has been in studying TV advertisements, because by using this method, it is possible to give fast feedback on the moment-to-moment performance of an advertisement to guide fine-tuning editing (Bridger 2015).

The limitations of the face reading is that despite the emotional facial expression tends to reflect associated internal emotional states, sometimes people can experience internal emotional states without displaying the accompanying facial expressions (Genco et al. 2013). In addition, humans' faces react in response to the emotional intensity of a stimulus called emotional arousal, not its emotional direction. Although it is safe to conclude that positive emotions relating to materials or marketing messages under considerations are prerequisite for messages' success in the real-life contexts, it is still unclear whether this is a sufficient condition for success. Thus, we do not know how well the outcomes of facial reading predict peoples' behaviour in real-life context.

Skin Conductance

Skin conductance is a method that measures the action of nervous system from sweat glands, from hands or from the soles of feet (Bridger 2015; Genco et al. 2013). The palms of an individual's hand could become noticeably sweaty when she/he is nervous. However, with skin conductance method, it is possible to measure the increased activation of sweat glands before noticing sweat. Thus, when the levels of sweat increases, the change of electrical conductivity in the skin is the first indicator of this change, measured by skin conductance (Bridger 2015). In typical skin conductance study, two electrodes are placed on the person's hand.

Brand Preference Study Example 1

In the brand preference study (Walla et al. 2011), 29 participants evaluated 300 common brand names online based on the dislike-like dimension. Based on this online survey, ten most liked and most disliked brand names were chosen as stimuli for physiological experiment. By measuring skin conductance when participants saw the most liked and most disliked brand names, they found a significantly reduced skin conductance value associated with visual presentations of liked brand names compared to disliked brand names (Walla et al. 2011). Thus, in the study (Walla et al. 2011), skin conductance was significantly reduced in case of viewing liked brand names compared to viewing disliked brand names. Thus, the visual perception of liked brand names elicited a more relaxing state compared to disliked brand names.

The Benefits and Drawbacks of Skin Conductance

The benefit of skin conductance is that the data collection and analysis are easy to do. In addition, it is a cheap biometric method compared to other technologies. Whereas most of the researchers argue that skin conductance is primarily a measure of arousal and does not differentiate between positive and negative valence (Critchley et al. 2000), previously described example (Walla et al. 2011) indicates that it may also be a promising objective indicator of valence and brand preference. Skin conductance is perhaps most relevant to marketing studies where it is important to measure the ongoing responses to longer videos, TV shows and movies (Bridger 2015).

Heart Rate

The heart activity contains a surprising amount of information about emotional arousal and attentional focus. In general, when someone's focus is outwards, their heart rate slows, and when it is inward, it quickens (Bridger 2015). Heart rate is the beating speed of the heart, and it can be an indicator of various physiological reactions and cognitive and physical effort (Genco et al. 2013). Usually, when attention increases, heart rate has been found to slow down in the short term. In these cases, the deceleration is associated with an orienting response. In addition, when an individual is performing some internal complex mental task, heart rate has been shown to speed up (Bridger 2015). Respectively, heart rate has been found to speed up in the long term when experiencing emotional arousal. This emotional arousal is associated with a defensive response. Also, heart rate may go up when an individual watches TV advertisements and when there are changes of scene in the ad.

Despite the fact that heart rate measurement is cheap and easy to use, the interpretation of the heart rate as an indicator of peoples' emotional state is still contradictory. It has been found that in the short-term, deceleration of heart rate has been found to be associated with both positive and negative emotions, whereas in the longer time period, positive stimuli speed up heart rate and negative stimuli decrease heart rate (Genco et al. 2013). Thus, it is important to use heart rate with other behavioural, biometrical and neuroimaging methods.

Brand Preference Study Example 2

In the previous skin conductance chapter, the example of Walla et al. (2011) has been presented. They measured heart rate in the same experiment too and found a strong trend towards a significantly reduced heart rate associated with liked brand names compared to disliked brand names. Thus, brand preferences have associations to low heart rate level. In the same study, they also used heart rate measurement. They found a strong trend towards a significantly reduced heart rate associated with visual presentations of liked brand names compared to disliked brand names. Thus, the study showed that decreased heart rate is an indicator for positive preference.

Benefits and Drawbacks of Heart Rate

Heart rate measurement is very sensitive for contextual and emotional factors, and the outcome data is difficult to interpret. However, it is an easy and cheap method to use, and it is typical tool to use in combination with other metrics.

Other Biometric Methods

Whilst previously described biometric methods are the most frequently used in neuromarketing and are thus most suitable methods for neuromanagement, there are several other measures that are also sometimes used.

Pupillometry measures the size and changes in the pupil's size in the eye. Pupillometry measurements are usually byproduct of most eye-tracking system. Thus it is possible to collect pupillometry data at the same time with eye tracking. It is a reliable and sensitive measure of emotional and cognitive reactions to stimuli in real time (Genco et al. 2013). Voice analysis focuses on the emotional sound patterns of the peoples' voices, and these patterns are amenable to automated computer analysis. Despite the fact that this analysis is not currently a mainstream neuromarketing method, advantage of voice analysis is that it can be done online without any requirement of the additional overheads of lab-based studies (Barden 2013).

Respiration rate is an indicator of voluntary or involuntary emotional arousal, and it can be measured with a chest strap. This measure is not used often in neuromarketing research.

In future, hormone level, neurotransmitter, and even gene testing will be spread as part of the neuroscientific tools, and their applications to the neuromarketing and neuromanagement will be very promising.

Biometric: Conclusion

The benefits of biometrics are that they can provide tools to measure a person's emotional responses to TV-, video-, online-, print- and other advertisements and messages. Especially these methods are capable to provide a moment-by-moment reading of person's emotional responses online, on video and on TV materials. These methods are comparatively cheap to run, and data is easy to analyse (Bridger 2015; Genco et al. 2013). Despite the fact that most of studies have been conducted by bringing people into a certain location, they have potential for being used in many real-world situations when technology develops more. In addition, the benefits of biometric methods are that they are easy to combine with other measures such as biometric and neuroimaging methods.

The biometric tools measure levels of arousal, emotional valence or emotional engagement rather than preferences itself. By using these tools, it is possible to lop off stimuli, which produce negative emotions. However, positive emotions do not necessarily mean that people change their behaviour according the message/advertisements. Thus, it is important to evaluate whether these biometric measures actually translate into real-life advertising success (Venkatraman et al. 2015). Although there has been considerable application in companies using biometric measures to better understand consumers' responses to new brands, services, products and advertisements, there is not clear evidence, which is the relationship between arousal, emotional valence and emotional engagement to the real-life success. After describing neuroimaging methods, we will concentrate on these issues more specifically in the conclusion chapter.

Neuroimaging Methods

The major work of the brain is processing and transfer of information. The generation of action potential is the main type of neuronal information processing, and the neuroimaging methods measure the signals of an action potentials' changes from the brain. When a neuron fires, it sends a signal called action potential down its axon to one or more other neurons. The action potential evokes the release of neurotransmitters at the synapse, and this process causes its membrane potential to change more positive or negative (Ruff and Huettel 2014). These processes require substantial energy in the form of oxygen and sugars. Despite the fact that the brain represents only about 3% of the mass of the human body, over 20% of the oxygen and sugar we employ as humans are used in the brain (Glimcher 2014). Those energy consumptions themselves are not involved in neuronal signalling, but they serve as important markers that information processing has increased within a specific brain region or brain networks in the brain (Ruff and Huettel 2014).

The action potential occurs in the two forms of signals that scientists have exploited in their efforts to get a better look at what is going on inside the brain. The first is electrical activity in the brain and the second is blood flow in the brain (Genco et al. 2013). The action potentials generate short coherent electrical changes in the neurons that can be measured via detectors on the scalp with EEG. In other words, EEG measures patterns of rhythmic electrical activity arising mainly from the cortex. The measurement of blood flow instead is based on fMRI. It measures the blood oxygen level-dependent (BOLD) signal, which is a measure of the ratio of oxygenated to deoxygenated haemoglobin (Ashby 2011). This measurement is based on the principle that active brain areas consume more oxygen and sugar than do inactive areas (Ashby 2011).

Electroencephalography (EEG)

The signals to a neuron change the electrical potential of its cell membrane and could produce action potentials. When many neurons express similar action potentials and if those neurons share a similar spatial location, then the coherent electrical current can be detected by electrodes positioned on the scalp (Ruff and Huettel 2014). The EEG can measure these electrical current.

Typical EEG studies record changes in electrical potential using 64, 128 or even 256 electrodes positioned on the scalp. A common research approach in modern consumer neuroscience and neuromarketing is to investigate how marketing materials or advertisements influence the amplitude of the electrodes positioned on the scalp.

A conventional way to interpret the electrical potential is to classify commonly observed signals into frequency bands named after Greek letters. Frequency bands describe neurons mean firing rate during one second (Hertz [Hz]). Delta bands mean that the frequency is 4 Hz or less. This is typical in dreamless sleep (Genco et al. 2013). Theta bands refer to the 4–8 Hz activation, and this is associated with internally focused information processing, such as memory activation. Alpha is activation between 8 and 12 Hz. This is typical for the brain's default frequency, dominant when the brain is in a relaxed state, like when eyes are closed. Beta bands refer to 13–30 Hz. This activation type is associated with active attention and alertness. In addition, the frontal beta bands have associations with reward processing (Boksem and Smidts 2015). Finally, Gamma activation (more than 30 Hz) is associated with emotional processing and learning (Genco et al. 2013).

The metrics power and coherence are commonly used to measure brain-wave frequencies. Power measures the degree of activity within a particular frequency band over a specified period of time (Genco et al. 2013). Coherence measures the consistency of brain-wave frequencies across different regions of the brain. Thus, EEG provides non-invasive access to the electrical activity of the brain and is a common choice for scientific laboratories and companies which make neuromarketing and neuromanagement studies (Ruff and Huettel 2014).

Movie Sale Example

Whereas traditional behavioural measurements of consumers' preferences and predicting their future behaviour are very limited (Dijksterhuis et al. 2006), the hope that neuroimaging methods will give essential adding value for consumer neuroscience has been increased (Ariely and Berns 2010). In order to test the benefits of neuroimaging as predictive tool in consumer neuroscience, Boksem and Smidts (2015) obtain both stated preference measures and neural measures by the EEG in response to movie trailers to probe their potential to provide insight into consumers' individual preferences as well as movie sales in the general population. Thus, they investigate whether neural measures could make contribution to the prediction of commercial success beyond traditional behavioural measures under naturalistic viewing conditions (Boksem and Smidts 2015).

Dutch University students (n = 29) viewed 18 movie trailers in random order while their EEG was recorded. Each trailer lasted 2.5 to 3 minutes. After the participant viewed the trailer, she/he indicated how much she/he liked the movie and how much she/he would be willing to pay for the whole movie DVD. In this way, it was possible to measure brain reactions to movie trailer and state preference for the movie. Movies were selected on the basis of their US box office results, and the researchers constructed a set of movies that varied considerably in commercial success. The range of the success of the movies these trailers promoted was between $4.4 million and $121 million. In addition, the movies which an individual has seen previously have been excluded from trailer set in experiment. Thus, each subject only saw trailers, which were unknown to them previously (Boksem and Smidts 2015).

The results showed that high beta activity (16–18 Hz) during viewing of the trailer was related to a high stated preference for that movie. On the group level, the EEG data revealed a cluster of EEG activity that was significant predictor of US box office in the gamma range (60–100 Hz). Thus, the higher this gamma activity in the participants during the viewing of the trailers, the more money this movie generated at the box office (Boksem and Smidts 2015). The researchers (Boksem and Smidts 2015) argued that these results indicate that there is unique information in the EEG measures that is not captured by traditional behavioural measurements. They emphasized that including both EEG and behavioural measurement significantly improved prediction of stated preference. In addition, the higher the frontal gamma activity in participants during the viewing of the trailers, the more money this movie generated at the box office (Boksem and Smidts 2015). Thus, the EEG gamma activity in response to viewing a trailer for a particular movie trailer significantly enhances predictions of commercial success for movies that these trailers promoted. The study importantly showed that based on small sample's brain activation profile during viewing advertisements (in this case trailers), it is possible to predict the success of this product (in this case movie) in the real market.

Benefits and Drawbacks of EEG

Whereas biometric methods can produce information about emotional engagement, valence and attitudes of the participants to the stimuli under considerations, the EEG could get unique data, which can be used as predicting success of the products in the real market.

The EEG machinery is readily available from multiple vendors, and it is only moderately expensive. In addition, the EEG is a non-invasive and silent technology that is directly sensitive to neuronal activity (Genco et al. 2013). The time resolution of the EEG is excellent, and it can record a voltage every one to three milliseconds almost at the speed of cognition. In addition to its high-time resolution, the EEG allows for relatively naturalistic viewing condition, in which participants can be seated in comfortable chair while viewing a large screen with high-quality audio (Boksem and Smidts 2015). Thus, it is a suitable tool for neuromarketing and neuromanagement. As Ruff and Huettel (2014) state, more researches have been published using the EEG than any other neuroimaging or biometric researches.

A significant limitation of EEG measurement compared to fMRI is that the EEG cannot localize activity inside the brain with the same precision as fMRI. The EEG signal expands throughout the brain but is attenuated greatly by the skull and the scalp, and the activity recorded by scalp electrodes could have been generated by any of an infinite number of potential sources (Ruff and Huettel 2014). Whereas the brain networks below cortex have essential role in the human choice (Levy and Glimcher 2012), the EEG could not capture these signals in the exact way. However, as the example above (Boksem and Smidts 2015) showed, there are promising results based on the EEG, which show that in the future, it is possible to predict customers' choice based on the EEG signals.

Functional Magnetic Resonance Imaging (fMRI)

A mental activity increases demand for oxygen and sugar in regions in the brain that participate in a mental activity, and this need is met by increased blood flow to a region (Genco et al. 2013). FMRI measures the BOLD signal, which varies by region in the brain because blood delivered to an active brain region requires more oxygen than blood delivered to an inactive region, and this oxygenated blood produces a stronger magnetic field than non-oxygenated blood (Genco et al. 2013). Whereas the fMRI scanner generates a very strong static magnetic field, it can reveal the changes of blood flows when a participant is lying inside a large camber. FMRI provides researchers the opportunity to study neural activity in the human brain almost real time (Ashby 2011). It is no wonder that fMRI has grown to become the dominant measurement technique in the cognitive neuroscience and neuroeconomics (Ruff and Huettel 2014).

Health Message Example

Whereas most of the neuroimaging studies try to find associations between different stimuli and the brain activation patterns (Heinonen et al. 2016; Plassmann et al. 2008), recent brain-as-predictor approach has demonstrated that brain measures can predict purchasing decisions, clinical outcomes and many other domains over longer timescales in ways that go beyond what was previously possible with behavioural data alone (Berkman and Falk 2013). The intervention study (Falk et al. 2012) asked whether neural responses of a small group of smokers can predict the quitting of smoking on population level.

The researchers used ads from three campaigns in fMRI investigation conducted in a separate location where the ads were aired (Falk et al. 2012). Participants' (n = 30) task during the fMRI session consisted of viewing three TV ad campaigns (A, B and C) which were designed to help smokers quit smoking. Each campaign comprised three separate ads, and each campaign ended by displaying the National Cancer Institute's Smoking Quitline phone number. After the fMRI session, participants ranked the effectiveness of the ads they viewed in the scanner behaviourally.

Based on behavioural data, the superiority order of the ads was B > A > C. In addition, industry experts who were familiar with the campaigns ranked the ads in similar ways. However, the participants' mean neural activity in the brain's valuation and self-referential region, namely Medial Prefrontal Cortex (MPFC), during ad exposure suggested campaign order as C > B > A (Falk et al. 2012). Probably, the reason for the success of the campaign C was that it included more self-affirmative and self-referential material than B and A campaign.

The population-level success of each advertising campaign has been compared with Quitline call volume in different state. Thus, the fMRI sample and real advertising campaign were in different states. Unexpectedly, the superiority order of the ads on population level was C > B > A. Thus, this intervention study (Falk et al. 2012) shows that the average and most frequently observed neural responses in the MPFC correctly ordered the success of the ad groups at population level, whereas participants' self-reports and evaluations of industry experts did not. Thus, the activation pattern in the MPFC during viewing messages of advertisements predicts what people will do in future.

Benefits and Drawbacks of fMRI

The main benefit of the fMRI measure is that its spatial resolution is excellent. In this way, it is possible to get critical data associated with human behaviour, which is otherwise not possible to do. As previous intervention study showed, our brains have hidden signals which have essential effect for our purchase (Berns and Moore 2010), healthy (Berkman and Falk 2013) and even social behaviour (Falk et al. 2013) under conscious level. Currently, it is possible to uncover these signals by using fMRI.

However, there are also many limitations relating to the fMRI. First, its time resolution is not very good compared to EEG. On practical level, this means that data analysis is difficult and time-consuming. However, today it is easier to automatize the data analysis processes. Thus in future, the fMRI study will be easier to do on practical level. Second limitation is that measurement situation is not very natural for a participant. The participants have to keep very still and lay supine in the scanner in unnaturalistic environment. Finally, the fMRI measurement is expensive, and this could be the reason why so few companies use the fMRI in neuromarketing and neuromanagement.

Other Neuroimaging Methods

In addition to the EEG and fMRI, there are other neuroimaging methods. Because they are not so commonly used, they will be described only shortly here. MEG measures brain activity via magnetic field changes by using external sensors near participants' skull. It has the same temporal properties as EEG (Ruff and Huettel 2014).

PET measures the motion of a radioactive isotope in the brain when a participant participates in the experiments. This radioactive isotope has been injected to a participant's blood before experiment, and this isotope will become more prevalent in the brain areas with increased sugar metabolism (Ruff and Huettel 2014).

Functional near-infrared spectroscopy (fNIRS) measures the hemodynamic response to the cerebral cortex using near-infrared light (Naseer and Hong 2013). Whereas fMRI can measure the hemodynamic activation from whole brain, the signals are detected by fNIRS from the cerebral cortex only. Because fNIRS is more practical to use than EEG or fMRI, it has very promising tool to measure the brain activation in the real-life context like when a participant uses computer (Naseer et al. 2014). However, this technology is new, and there are only a few scientific studies based on this technology.

Neuroimaging: Conclusion

EEG and fMRI are the most widely spread neuroimaging methods in neuropsychology, consumer neuroscience and neuromarketing. Thus, they have a lot of potential applications to the management, too.

The benefits of these methods are that they can detect both emotional signals and preference signals from the brain. The new brain-as-prediction approach has given new opportunities to optimize advertisements, healthy and other campaigns by using neuroimaging. In addition, this approach gives a new opportunity to optimize new innovative products and services (Ariely and Berns 2010).

EEG and fMRI are not so practical to use compared with biometric methods. In addition, the use of EEG and fMRI is more expensive than the use of biometric methods. In future, when the technology develops, there is very high probability that the applications of EEG and fMRI will increase in management.

Summary

This chapter has looked at a number of neuroscientific methods that underpin how scientists, marketers and managers are using this data in order to better explain human behaviour. However, each of them has distinct and often complementary strengths and weakness (Ruff and Huettel 2014). For instance, biometric methods can uncover subconscious emotional responses—emotional arousal, emotional valence and emotional engagement—in a relatively cheap way. The data collection and analysis are easy and fast to do. By using biometric, it is possible to optimize marketing and management messages and advertisements by looping elements, which have associations to the strong negative emotions.

Neuroimaging methods are more expensive and difficult to use than biometrics. However, by using neuroimaging, it is possible to find a hidden information in the brain that can help scientist, marketers and managers to figure out things that are hard to figure out otherwise (Falk et al. 2015).

Converging evidence demonstrates that activity within brain's valuation network can predict individual behaviour in response to marketing and other messages. In addition, growing evidence also suggests that neural activity in this valuation network can forecast population-level campaign outcomes above and beyond what is explained by participants' self-reports of their attitudes towards the messages in question and their intentions to change their behaviour (Falk et al. 2010, 2015; Venkatraman et al. 2015).

One way to overcome the limitations of each neuroscientific method is to combine several complementary research techniques in order to provide more explanatory power of the datasets (Ruff and Huettel 2014). For instance, it is possible to use EEG and fMRI at the same time. The benefit of this combination is that it is possible to get simultaneously both exact spatial (by fMRI) and temporal (by EEG) data from the brain. It is also possible to use eye tracking simultaneously with neuroimaging techniques and other biometric techniques (Ruff and Huettel 2014). In addition, it is often a routine protocol to use heart rate measurement during fMRI experiments (Heinonen et al. 2016; Suomala et al. 2012) in order to control participant's activity level during experiment.

Whereas management is decision-making based on the best relevant data, information and knowledge of the neuroscientific methods will give new and promising avenues to expand current methods in management. The hope is that this chapter has illustrated the main neuroscientific techniques suitable for management community.

References

Ackerman, J. M., Nocera, C. C., & Bargh, J. A. (2010). Incidental Haptic Sensations Influence Social Judgments and Decisions. *Science, 328*(5986), 1712–1715. https://doi.org/10.1126/science.1189993.

Ariely, D., & Berns, G. S. (2010). Neuromarketing: The Hope and Hype of Neuroimaging in Business. *Nature Reviews. Neuroscience, 11*(4), 284–292. https://doi.org/10.1038/nrn2795.

Ashby, F. G. (2011). *Statistical Analysis of fMRI Data*. Cambridge, MA: MIT Press.

Barden, P. (2013). *Decoded: The Science Behind Why We Buy* (1st ed.). Chichester: Wiley.

Bargh, J. A. (2013). Our Unconscious Mind. *Scientific American, 310*(1), 30–37. https://doi.org/10.1038/scientificamerican0114-30.

Bargh, J. A., Schwader, K. L., Hailey, S. E., Dyer, R. L., & Boothby, E. J. (2012). Automaticity in Social-Cognitive Processes. *Trends in Cognitive Sciences, 16*(12), 593–605. https://doi.org/10.1016/j.tics.2012.10.002.

Berkman, E. T., & Falk, E. B. (2013). Beyond Brain Mapping Using Neural Measures to Predict Real-World Outcomes. *Current Directions in Psychological Science, 22*(1), 45–50. https://doi.org/10.1177/0963721412469394.

Berns, G., & Moore, S. E. (2010). *A Neural Predictor of Cultural Popularity* (SSRN Scholarly Paper ID 1742971). Rochester: Social Science Research Network. http://papers.ssrn.com/abstract=1742971

Boksem, M. A. S., & Smidts, A. (2015). Brain Responses to Movie Trailers Predict Individual Preferences for Movies and Their Population-Wide Commercial Success. *Journal of Marketing Research, 52*(4), 482–492. https://doi.org/10.1509/jmr.13.0572.

Bridger, D. (2015). *Decoding the Irrational Consumer: How to Commission, Run and Generate Insights from Neuromarketing Research* (Marketing Science Series). London/Philadelphia: Kogan Page.

Critchley, H. D., Elliott, R., Mathias, C. J., & Dolan, R. J. (2000). Neural Activity Relating to Generation and Representation of Galvanic Skin Conductance Responses: A Functional Magnetic Resonance Imaging Study. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 20*(8), 3033–3040.

Dijksterhuis, A., Bos, M. W., Nordgren, L. F., & van Baaren, R. B. (2006). On Making the Right Choice: The Deliberation-Without-Attention Effect. *Science, 311*(5763), 1005–1007. https://doi.org/10.1126/science.1121629.

Ekman, P., & Friesen, W. V. (1971). Constants Across Cultures in the Face and Emotion. *Journal of Personality and Social Psychology, 17*(2), 124–129. https://doi.org/10.1037/h0030377.

Evans, W. (2012). Eye Tracking Online Metacognition: Cognitive Complexity and Recruiter Decision Making. *The Ladders, 1*(1), 5.

Falk, E. B., Berkman, E. T., Mann, T., Harrison, B., & Lieberman, M. D. (2010). Predicting Persuasion-Induced Behavior Change from the Brain. *The Journal of Neuroscience, 30*(25), 8421–8424. https://doi.org/10.1523/JNEUROSCI.0063-10.2010.

Falk, E. B., Berkman, E. T., & Lieberman, M. D. (2012). From Neural Responses to Population Behavior Neural Focus Group Predicts Population-Level Media Effects. *Psychological Science, 23*(5), 439–445. https://doi.org/10.1177/0956797611434964.

Falk, E. B., Morelli, S. A., Locke Welborn, B., Dambacher, K., & Lieberman, M. D. (2013). Creating Buzz the Neural Correlates of Effective Message Propagation. *Psychological Science, 24*(7), 1234–1242. https://doi.org/10.1177/0956797612474670.

Falk, E. B., O'Donnell, M. B., Tompson, S., Gonzalez, R., Cin, S. D., Strecher, V., Cummings, K. M., & An, L. (2015, September). Functional Brain Imaging Predicts Public Health Campaign Success. *Social Cognitive and Affective Neuroscience*, nsv108. doi:https://doi.org/10.1093/scan/nsv108.

Genco, S. J., Pohlmann, A. P., & Steidl, P. (2013). *Neuromarketing for Dummies*. Mississauga: John Wiley & Sons.

Glimcher, P. W. (2014). Introduction to Neuroscience. *Neuroeconomics*, 63–75. Elsevier. http://linkinghub.elsevier.com/retrieve/pii/B978012416008800005X.

Heinonen, J., Numminen, J., Hlushchuk, Y., Antell, H., Taatila, V., & Suomala, J. (2016). Default Mode and Executive Networks Areas: Association with the Serial Order in Divergent Thinking (E. A. Stamatakis, Ed.). *PLOS ONE, 11*(9), e0162234. doi:https://doi.org/10.1371/journal.pone.0162234.

Iyengar, S. S., & Lepper, M. R. (2000). When Choice Is Demotivating: Can One Desire Too Much of a Good Thing? *Journal of Personality and Social Psychology, 79*(6), 995–1006. https://doi.org/10.1037//0022-3514.79.6.995.

Levy, D. J., & Glimcher, P. W. (2012). The Root of All Value: A Neural Common Currency for Choice. *Current Opinion in Neurobiology, 22*(6), 1027–1038. https://doi.org/10.1016/j.conb.2012.06.001.

Mattson, J., & Simon, M. (1996). *The Pioneers of NMR and Magnetic Resonance in Medicine: The Story of MRI*. Ramat Gan/Jericho: Bar-Ilan University Press; published in the U.S.A. by Dean Books Co.

Naseer, N., & Hong, K.-S. (2013). Classification of Functional Near-Infrared Spectroscopy Signals Corresponding to the Right- and Left-Wrist Motor Imagery for Development of a Brain–Computer Interface. *Neuroscience Letters, 553*(October), 84–89. https://doi.org/10.1016/j.neulet.2013.08.021.

Naseer, N., Hong, M. J., & Hong, K.-S. (2014). Online Binary Decision Decoding Using Functional Near-Infrared Spectroscopy for the Development of Brain–Computer Interface. *Experimental Brain Research, 232*(2), 555–564. https://doi.org/10.1007/s00221-013-3764-1.

Plassmann, H., O'Doherty, J., Shiv, B., & Rangel, A. (2008). Marketing Actions Can Modulate Neural Representations of Experienced Pleasantness. *Proceedings of the National Academy of Sciences, 105*(3), 1050–1054. https://doi.org/10.1073/pnas.0706929105.

Rayner, K. (1998). Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin, 124*(3), 372–422. https://doi.org/10.1037/0033-2909.124.3.372.

Ruff, C. C., & Huettel, S. A. (2014). Experimental Methods in Cognitive Neuroscience. *Neuroeconomics,* 77–108. Elsevier. http://linkinghub.elsevier.com/retrieve/pii/B9780124160088000061

Shapiro, S., MacInnis, D. J., & Heckler, S. E. (1997). The Effects of Incidental Ad Exposure on the Formation of Consideration Sets. *Journal of Consumer Research, 24*(1), 94–104. https://doi.org/10.1086/209496.

Suomala, J., Palokangas, L., Leminen, S., Westerlund, M., Heinonen, J., & Numminen, J. (2012, December). Neuromarketing: Understanding Customers' Subconscious Responses to Marketing. *Technology Innovation Management Review*, *2*, Recent Research, 12–21.

Venkatraman, V., Clithero, J. A., Fitzsimons, G. J., & Huettel, S. A. (2012). New Scanner Data for Brand Marketers: How Neuroscience Can Help Better Understand Differences in Brand Preferences. *Journal of Consumer Psychology, 22*(1), 143–153. https://doi.org/10.1016/j.jcps.2011.11.008.

Venkatraman, V., Dimoka, A., Pavlou, P. A., Vo, K., Hampton, W., Bollinger, B., Hershfield, H. E., Ishihara, M., & Winer, R. S. (2015). Predicting Advertising Success Beyond Traditional Measures: New Insights from Neurophysiological Methods and Market Response Modeling. *Journal of Marketing Research, 52*(4), 436–452. https://doi.org/10.1509/jmr.13.0593.

Walla, P., Brenner, G., & Koller, M. (2011). Objective Measures of Emotion Related to Brand Attitude: A New Way to Quantify Emotion-Related Aspects Relevant to Marketing (W. El-Deredy, Ed.). *PLoS ONE, 6*(11), e26782. doi:https://doi.org/10.1371/journal.pone.0026782.

Zurawicki, Leon. 2010. *Neuromarketing*. Berlin/Heidelberg: Springer Berlin Heidelberg. http://link.springer.com/10.1007/978-3-540-77829-5

7

Benefits of Neuromarketing in the Product/Service Innovation Process and Creative Marketing Campaign

Jyrki Suomala

Introduction

Innovation is a novel idea that has been implemented and has been profitable from commercial or peoples' welfare point of view. It consists of the generation of a new idea (the fuzzy front end (FFE) of innovation and its implementation as a new product, service, or entertainment shows), leading to the dynamic growth of the enterprise and profit creation (Taatila et al. 2006; Suomala et al. 2006). New idea generation is creative and uncertain and is often called the FFE of innovation (Santonen and Hytönen 2015). The capability to create innovations is one of the most critical factors for companies and organizations in today's highly competitive marketing environment. The companies spend millions of dollars to develop and create new products, services, and entertainment shows for the market. Therefore, it is not surprising that each year tens of thousands of new innovative products, services and TV shows enter into

J. Suomala (✉)
Laurea University of Applied Sciences, NeuroLab, Finland

L. Moutinho, M. Sokele (eds.), *Innovative Research Methodologies in Management*,
DOI 10.1007/978-3-319-64400-4_7

market. In the same time, companies pay millions of dollars to advertising and creative agencies, which try to produce creative campaigns in order to increase consumers' willingness to buy these new goods.

Whereas the range of products and services increases constantly, over half of new products fail commercially (Iyengar 2011), less than a quarter of new show episodes become full shows, and majority of those shows are canceled within the first years (Berkman and Falk 2013). In the similar way, companies spend millions of dollars each year in order to create successful advertising campaigns (Venkatraman et al. 2015). However, it is not guaranteed that the advertising campaigns are successful. For example, Republican candidate Jeb Bush spent over $80 million on TV and radio advertising to be nominated for the US presidential elections , whereas Donald Trump spent only $11 million in the one-week period, March 28–April 5 (Galloway 2016). However, Trump was elected as the Republican candidate for the US presidential elections.

Usually, the new innovation processes and advertising optimizing include pretesting copy alternatives of new product concepts, prototypes, and marketing campaign ideas (Venkatraman et al. 2012; Falk et al. 2013a). Many methods have been developed for these purposes ranging from self-reported measures such as recall, liking, purchase intent, surveys, interviews, and focus groups (Venkatraman et al. 2015; Falk et al. 2013a). The companies use these methods under the assumption that people are capable of accurately reporting what they like and do not like to buy. Despite considerable advancement in the behavioral research methods, predicting consumers' preferences and behavior change in response to new product concepts and persuasive message exposure remains a difficult task (Cascio et al. 2013). It has turned out that people cannot fully articulate their preferences when asked them explicitly (Boksem and Smidts 2015). The neuroscientific studies have shown that many important mental processes occur on subconscious level (Dijksterhuis et al. 2006), leaving consumers very limited in their ability to predict their own future behavior and to accurately produce their mental states through verbal or written self-reports (Cascio et al. 2013).

The development of neuroscientific technologies and the applications of these technologies to product marketing—neuromarketing—give new opportunities for innovators and marketers to solve the problems that marketers and innovators face with traditional methods (Ariely and Berns

2010; Venkatraman et al. 2015). During the last decade, there is growing evidence that neuroimaging will reveal information about consumer preferences that is unobtainable through conventional methods (Falk et al. 2016; Boksem and Smidts 2015; Venkatraman et al. 2012).

The goal of this chapter is to describe how neuromarketing can provide less biased data of consumers' preferences in order to get more accurate information for innovation processes and for successful marketing campaign. The essential assumption of the chapter is that neuromarketing could illuminate not only what consumers like, but also what they will buy in future. After the introduction, the brain's valuation and mentalizing regions will be described. These regions are essential in order to understand how to apply neuroimaging data for successful innovation processes and marketing campaign. Whereas most of the applications of neuromarketing have focused on post-design applications in order to optimize the effect of advertising campaign, it has been argued that the neuromarketing can be more effective, if it is applied to early design processes (Ariely and Berns 2010). The brain's valuation and mentalizing regions have essential role in both design and post-design phases. After that the examples of the neuromarketing in the effective marketing campaigns and in the effective innovative process by using neuromarketing will be presented. After that, an example about of the effective innovative process by using neuromarketing will be presented. Furthermore the model, in which creative innovators and creative marketing campaign designers work together by using neuromarketing in order to find suitable products and advertisement to the market will be decribed. Finally, the chapter ends with the conclusion. The goal of this chapter is to give taste of neuromarketing to managerial scientists and business managers in order to increase their capability to make better decisions relating to product development and marketing campaigns.

The Brain's Regions Essential to Neuromarketing

In a current high-choice market environment, a consumer is exposed to the diverse messages and content every day. Thus, it is important to understand how consumers compare different goods and make decisions in the market. The growing evidence from behavioral sciences and consumer neurosciences shows that there are general decision networks in the brain, which "take care" of the valuation of different items in the market as well as execute the behavior according to most valuable items.

The technological development of neurophysiological tools gives researchers the opportunity to measure brain activation, when the subject participates in experiments. In this way, it is possible to study the role of brain activation for the human behavior. Neurophysiological tools can be used to produce activation maps showing which parts of the brain are involved in a particular mental process, for example, when a subject sees pictures of different goods or marketing messages. Especially by functional magnetic resonance imaging (fMRI), it is possible to measure the hemodynamic response related to neural activity, when subjects see different stimuli in fMRI experiments. By using fMRI, it is possible to measure the blood oxygen level-dependent (BOLD) signal, which is a measure of the ratio of oxygenated to deoxygenated hemoglobin (Ashby 2011). FMRI measurement is based on a principle that active brain areas consume more oxygen and sugar than do inactive areas. However, it is important to connect the neurophysiological data to real behavior, because the human behavior is the result of coevolution of neurophysiological, biological, and social issues.

Brain's Valuation Region

A growing consensus in consumer neuroscience suggests that the brain assigns values to stimuli using a common currency (Glimcher 2011; Hytönen et al. 2010; Levy and Glimcher 2012; McNamee et al. 2013). Thereby a human brain makes decisions by assigning values to the stimuli under consideration, which are then compared to make a choice (Lim et al. 2013). This valuation process of the stimuli is represented in a small number of well-identified networks in the brain. Most of studies have found that value encoding regions in the medial prefrontal cortex (MPFC), in the striatum, and in the precuneus represent consistently the value of nearly all stimulus types on a common scale. The activation change in these brain regions has been shown to reliably correlate with stimulus values using a wide class of objects from biological needs like food (Levy and Glimcher 2012), to clothes (Lim et al. 2013), and to the abstract cultural values like money (Glimcher and Fehr 2013) and charitable donations (Hare et al. 2010). In addition, the activation changes in the regions of the MPFC, the striatum, and the precuneus during stimulus presentation in the experiments predict behaviorally observed decision-making and choice (Berkman and Falk 2013; Levy and Glimcher 2012).

Thus consumer neuroscience has successfully uncovered the relationships between neuropsychological processes and behavioral outcomes, yielding a new understanding of human behavior. There are studies which have shown that relatively small group's ($n = 30$) brain activation profile in neuroscientific experiment can predict real behavioral chance in real context (Falk et al. 2015). For example, the sunscreen study showed that the neural signals in the MPFC—when subjects are exposed to persuasive messages concerning sun exposure—can predict changes in sunscreen use one week following the experiment (Falk et al. 2010). These neural signals in the MPFC predicted the better variability in behavior than self-report measures like intentions and attitudes measures explained alone. Hereby an activation pattern of neural focus group's brain uncovers hidden part of preferences, which predicts their behavioral change better than information based on traditional behavioral methods.

In the same vein, Falk et al. (2011) examined smokers' neural responses to antismoking advertisement campaigns and subsequent smoking behavior. Consistent with the findings of the sunscreen study, the MPFC activation patterns in the subjects brain, when they exposure to anti-smoke message in the scanner, expect better a quit from smoking in the future than behavioral responses of these subjects (Falk et al. 2011). Therefore, activation of the critical valuation area in the brain (MPFC) may serve as an indirect marker of future behavior change. In addition, activity in the same region of the MPFC that predicted individual behavior change during message exposure predicted population-level behavior in response to health messages and provided information that was not conveyed by participants' self-reports (Falk et al. 2012). Thus, incorporating neural data with self-report measures may provide additional information to develop predictive models. These results extend the use of neuroimaging to predict other types of behavior, as opposed to simply predicting immediate effects (Berkman and Falk 2013).

The above-described studies suggest that the neural responses in the valuation regions to messages and goods are not only predictive of purchase decisions for those individuals actually scanned, but such responses generalize to the population at large and may be used to predict the success of the sales of the products and effectiveness of messages. Therefore the responses of the MPFC in the time of decision are similar for multiple classes of items matched for value (Chib et al. 2009). The activation profiles in the brain's valuation regions detected during fMRI experiments not only predict behavioral change in real life beyond and above traditional behavioral methods, but the results from sample's brain activation profile could be generalized to the population level too. Activity in an a priori valuation circuit in the brain clearly predicts the real-world success of different advertising campaigns at the population level, whereas self-reports do not (Falk et al. 2016). Thus, it is safe to conclude that, at the time of decision, activity in the valuation regions in the brain (the MPFC, the striatum, and the precuneus) correlated with the net value of the stimulus (Levy and Glimcher 2012; Lim et al. 2013).

The Brain's Valuation Regions as Predictor of Cultural Popularity

The study of Berns and Moore (2012) shows that specific brain's activation patterns' of the sample not only predict sample's consuming of popular songs, but by using FMRI measurement, it is possible to predict the success of new songs in the market on population level.

In the study (Berns and Moore 2012), participants did listen to 60 previously unknown popular music clips in the fMRI scanner. Participants (*n* = 28) were teenagers between the ages of 12 and 17.9. According to The Recording Industry Association of America (RIAA), this age range of the study cohort accounts for approximately 20% of music sales (Berns and Moore 2012). Thus, this neurofocus group is representative of the population that is target of a brand campaign. Each clips lasted 15 seconds. Songs were downloaded from MySpace.com, and songs from unsigned musicians and relatively unknown artists were used (Berns and Moore 2012). In this way, it was possible to test the effect of new songs on the sample's brain, and the experiment was truly to test whether the neural signals are predictive of success of songs in the real market, because the songs were not on the air during experiment. In addition, the researchers used Nielsen SoundScan data as the source of post factum popularity information over the three years since the fMRI study was executed (Berns and Moore 2012).

The genres of musical clips were Rock, Country, Alternative/Emo/Indie, Hip-Hop/Rap, Jazz/Blues, and metal, all identified by the MySpace category. After listening each song, participants were required to rate the song based on how familiar it was and how much they liked it. In this subjective rating, the participants used scale ranging from 1 to 5. In this way, the researchers get both behavioral and neurophysiological data from participants' preferences of new songs.

Berns and Moore (2012) found that the vast majority of songs in the study were not commercially successful. Only three songs' albums sold over 500,000 units, which is the industry standard for "gold". The study showed that the correlation between behavioral subjective song ratings with sales data was near zero ($r = 0.11$). However, the activation within the striatum—one essential region of valuation network in the brain—was significantly correlated to the sales.

The study demonstrates that not only are signals in valuation-related networks of the human brain predictive of one small sample's purchase decisions, they are also predictive of population effects. While the main argument of neuromarketing has made claims to this effect (Ariely and Berns 2010), the data of the study suggest validity to these claims.

The Brain's Mentalizing Region

In today's complex marketing environments, it is not enough to advertise in the traditional mass media. It is more and more important that consumers like, share, and comment about the features of new goods and their experiences about a new product, service, marketing message, and TV shows in the social media. Thus, the important question is what kind of messages will go viral on social media. For example, Tesla has increased its market share every year, and this is a very profitable company today without any advertising campaigns on the TV and radio. Instead, consumers share news, personal experiences, and their comments about Tesla actively in social media more than about traditional car companies (Galloway 2016). Thus, it is important to know which messages will go viral on social media. Neuroscience has increased our understanding about effective marketing message from social media point of view. Nowadays it is possible to use neuroscience in order to find suitable ads and marketing messages, which will go viral on social media (see Falk et al. 2013b).

Whereas the brains' valuation system has important role in an individual's decision-making, the social aspect of decision-making is more and more important in today's marketing environment. The temporal parietal junction (TPJ) and the dorsomedial prefrontal cortex (DLPFC) form the brain's mentalizing region (Falk et al. 2013b). This mentalizing region has been shown to be essential during successful communication, indicating that speakers' and listeners' brains exhibit joint response patterns during this successful communication (Stephens et al. 2010).

The Brain's Mentalizing Regions as Predictor of Cultural Popularity

A large body of scientific literature has concentrated on persuasion and social influence from perspective of message recipient (Falk et al. 2013b; Weber et al. 2015; Levy and Glimcher 2012). These studies have successfully found the important role of many social-cognitive processes to the valuation and choice like conformity (Klucharev et al. 2009), responsiveness to social tagging of stimuli (Plassmann et al. 2008), and other persuasive inputs (Falk et al. 2010). However, little is known about the mechanisms that motivate communicators to share ideas effectively (Falk et al. 2013b). In order to better understand these processes, Falk and her colleagues (2013a, b) used fMRI to investigate the neurocognitive processes in the minds of message communicators in a situation in which they needed to spread ideas successfully to other individuals. Thus, the researchers investigated of the neural bases of social influence from the perspective of the influencer, rather than the influenced.

In the study by Falk et al. (2013b), message communicators (*n* = 19) pretended to be interns at a TV studio, in which their work was to provide recommendations to their boss ("the producer") about which shows should be considered for further development and production. They viewed 24 ideas for TV pilots during an fMRI scanning session, and their task was to consider whether they would pass the ideas on to producers (*n* = 79) for further consideration. In addition, the interns rated how likely they would be to recommend the idea to the producer on a scale from 1 to 4. These ideas were generated by an independent group of undergraduates, who pretended pitching a new TV show idea to a network, and from this pool of show descriptions, 24 show ideas were selected as final stimuli for fMRI study by the research team.

Both message communicators and producers were undergraduate students, and none of the producers knew the message communicators. After scanning, message communicators gave video interviews about each pilot-show idea. These video interviews were then shown to producers in a separate behavioral testing session. On the basis of these videotaped interviews, producers indicated whether they would pass the idea on to other individuals (Falk et al. 2013b).

The researchers monitored neural activity in each intern's brain while the intern was presented with 24 ideas—one at the time—to recommend to the producer. After the fMRI scan, each intern participated in video interview, and these video materials were stimuli for the producers, and the producers rated their intentions to further recommend the show idea on a scale from 1 (definitely would not) to 5 (definitely would) (Falk et al. 2013b).

The association between interns' behavioral rating (the interns rated how likely they would be to recommend the idea to the producer) and producers' intentions to further recommend the show idea was negative and there were a lot of variation in this association, indicating that interns have varied widely in their ability to persuade their producers to share their views (Falk et al. 2013b). However, the intention to share ideas to producers had associated with activity in the interns' bilateral TPJ, which is a primary component of mentalizing network. Therefore, the neural activation profile in the brain's TPJ at the moment when message communicators are initially processing the new ideas predicts the success of sharing the ideas with other people.

The Brain Computes Total Value of the Goods

Our brain valuates things from subjective point of view, and our preferences are very biased. This subjectivity applies both to experts and laymen. Experts overgenerate their personal experiences and use the information, which supports their own view (Bazerman 2006). Laymen are often unable to describe their preferences on conscious level, thus the collected behavioral data are often not valuable from consumers' behavioral point of view (Venkatraman et al. 2012). Thus, people are not omniscient but very sensitive to the subtle emotional cues and information in the environment. Asking the consumer how much they like something requires several cognitive operations. These operations include the initial processing of the stimulus, referencing similar items with which the individual has experience, and projection of future benefits, all of which may be subject to framing and social effects of the situation (Berns and Moore 2012). In contrast, brain responses in valuation regions are likely to reflect subconscious processes and may yield measurements that are less subject only to cognitive control. This would be especially true during the consumption of new products, services, and entertainments in market.

Therefore, while the act of rating something by using behavioral methods requires metacognition, the brain response during the consumption of the good does not, and the latter may prove superior to rating approaches. In the high-choice marketing environment, companies try to reach the audience by producing attractive new products and by advertising these products in framing the marketing messages by attractive features. Because the brain computes valuation all the time on the conscious and subconscious levels, this valuation is a more accurate predictor of real preferences than behavioral-based data only. Whereas people can concentrate only on a few variables on the conscious level, the brain computes total value of the goods based on many attributes relating to the products

(Rangel et al. 2008). Thus, the valuation network's reaction to the product or services reflects preferences more accurately than behavioral measurements only. The next chapter describes the possibilities of the neuromarketing in the design and post-design situations.

Experts in the companies will benefit the neuromarketing, if their attitude to the objective data-driven decision-making is positive. Unfortunately, usually the managers and other experts in the companies do not use objective source of information as source for their decision-making (Silver 2015). Their judgments and decisions include similar biases than other people's judgments and decisions (Silver 2015; Bazerman 2006). The next subchapter describes how companies could benefit from neuromarketing in their innovation and advertisement processes.

The Art and Neuroscience of Prediction

Prediction is indispensable to the brain's activity (Hawkins and Ahmad 2016). Thus, the brain computes constantly values of different aspects of the inner and outer environment of an individual. This computation is subjective and most of the time automatic. Neuroscientific data suggest that choices are made by computing the value of the different attributes and then integrating them into an overall stimulus value signal (Limb and Braun 2008; Levy and Glimcher 2012). By using suitable neuroscientific tools, it is possible to uncover these processes, and companies could use this previously hidden information in order to consider these consumers' hidden preferences to make optimal decision in product development and advertisement campaigns. Whereas most neuroscientific research has focused on finding associations between specific stimulus and the brain's activation profiles (Suomala et al. 2012; Heinonen et al. 2016), current neuroscientific studies have focused increasingly on the associations between the participants brain's activation and behavioral change in the future (Weber et al. 2015; Falk et al. 2016; Venkatraman et al. 2012).

The previous chapter described how the brain's valuation and mentalizing region's activation profile measured by fMRI in the moment when participants are initially processing a stimulus can predict the behavioral change of participants in the future. These stimuli could be any marketing material in different media format and can include videos, still pictures, texts, and auditive materials.

Therefore, neuromarketing will have a central role in the design and post-design projects of the companies, because by using neuroimaging methods, it is possible to provide designers and engineers in the research, development and innovation (RDI) department with the information that is not obtainable through conventional behavioral methods (Ariely and Berns 2010). Used in conjunction with neurophysiological methods, the behavioral experiments and interviews form a framework that enables the companies to make better predictions about the success of new products and services as well as successful advertising campaigns in the ways that were not possible previously.

Neuromarketing in the Product Design

The FFE of innovation is the early phase of innovation process, which includes stages from idea generation to decision on further development (Santonen and Hytönen 2015; Jetter 2003). After the decision for further development has been done, the development process could begin. Usually, the actual implementation process is more expensive than FFE, and this process is more linear than FFE. Thus the knowledge about consumers' preferences is more beneficial during FFE than the development process. Because multiple early decisions in FFE define central features of the final product, selecting the best ideas for further development is critical for business success, since a great majority of whole life cycle costs and features are defined at the FFE stage (Santonen and Hytönen 2015; Hytönen et al. 2016).

Experts who participated in FFE are usually designers and engineers who understand the limitations and possibilities of the technologies in their area of expertise. One explanation for the success of the Apple is, Steve Jobs emphasizes, that engineers and designers collaborated in each phase of the product development, from idea generation to the product launching in the market. This approach sees the innovation as the crossroads of liberal arts and technology (Isaacson 2011). In the same vein, when neuromarketing is applied to product and service innovations, it is a crossroads of neuroscience, technology, and art.

The innovation starts by idea generation. Architects, industrial designers, fashion designers, and other creative people—depending on the companies' productions and services—have usually essential role in this phase. When these groups begin to apply neuromarketing in this process, they need to trust neuroscientist or other neuromarketers in the way that they give their ideas exposed to objective evaluation. Thus the first step will be that designers and neuromarketers begin to collaborate. This collaboration does not denote that the original process of designers changes. In every case, the designers create many ideas during FFE. The novelty in this process is that they only learn to trust neuromarketers in order to give their outputs for review to neuromarketers. The outputs could be pictures, 3D models, videos, and audios.

The neuromarketing researchers in the Laurea's NeuroLab (http://neuroeconomics.laurea.fi/) collaborated with a construction company which has planned to build new apartment houses in Finland (Hytönen et al. 2016). In this NeuroService project, architects from the construction company planned different decors for rooms and different views of the apartments from outside. After that the brain activation of the neurofocus group's (n = 26) brain activation was measured by fMRI (Hytönen et al. 2016). The study showed that when the subjects viewed the apartments with large windows, the fMRI data revealed increased brain activation in the valuation regions (the striatum and the MPFC), indicating that the participants prefer more apartments with large windows than medium or small size. Despite the fact that the applications of neuroscience for FFE are still rare, there is growing evidence that designers, engineers, and neuroscientist work together in order to make products and services which match with consumers' preferences.

Neuromarketing in the Advertising Campaign Design

Most of the neuromarketing applications are related to advertisements optimization (Bridger 2015). Whereas most of the neuromarketing companies use biometric methods and make consultants work (Zurawicki 2010; Genco et al. 2013; Bridger 2015), there is a growing tendency to apply brain scanning—both electroencephalography (EEG) and fMRI—to advertising optimizations. Thus, neuroscientific evidence that human brain can give unique information about consumers' preferences (Glimcher and Fehr 2013; Falk et al. 2016; Boksem and Smidts 2015) has encouraged business people to apply these results for solving real marketing problems. In addition, the pioneers of neuromarketing have developed fruitful collaborative models, in which scientist and designers and creative agencies collaborate together.

A representative example of how to combine creativity with neuromarketing is the collaboration between the creative agency Saddington Baynes and neuromarketing company NeuroStrata (Noble and Reid 2016). They use real implicit measure in order to optimize advertisements in a way that consumers' preferences to the products and brands will increase during marketing campaign. The real implicit measurement can uncover participants' emotional engagement toward the stimuli. Despite the emotional engagement measurement having limitations of its own, the collaboration model of NeuroStrata and Saddington Baynes is a good example of how creative agencies and neuroscientists could collaborate in order to optimize advertisement campaign from consumers' preference point of view.

The Dutch company Neurensics is the first to make fMRI commercially available for market research (De Munnik and Vd Leij 2016). They have tested TV and radio commercials for tens of media companies. Neurensics has shown that there are not too many technical constrains to apply fMRI in order to solve marketing problems. The example of the large-scale neuromarketing company is NielsenNeuro, which is one branch of the Nielsen Media Research. NielsenNeuro uses mainly EEG in order to optimize their customers' companies' advertisements campaigns (Smith 2016). Therefore, it is safe to expect that the applications of neuroscience will be increased in commercial context.

Conclusion

Neuroscientific studies have found that the brain valuation regions assign values to goods using a common currency (Glimcher 2011; Hytönen et al. 2010; Levy and Glimcher 2012; McNamee et al. 2013). This brain-as-prediction approach has found that value encoding areas in the MPFC, in the striatum, and in the precuneus represent consistently the value of nearly all stimulus types on a common scale (Falk et al. 2015; Levy and Glimcher 2012). In addition, there is promising evidence that a human brain's mentalizing region's activation pattern indicates how well goods under consideration will spread in the social networks in the future (Falk et al. 2013b). These brain activation changes during a good presentation in the experiments predict behaviorally observed decision-making and choice (Falk et al. 2016; Levy and Glimcher 2012).

Nowadays the technological barrier of neuromarketing is not an essential problem for the spreading of neuromarketing. On the contrary, the most essential barrier is the difficulty to different stakeholders to collaborate in the innovation and marketing campaigns processes. Thus, the managers of innovation and marketing need to understand two important issues. First, they need to understand that their decisions are based on subjective biases, like other people's decisions. Second, they need to learn to apply objective data in the right way in order to make better decisions. And finally, they need to begin to trust neuromarketers as essential source to produce valuable information about consumers' hidden preferences. This chapter has described essential examples of neuroscientific studies behind neuromarketing. The hope is that managers can apply these ideas in their own work in order to make better prediction of consumers' behavior in the future.

References

Ariely, D., & Berns, G. S. (2010). Neuromarketing: The Hope and Hype of Neuroimaging in Business. *Nature Reviews. Neuroscience, 11*(4), 284–292. https://doi.org/10.1038/nrn2795.

Ashby, F. G. (2011). *Statistical Analysis of fMRI Data*. Cambridge, MA: MIT Press.

Bazerman, M. H. (2006). *Judgment in Managerial Decision Making* (6th ed.). Hoboken: J. Wiley.

Berkman, E. T., & Falk, E. B. (2013). Beyond Brain Mapping Using Neural Measures to Predict Real-World Outcomes. *Current Directions in Psychological Science, 22*(1), 45–50. https://doi.org/10.1177/0963721412469394.

Berns, G. S., & Moore, S. E. (2012). A Neural Predictor of Cultural Popularity. *Journal of Consumer Psychology, 22*(1), 154–160. https://doi.org/10.1016/j.jcps.2011.05.001.

Boksem, M. A. S., & Smidts, A. (2015). Brain Responses to Movie Trailers Predict Individual Preferences for Movies and Their Population-Wide Commercial Success. *Journal of Marketing Research, 52*(4), 482–492. https://doi.org/10.1509/jmr.13.0572.

Bridger, D. (2015). *Decoding the Irrational Consumer: How to Commission, Run and Generate Insights from Neuromarketing Research* (Marketing Science Series). London/Philadelphia: Kogan Page.

Cascio, C. N., Cin, S. D., & Falk, E. B. (2013). Health Communications: Predicting Behavior Change from the Brain. In P. A. Hall (Ed.), *Social Neuroscience and Public Health* (pp. 57–71). New York: Springer New York. http://link.springer.com/10.1007/978-1-4614-6852-3_4

Chib, V. S., Rangel, A., Shimojo, S., & O'Doherty, J. P. (2009). Evidence for a Common Representation of Decision Values for Dissimilar Goods in Human Ventromedial Prefrontal Cortex. *The Journal of Neuroscience, 29*(39), 12315–12320. https://doi.org/10.1523/JNEUROSCI.2575-09.2009.

De Munnik, M., & Leij, A. V. (2016). *Measure Valued Brand Associations*. Presented at the Neuromarketing World Forum, Dubai, April 4–6.

Dijksterhuis, A., Bos, M. W., Nordgren, L. F., & van Baaren, R. B. (2006). On Making the Right Choice: The Deliberation-Without-Attention Effect. *Science, 311*(5763), 1005–1007. https://doi.org/10.1126/science.1121629.

Falk, E. B., Berkman, E. T., Mann, T., Harrison, B., & Lieberman, M. D. (2010). Predicting Persuasion-Induced Behavior Change from the Brain. *The Journal of Neuroscience, 30*(25), 8421–8424. https://doi.org/10.1523/JNEUROSCI.0063-10.2010.

Falk, E. B., Berkman, E. T., Whalen, D., & Lieberman, M. D. (2011). Neural Activity During Health Messaging Predicts Reductions in Smoking Above and Beyond Self-Report. *Health Psychology: Official Journal of the Division of Health Psychology, American Psychological Association, 30*(2), 177–185. https://doi.org/10.1037/a0022259.

Falk, E. B., Berkman, E. T., & Lieberman, M. D. (2012). From Neural Responses to Population Behavior Neural Focus Group Predicts Population-Level Media Effects. *Psychological Science, 23*(5), 439–445. https://doi.org/10.1177/0956797611434964.

Falk, E. B., Hyde, L. W., Mitchell, C., Faul, J., Gonzalez, R., Heitzeg, M. M., Keating, D. P., et al. (2013a). What Is a Representative Brain? Neuroscience Meets Population Science. *Proceedings of the National Academy of Sciences of the United States of America, 110*(44), 17615–17622. https://doi.org/10.1073/pnas.1310134110.

Falk, E. B., Morelli, S. A., Locke Welborn, B., Dambacher, K., & Lieberman, M. D. (2013b). Creating Buzz the Neural Correlates of Effective Message Propagation. *Psychological Science, 24*(7), 1234–1242. https://doi.org/10.1177/0956797612474670.

Falk, E. B., O'Donnell, M. B., Tompson, S., Gonzalez, R., Cin, S. D., Strecher, V., Cummings, K. M., & An, L. (2015). Functional Brain Imaging Predicts Public Health Campaign Success. *Social Cognitive and Affective Neuroscience*, nsv108. doi:https://doi.org/10.1093/scan/nsv108.

Falk, E. B., O'Donnell, M. B., Tompson, S., Gonzalez, R., Cin, S. D., Strecher, V., Cummings, K. M., & An, L. (2016). Functional Brain Imaging Predicts Public Health Campaign Success. *Social Cognitive and Affective Neuroscience, 11*(2), 204–214. https://doi.org/10.1093/scan/nsv108.

Galloway, S. (2016). *Death of the Industrial Advertising Complex*. https://www.youtube.com/watch?v=yOpSpQAxCHU

Genco, S. J., Pohlmann, A. P., & Steidl, P. (2013). *Neuromarketing for Dummies*. Mississauga: Wiley.

Glimcher, P. W. (2011). *Foundations of Neuroeconomic Analysis*. New York: Oxford University Press.

Glimcher, P. W., & Fehr, E. (2013). *Neuroeconomics: Decision Making and the Brain*. London: Academic Press.

Hare, T. A., Camerer, C. F., Knoepfle, D. T., O'Doherty, J. P., & Rangel, A. (2010). Value Computations in Ventral Medial Prefrontal Cortex During Charitable Decision Making Incorporate Input from Regions Involved in Social Cognition. *The Journal of Neuroscience, 30*(2), 583–590. https://doi.org/10.1523/JNEUROSCI.4089-09.2010.

Hawkins, J., & Ahmad, S. (2016). Why Neurons Have Thousands of Synapses, a Theory of Sequence Memory in Neocortex. *Frontiers in Neural Circuits, 10*(March). https://doi.org/10.3389/fncir.2016.00023.

Heinonen, J., Numminen, J., Hlushchuk, Y., Antell, H., Taatila, V., & Suomala, J. (2016). Default Mode and Executive Networks Areas: Association with the Serial Order in Divergent Thinking (E. A. Stamatakis, Ed.). *PLOS ONE, 11*(9), e0162234. doi:https://doi.org/10.1371/journal.pone.0162234.

Hytönen, K., Sanfey, A. G., Cochran, J. J., Cox, L. A., Keskinocak, P., Kharoufeh, J. P., & Smith, C. (2010). Neuroeconomics Insights for Decision Analysis. In *Wiley Encyclopedia of Operations Research and Management Science*. Wiley. http://onlinelibrary.wiley.com/doi/10.1002/9780470400531.eorms0567/abstract

Hytönen, K., Heinonen, J., Maunula, S., Minkkinen, L., & Suomala, J. (2016). Behavioral and Neural Preference Signals of Architectural Design. In *Poster Presentation in Neuroeconomics Conference*, Berlin, August 28–30.

Isaacson, W. (2011). *Steve Jobs*. New York: Simon & Schuster.
Iyengar, S. (2011). *The Art of Choosing* (Reprint Ed.) New York: Twelve.
Jetter, A. J. (2003). Educating the Guess: Strategies, Concepts and Tools for the Fuzzy Front End of Product Development. In *Proceedings of PICMET 2003: Technology Management for Reshaping the World*, Portland.
Klucharev, V., Hytönen, K., Rijpkema, M., Smidts, A., & Fernández, G. (2009). Reinforcement Learning Signal Predicts Social Conformity. *Neuron, 61*(1), 140–151. https://doi.org/10.1016/j.neuron.2008.11.027.
Levy, D. J., & Glimcher, P. W. (2012). The Root of All Value: A Neural Common Currency for Choice. *Current Opinion in Neurobiology, 22*(6), 1027–1038. https://doi.org/10.1016/j.conb.2012.06.001.
Lim, S.-L., O'Doherty, J. P., & Rangel, A. (2013). Stimulus Value Signals in Ventromedial PFC Reflect the Integration of Attribute Value Signals Computed in Fusiform Gyrus and Posterior Superior Temporal Gyrus. *The Journal of Neuroscience, 33*(20), 8729–8741. https://doi.org/10.1523/JNEUROSCI.4809-12.2013.
Limb, C. J., & Braun, A. R. (2008). Neural Substrates of Spontaneous Musical Performance: An fMRI Study of Jazz Improvisation. *PLOS ONE, 3*(2), 1–9. https://doi.org/10.1371/journal.pone.0001679.
McNamee, D., Rangel, A., & O'Doherty, J. P. (2013). Category-Dependent and Category-Independent Goal-Value Codes in Human Ventromedial Prefrontal Cortex. *Nature Neuroscience, 16*(4), 479–485. https://doi.org/10.1038/nn.3337.
Noble, T., & Reid, A. (2016). *Design Optimizing Using Neuroscience*. Presented at the Neuromarketing World Forum, Dubai, April 4–6.
Plassmann, H., O'Doherty, J., Shiv, B., & Rangel, A. (2008). Marketing Actions Can Modulate Neural Representations of Experienced Pleasantness. *Proceedings of the National Academy of Sciences, 105*(3), 1050–1054. https://doi.org/10.1073/pnas.0706929105.
Rangel, A., Camerer, C., & Read Montague, P. (2008). A Framework for Studying the Neurobiology of Value-Based Decision Making. *Nature Reviews. Neuroscience, 9*(7), 545–556. https://doi.org/10.1038/nrn2357.
Santonen, T., & Hytönen, K. (2015). Managing Human Factor at the Fuzzy Front-End of Innovation. In *Changing the Innovation Landscape. Proceedings of the ISPIM Innovation Summit*, Brisbane.
Silver, N. (2015). *The Signal and the Noise: Why So Many Predictions Fail – But Some Don't*. New York: Penguin Books.
Smith, M. (2016). *Reproducibility and Predictive Validity in Consumer Neuroscience*. Presented at the Neuromarketing World Forum, Dubai, April 4–6.

Stephens, G. J., Silbert, L. J., & Hasson, U. (2010). Speaker-Listener Neural Coupling Underlies Successful Communication. *Proceedings of the National Academy of Sciences, 107*(32), 14425–14430. https://doi.org/10.1073/pnas.1008662107.

Suomala, J., Taatila, V., Siltala, R., & Keskinen, S. (2006). Chance Discovery as a First Step to Economic Innovation. In *Proceedings of the CogSci 2006, 28th Annual Conference of the Cognitive Science Society in Cooperation with the 5th International Conference of the Cognitive Science* (pp. 2204–2209), Lawrence Erlbaum Ass, Vancouver, July 26–29.

Suomala, J., Palokangas, L., Leminen, S., Westerlund, M., Heinonen, J., & Numminen, J. (2012). Neuromarketing: Understanding Customers' Subconscious Responses to Marketing. *Technology Innovation Management Review*, *2*: Recent Research, 12–21.

Taatila, V. P., Suomala, J., Siltala, R., & Keskinen, S. (2006). Framework to Study the Social Innovation Networks. *European Journal of Innovation Management, 9*(3), 312–326. https://doi.org/10.1108/14601060610678176.

Venkatraman, V., Clithero, J. A., Fitzsimons, G. J., & Huettel, S. A. (2012). New Scanner Data for Brand Marketers: How Neuroscience Can Help Better Understand Differences in Brand Preferences. *Journal of Consumer Psychology, 22*(1), 143–153. https://doi.org/10.1016/j.jcps.2011.11.008.

Venkatraman, V., Dimoka, A., Pavlou, P. A., Vo, K., Hampton, W., Bollinger, B., Hershfield, H. E., Ishihara, M., & Winer, R. S. (2015). Predicting Advertising Success Beyond Traditional Measures: New Insights from Neurophysiological Methods and Market Response Modeling. *Journal of Marketing Research, 52*(4), 436–452. https://doi.org/10.1509/jmr.13.0593.

Weber, R., Richard, H., Michael Mangus, J., Westcott-Baker, A., & Turner, B. O. (2015). Neural Predictors of Message Effectiveness During Counterarguing in Antidrug Campaigns. *Communication Monographs, 82*(1), 4–30. https://doi.org/10.1080/03637751.2014.971414.

Zurawicki, L. (2010). *Neuromarketing*. Berlin/Heidelberg: Springer Berlin Heidelberg. http://link.springer.com/10.1007/978-3-540-77829-5

8

Neuromarketing

Robin Chark

Neuromarketing is the study of marketing using neuroscientific methods. As introduced by Plassmann et al. (2012), it is an approach "combining neuroscience with consumer psychology" (p. 18). This interdisciplinary approach is enabled by the accumulating knowledge about how the brain works and the increased availability of neuroscientific methods worldwide (Shiv et al. 2005). While the cost of running neuroimaging studies is high relative to behavioral studies, it is not difficult to find research-dedicated brain scanners at many research-oriented universities. Scanners for clinical use are also made available to researchers in other institutions. It is not surprising to see that some 40,000 papers catalogued in PubMed have been published using functional magnetic resonance imaging (fMRI) since its introduction 25 years ago (Eklund et al. 2016).

While the use of fMRI has taken root in other fields, for example, psychology and economics, its application to marketing is relatively new and is paralleled by similar developments in other disciplines in business

R. Chark (✉)
Faculty of Business Administration, University of Macau, Macau, China

L. Moutinho, M. Sokele (eds.), *Innovative Research Methodologies in Management*,
DOI 10.1007/978-3-319-64400-4_8

schools such as neurofinance. It was argued that the application of neuroimaging might extend our observational power by providing real-time measures of implicit, automated, and non-conscious processes that are not feasible to detail from behavioral studies because participants may not be aware of, nor be able to articulate, these processes in their self-reports retrospectively (Camerer and Yoon 2015; Hedgcock and Rao 2009. Neuroimaging techniques also have the potential to enable our understanding of the role of emotion in decision-making by dissociating it from competing cognitive processes taking place simultaneously (Shiv 2007). Furthermore, the neural level measure may serve as the mediator for studying the underlying process and testing consumer theories (Shiv et al. 2005). On the practical side, there are hopes that neuroimaging may help "streamline marketing process and save money…[and] reveal information about consumer preferences that is unobtainable through conventional methods" (Ariely and Berns 2010, p. 284). Some even argue that neuromarketing might be used to "measure what consumers really think and feel, rather than simply what they state… [and] unlock what your customer really thinks" (Fisher et al. 2010, p. 233).

There are disagreements among marketing researchers and neuroscientists on the use of the term "neuromarketing." For instance, some argue that the "critical distinction is between consumer neuroscience, which refers to academic research at the intersection of neuroscience and consumer psychology, and neuromarketing, which refers to practitioner and commercial interest in neuropsychological tools, such as eye tracking, skin conductance, electroencephalography (EEG), and functional magnetic resonance imaging (fMRI), to conduct company-specific market research" (Plassmann et al. 2012, p. 19). Others use different terms to refer to different areas with differing scopes. For instance, Shiv et al. (2005) termed the application of neuroscience in the study of behavioral decision-making *decision neuroscience*.

We will use consumer neuroscience and neuromarketing interchangeably (e.g., Lee et al. 2007; Morin 2011; Reimann et al. 2011) for a few reasons. First, neuromarketing is more familiar to most readers. A distinction seems to confuse them without adding much value. Second, it is sometimes easier defining such a distinction than actually making one. For example, Dietvorst et al. (2009) developed a sales force-specific

Theory-of-Mind (ToM) scale and validated the questionnaire-based scale with neuroimaging data. The authors found that those who scored high in the scale showed more activation in brain areas that are associated with ToM or mentalization in interpersonal situations than those who scored low in the scale. While the development of the scale is guided by consumer theory, the scale itself has great potential for commercial use. It is not clear to us whether this advancement should be classified as consumer neuroscience or neuromarketing.

Another debate is whether neuromarketing should be treated as a branch of neuroscience studying human behavior by borrowing the theoretical perspectives from consumer research, or if it is more accurately a part of consumer research trying to contribute to the development of consumer theories by utilizing neuroscientific methodologies (Plassmann et al. 2015). While the former may contribute to the study of science generally, or neuroscience more specifically, the focus of this chapter is on the latter, as this book is written for readers who are interested in recent advancements in research in management. In order to contribute to consumer research, Chen et al. (2015) proposed that neuromarketing researchers should move on from answering the "where"-type questions to the "what"-type questions. The former refers to the stage when research focuses on physically localizing different psychological processes. Shifting the focus involves answering questions such as what process is involved, and what and how information is being processed. More specifically, Shiv et al. (2005) proposed a few ways that neuroscientific methods might contribute to consumer research by:

1. providing confirmatory evidence about the existence of a phenomenon,
2. generating a more fundamental (i.e., a neural level) conceptualization and understanding of underlying processes,
3. refining existing conceptualizations of various phenomena, and
4. providing methodologies for testing new as well as existing theories.

More than a decade has lapsed since Shiv et al.'s proposal. It may be a good time to review the literature and see how much has been achieved. To this end, we discuss a few neuromarketing papers in the next section.

This is followed by the proposal of a neurobiological framework that could broaden our scope of investigation.

What Has Been Achieved?

This section discusses a few fMRI papers in terms of their contribution to consumer research in the four ways proposed by Shiv et al. (2005). Given the space limit, the focus is on papers using fMRI, which is the most popular tool, although studies abound using other neuroimaging methods such as magnetoencephalography (MEG) and EEG. fMRI measures the blood oxygen level-dependent (BOLD) signals in the brain. It is a proxy of the neural activities since it has been shown that the neural activities result in oxygen consumption, which triggers an influx of oxygenated blood. Such a hemodynamic response allows us to compare the brain processes we are interested in and similar ones for which a key component of the process is missing.[1] This allows us to localize the neural activities of a particular process of interest and is crucial to answering questions that have not been solved by behavioral studies alone.

Providing Confirmatory Evidence About the Existence of a Phenomenon

As discussed in Shiv et al.'s (2005) work, medical research has used fMRI to validate the placebo effect, in which the patients' belief in the efficacy of placebo analgesia, despite not containing the medical ingredient, apparently alleviates pain. Wager et al. (2004) validated the placebo with neuroimaging data in the sense that "placebo analgesia actually induces decreased brain activity in pain-sensitive brain regions" (Shiv et al. 2005, p. 376). More interestingly, it has been demonstrated that people find more expensive placebos more effective in killing pain than less expensive ones (Waber et al. 2008). Similarly in consumer research, it has been questioned whether pricing, a marketing action, may change consumers' subjective perception similar to the case of placebo effect. It was often found that perceived quality positively correlated with price (Rao and Monroe 1989). However, there is an alternative explanation that participants may be susceptible to the demand effect in which they may consider it inappropriate to report a lower satisfaction when they see a higher price. Thus it was an open question whether this price-perceived quality relationship was real. Without a proper implicit measure of the effectiveness, this question could not be resolved.

To this end, Plassmann et al. (2008) examine this price-perceived quality relationship in an fMRI. The authors hypothesize that the price of a product not only influences the reported perceived quality but also the neural representation of the subjective enjoyment of the product. A notable feature of this study was that participants were asked to taste wines labeled with different price tags. The authors found that prices of the wines positively correlated with the reported pleasantness. This behavioral pattern is consistent with previous results. The authors also found greater brain activity in the bilateral medial orbitofrontal cortex (mOFC) and roastral anterior cingulate cortex (rACC) in the high price condition. In particular, higher taste expectations as a result of higher price would lead to higher activity in the mOFC, which was previously associated with the encoding of subjective value. This finding validates the perceived quality-price relationship, which would have been difficult had neuroscientific methodology not been accessible to consumer researchers.

The major contribution of the paper is that while there are numerous examples in the literature that various marketing variables can influence consumers' experienced pleasantness, this is the first time such a variable, namely, price, is shown to modulate the neural representations of this signal. While prices can modulate the representations of experienced pleasantness, the encoding of the taste in the primary gustatory cortex is independent of price. Thus the perceived quality-price relationship is more complicated. The increase in experienced pleasantness occurs through the increase in taste expectations as a result of consumers' price perceptions.

Generating a More Fundamental Conceptualization and Understanding of Underlying Processes

In a pioneering paper, McClure et al. (2004) studied the process through which branding, another marketing action, might shape our preferences. It has been demonstrated that subjective enjoyment of the consumption of a product could depend on the name of its brand. However, it is very difficult, if not impossible, to study the underlying process of how branding may shape our subjective perceptions using behavioral methodologies alone, because it is likely that the process is automatic and not conscious to consumers.

McClure et al. (2004) were interested in the question of why people displayed strong preference for Coke or Pepsi despite the fact that the drinks have nearly identical chemical compositions. They hypothesized that the strong brand preference could be related to the modulating role of cultural messages by the brand in shaping perceptions. They argued that our preferences for food and beverages are "modulated by an enormous number of sensory variables, hedonic states, expectations, semantic priming, and social context… many levels of social, cognitive, and cultural influences combine to produce behavioral preferences for food and drink" (p. 379). The cultural information might be used as input to modulate reward-related brain response. For example, Erk et al. (2002) found greater activation in the brain's reward circuits (e.g., ventral striatum and orbitofrontal cortex and anterior cingulate) when subjects looked at sports cars (symbols of social status) than when they viewed sedans or compact cars. It was further found that brand names of the cars might modulate reward-related brain responses (Schaefer et al. 2006; Schaefer and Rotte 2007).

McClure et al. (2004) used a simple taste test and fMRI to study the question of whether a relationship exists between the expressed preference and the preference-related neural response. Consistent with previous behavioral findings, the authors found that stated and behavioral preferences were not correlated in an anonymous taste test, demonstrating that stated preference was affected by factors other than taste alone. Under anonymous conditions, brain responses in the ventromedial prefrontal cortex (vmPFC), which was implicated in signaling basic appetitive aspects of reward, were observed to correlate with behavioral preferences but not stated preferences.

Under semianonymous conditions, subjects knew the brand of one of the drinks only (e.g., they were told that one of the two drinks was Coke and the other could be either Coke or Pepsi), although the other cup was also filled with the same brand. Despite the fact that both cups contained Coke, subjects preferred Coke in the labeled cups significantly more than they did in the anonymous condition, and this increase in preference was also significantly greater than the same semianonymous-anonymous comparison when the brand was Pepsi. During the fMRI experiment, increased neural activations were found for the labeled Coke as compared to unlabeled Coke in the bilateral hippocampus, parahippocampus, dorsolateral prefrontal cortex (dlPFC), thalamus, left visual cortex, and midbrain. These areas have been associated with behavior modification through their roles in memory and affect. However, such brain activations were not found in the semianonymous condition of Pepsi. Brand knowledge of Coke versus Pepsi had differential effects on both behavioral preference and brain responses.

These findings suggest that the cultural aspects of a brand may help shape our preferences through emotional memory associated with the brands. In particular, two separate brain systems—namely, the vmPFC and the hippocampus/dlPFC/midbrain—may function independently in shaping our preferences based on sensory and cultural information, respectively. This work enhances our understanding of the underlying process of how brand shapes our preferences at a neural level.

Refining Existing Conceptualizations of Various Phenomena

Neuromarketing helps consumer researchers to refine existing conceptualizations. Brand personality is an important concept in the extant literature. A brand-personality scale was developed (Aaker 1997) that was quite successful in the past in conceptualizing consumers' perception of different brands with personalities akin to the ones used to characterize people, though it is not a given that consumers actually process brand and people personalities similarly. In particular, Azoulay and Kapferer (2003) argued against the validity of brands being thought of as having full-fledged human personalities. For example, both a friend and a car can be described as "reliable," yet this word may carry very different meanings (Yoon et al. 2006). Yoon et al. further argue that "objects and persons can have similar 'personality structures' but could well be processed using different neural systems" (p. 32).

Using fMRI, Yoon et al. (2006) investigated whether semantic judgments about products and persons were processed similarly. Instead of trying to resolve the controversy, their goal was to use neuroimaging techniques to investigate the implicit assumption that the brand qualities are processed the same way as people qualities. In their experiment, neural processing of the judgments of brand and human personalities was compared when participants were scanned with fMRI.

The authors found that brand and person judgments were processed in two distinct neural networks. In particular, comparisons of neural correlates of product versus human descriptor judgments indicated greater activation in the medial prefrontal cortex (mPFC) regions for judgments of humans, while the reverse contrast showed greater activation in the left inferior prefrontal cortex, an area associated with object processing. This implies that judgments about brands are processed in brain areas that are involved in semantic object processing, rather than areas that are involved in judgments about people. These results support the skepticism of Yoon et al. toward the brand-personality research. The authors further found an incongruence in the brain activation patterns between the subjects' self-concept and their brand associations. The differences in activation in mPFC between the judgments of the self-relevant and non-self-relevant brands were smaller than they were between the judgments of the self and other people. These results further undermine the validity of a direct application of personality theory to brand theory. The finding suggests a refinement of the existing conceptualization of brand personality. While brand personality is akin to human personality, consumer researchers should be careful in applying our knowledge of the latter to the study of the former.

Providing Methodologies for Testing New as Well as Existing Theories

Last but not least, consumer researchers may distinguish theories by forming specific predictions on the process—at the neural level in particular—and testing them with fMRI. One notable example of using fMRI to test consumer theories was given by Hedgcock and Rao (2009). Huber et al. (1982) demonstrated that an addition of an option (decoy) asymmetrically dominated by one of the two options, which were equally attractive, would increase the choice share of the dominating option by stealing share from the other option in the original choice set.

Before Huber et al. (1982), many people believed that the similarity effect is the major deviation from Luce's axioms of choice model. The original axioms require our choice probabilities to follow regularity and proportionality. The former refers to the requirement that the addition of a new option to a choice set will only hurt the original alternatives, and the latter requires that the new option hurts them proportionally. The similarity effect refers to the disproportional substitution effect predicting that options in the original choice set more similar to the added one will be hurt disproportionally more. Thus the similarity effect violates proportionality.

However, the attraction effect questioned not only the proportionality axiom but also the regularity condition, which might have sounded reasonable in the first place. The attraction effect is also incompatible with the similarity effect in which the neighboring option, instead of being hurt, is helped by the decoy. This effect spurred a new wave of studies in multiattribute decision-making and stimulated consumer researchers' interests in judgment and decision-making.

Despite the importance of the phenomenon in decision research, no conclusive explanation for the phenomenon was reached. Several possible explanations have been raised. One line of reasoning borrows from the range-frequency model originally developed in the 1960s (e.g., Parducci 1965). Due to the change in the attribute range and ranks of the alternatives on some attribute dimension when a decoy is introduced, the model predicts a change in the perceived values associated with the original alternatives (Huber et al. 1982; Wedell 1991; Wedell and Pettibone 1996). Another line of reasoning, instead of changing the perceived values, relies on a change in the weights associated with the attributes, resulting in a change in the relative positions of the overall attractiveness of the options in the original choice set (Wedell 1991; Wedell and Pettibone 1996). The third explanation suggests that the presence of a decoy results in an obvious dominance relationship and thus increases the justifiability of the choice of the dominating alternative (Huber et al. 1982; Simonson 1989; Wedell 1991; Wedell and Pettibone 1996).

Hedgcock and Rao (2009) advanced this discussion to another level by testing these theories using neuroimaging technique. In particular, they tested the role of emotions during the decision-making process. Luce et al. (2001) posited that making difficult tradeoffs could be emotionally taxing and was aversive to decision-makers. Thus when a decoy is introduced to the choice set, the decision-maker finds a way out of this difficult tradeoff by choosing the dominating option, which is easy to justify. The negative emotion is thus attenuated. Hedgcock and Rao hypothesized that the presence of a decoy might lead to decreased activation in the amygdala (which was associated with negative emotion), decreased activation of mPFC (which was associated with self-referential evaluation of preferences), increased activation in the anterior cingulate cortex (ACC) (which was associated with conflict monitoring), and increased activation in the dlPFC (which was associated with the use of decision rules).

As hypothesized, the presence of a decoy leads to lower decision difficulties and thus less tradeoff aversion as represented by reduced activations in brain regions such as the amygdala and mPFC and increased activation in the ACC and dlPFC. Hedgcock and Rao (2009) provided a good start to the testing of consumer theories by adopting neuroscientific methods. In particular, the involvement of emotion in the attraction effect gives support for the tradeoff aversion account of the phenomenon.

Summary

To sum up, the neuromarketing literature has evolved around the four potential contributions proposed more than a decade ago (Shiv et al. 2005). One notable feature of these neuromarketing studies pertaining especially to the third and fourth contributions about refining existing conceptualizations and testing theories is the use of a theoretically grounded, hypothesis-driven approach. Based on consumer theories and neuroscientific knowledge, *a priori* hypotheses at both the behavioral and neural levels are developed before the actual implementation of the study. Results are based on region-of-interest (ROI) analysis in which specific hypotheses about the neural activities are tested. This hypothesis-driven approach helps address the multiple comparisons problem that arises when a set of statistical inferences is being considered at the same time without proper correction for the resulting increased likelihood of false positives. Eklund et al. (2016) demonstrated empirically that the correction applied to multiple comparisons is erroneous, and thus statistical significance is often inflated in a few common statistics packages used in neuroimaging data analysis. The ROI analysis, instead of making statistical inferences from data of tens of thousands of voxels simultaneously, only makes inferences from activities of a few brain regions picked *a priori*, and thus greatly reduces the number of statistical tests and the likelihood of false positives. Relatedly, the interpretation of results often relies on the so-called *reverse inference* (Poldrack 2006). It is problematic when a process is inferred from the activity of a particular brain region, which is often involved in multiple processes. The hypothesis-driven approach helps alleviate this problem. The neuroimaging experiment is designed to single out the process of interest. Data analysis is guided by *a priori* hypotheses about which brain regions are likely to be involved. The interpretation of results thus relies less on the use of reverse inference.

What Is Next?

In the previous section, we review some fMRI studies in neuromarketing. In this section, we propose the application of a neurobiological framework to broaden the scope of investigation. A fifth contribution proposed later is the ability to help uncover individual differences in how the brain might process the same input differently (Plassmann et al. 2015; Venkatraman et al. 2012). In particular, the Venkatraman et al. argued that "cognitive neuroscience can identify both a homogeneous neural mechanism for different types of valuation and the heterogeneous output of that mechanism across individuals… [or it may be used] to understand when different individuals employ different choice strategies" (p. 145). For example, in a recent fMRI study, neuroimaging data were used to examine the individual differences in the perceived quality-price relationship (Plassmann and Weber 2015). This new direction of examining individual differences could benefit from broadening the scope of investigation, as suggested in recent developments in behavioral genetics and psychopharmacology.

In order to advance the field of neuromarketing to the next level, taken from the neurobiology literature, consumer researchers may broaden the scope by following the recent developments in behavioral genetics. Smidts et al. (2014) proposed that the study of genetics would be the next frontier of neuromarketing. In particular, these authors argued that "choice behavior, like all complex biological traits, is the result of interactions between genetic, environment, and developmental/epigenetic processes" (p. 262). More concretely, "neuroscience can illuminate the lesser-known fact that biological traits that engender those differences likely exist at many points (e.g., genes, hormones, brain regions) in the neurobiological mechanisms pertaining to decision making" (Venkatraman et al. 2012, pp. 145–146).

To better organize these factors, a framework is taken from the neurobiology literature. It is believed that our behaviors are results of psychological processes embodied physiologically in our brain and nervous system. The neurobiological influences may take root in our genes and shape the activities in our brain and thus behaviors through action on hormones and neurotransmitters. The heterogeneity of behavioral tendency may be the consequence of differences in the neural mechanisms in any intermediate steps and could be traced all the way to the individual differences in genotypes. Thus we propose that the Gene-Brain-Behavior relationship could provide a framework to organize the study of neurobiological factors more systematically. Under this framework, our genotypes may be interpreted as a measure of individual differences, while brain activities may be observed and taken as more direct measures of the underlying psychological process, or as the mediator in the whole neurobiological mechanism as proposed previously by Shiv et al. (2005). In the following, recent developments are reviewed in behavioral genetics and its application in marketing. Then the framework is extended to include the study of neurotransmitters.

Paralleling the development of the study of human behavior with neuroimaging, advancement in behavioral genetics has also come to the point that we may apply it in social sciences.[2] While there are quite a few twin studies[3] exploring the heritability effect of prosocial behavior (Knafo and Plomin 2006), financial risk-taking (Cesarini et al. 2010), and a set of anomalies in judgment and decision-making (Cesarini et al. 2012), the study of genetic factors in marketing is rather new. A notable example was Simonson and Sela (2011). In consumer research, there has been interest in studying whether our preferences are inherited or learned (e.g., Tu and Hsee 2016). Simonson and Sela (2011) went one step further by measuring "the magnitude of heritable effects on consumer decision making across a wide range of previously studied choice and judgment phenomena" (p. 952). The authors did not stop here. By comparing the magnitudes of the estimated heritable effects, they were able to generate a new hypothesis of a potential underlying genetic influence common across different judgments and choice processes. In particular, the identification of large heritable effects in "certain choice problems, such as those involving compromise, vice and virtue, risk and loss, and maximizing" (p. 962) suggests an underlying genetic influence related to prudence.

A few methodological advancements in behavioral genetics could help guide the developing application of this neurogenetic approach going forward. Here we identify a couple of paths that consumer researchers may take to further the application of neurogenetics in consumer research. First, Smidts et al. (2014) argued that "understanding the genetic and molecular basis of choice behavior can shed light on how preferences are formed" (p. 262). For instance, it has been shown that self-construal—whether it is independent or interdependent—affects consumer risk-taking (Mandel 2003), among other consumer behaviors. Kitayama et al. (2014) found that a dopamine D4 receptor gene (*DRD4*) moderated the cultural difference in the social orientation of viewing oneself as either more independent or interdependent. They used a candidate gene approach, akin to the ROI analysis in fMRI, in which a gene is picked *a priori* of the study and with the guidance of existing knowledge of its functional significance. In this case, *DRD4* is picked since it is known to modulate the efficiency of the neurotransmission in the dopamine pathways, which process reward signals and are active in reinforcement learning. A more efficient dopaminergic neurotransmission thus "may be expected to be associated with greater adoption of norm-congruous responses" (p. 1171). This example illustrates how molecular genetics may be applied to the study of consumer behavior.

A second development with potential for marketing insights is the so-called *imaging genomics* (Hariri and Weinberger 2003). While we have learned from behavioral genetics how individual genes might contribute to our behaviors, there are quite a few intermediate steps between genes' influence and our behavioral tendencies. Thus the effect size of the contribution is often low. For example, in a large-scale genome-wide association study (GWAS), a whole genome search identified three genes that significantly contributed to education attainment, a high-level, complex trait (Rietveld et al. 2013). However, these three genes together account for less than 3 percent of the total phenotypical variations of the trait. Under the Gene-Brain-Behavior framework, the genetic influence operates through its expression in the function of the brain. For instance, the amygdala is found to be more sensitive to fearful faces among carriers of the short allele of the serotonin transporter (5-HTT) promoter polymorphism compared with those homozygous for the long allele (Hariri et al. 2002). This finding fills the gap between the short allele genotype and the behaviors characterized by increased fear and anxiety. The latter paper also illustrates how consumer researchers may study the genotypic differences in how the brain processes the same input differently and gives rise to behavioral differences by applying neuroimaging in their studies, as suggested in Simonson and Sela (2011).

In order to complete the loop, we revise the neurobiological framework to include the role of neurotransmitters. Neurotransmitters are chemical messengers released by synaptic vesicles on the axon of a neuron to enable neurotransmissions of signals across synapses. They are received by receptors on the target cell, which could be another neuron, a muscle cell, or a gland cell. This role of neurotransmitters is critical in the neurochemical process of the whole neural system. For instance, dopamine, a neurotransmitter, modulates the dopaminergic system in processing the reward signal (O'Connell and Hofmann 2012).

The purpose of studying the role of neurotransmitters is twofold. First, it bridges the huge gap in our understanding of how genes are expressed in the brain either functionally or structurally, and thus how they give rise to the individual differences in behavioral tendencies. In particular, genes may contribute to the chronic level differences of neurotransmitters among individuals and in turn affect the efficiency of the neurotransmission of a particular signal type. The dopaminergic modulation of cultural differences in interdependent versus independent social orientation discussed previously is a good example (Kitayama et al. 2014).

Second, a better understanding of neurotransmitters may shed light on a "possible biological mechanism underlying the observed variability of consumer behavior" (Lichters et al. 2016, p. 183). Some external factors may trigger an acute change of the level of a neurotransmitter and thus lead to a change in an individual's behavioral tendency. For instance, as noted in Simonson and Sela (2011), being "an explicit, conscious compromise between conflicting preferences as a means for reducing conflict and anticipated regret" (p. 962), the compromise effect has a large heritable component compared with other context effects such as the asymmetric dominance effect. Lichters et al. (2016) conducted a psychopharmacology study to examine the downstream consequence of such genetic influence, namely, the serotonergic pathways. An amino-acid mixture that lacks tryptophan was orally administered to half of the participants. As a precursor of serotonin, tryptophan depletion leads to an acute reduction in circulating serotonin. Compared with those who received the placebo in the control group, the acutely tryptophan-depleted participants were less susceptible to the compromise effect as it required an explicit, conscious tradeoff that was made difficult for this group as a result of serotonin deficiency. The result gives support for the conscious and effortful tradeoff explanation to the compromise effect. The advantage of a typical psychopharmacology study is that since the target neurotransmitter is acutely manipulated, causal relationship can be established. The causality could be useful in explaining individual differences in consumer behavior as different population groups (e.g., age groups, genotypes) may differ in the chronic level of the neurotransmitter. This would allow new ways to test consumer theory.

In addition, two developments that fit into our framework but are not reviewed particularly in this chapter are hormonal and epigenetic factors. Plassmann et al. (2012), when describing how neuroscience may help generate new consumer theories, referred to the recent research studying how ovulation cycle influences female consumer behavior (e.g., Durante et al. 2011) and argued that such an "approach is of great potential for developing an interdisciplinary understanding of how consumers make decisions and may provide significant improvements in our understanding of preference formation and decision making" (p. 32). Other hormones such as testosterone (e.g., Aspara and Van Den Bergh 2014) are also demonstrated to influence consumer behavior significantly. Epigenetics is a very active area in biology. It may potentially uncover how genes interact with environmental factors to shape consumer behavior (Simonson and Sela 2011). We defer these topics to another review in the future.

Concluding Remarks

What did neuromarketing achieve in its first decade? To answer this question, we reviewed a few papers and grouped them according to the four contributions proposed by Shiv et al. (2005). While the number of fMRI papers in neuromarketing in the past ten years or so could not match the exponential growth in the sheer number of fMRI studies in other fields such as neuroeconomics, quite a few neuromarketing papers attracted attention from researchers within and outside the field of marketing. More importantly, they contributed to the development of consumer theories. In particular, some of them provided unprecedented neuroscientific evidence supporting the existence of certain marketing phenomena. Other studies helped us better understand consumer behavior by tapping into the underlying psychological processes as represented at the neural level. The technology also helped refine existing conceptualizations and test consumer theories using neuroimaging data. The next question is what the future of neuromarketing will be. To this end, a neurobiological framework is proposed for extending the scope of investigation. In particular, the Gene-Neurotransmitter-Brain-Behavior relationship may be used to organize these factors at different stages in the neurobiological system with varying proximity to behavior. To integrate them, we propose that researchers may combine different methodologies, as in the cases of imaging genomics and psychopharmacological fMRI studies.

Notes

1. Readers are referred to Reimann et al. (2011) and Shiv et al. (2005) for more elaborated discussions of the methodology of fMRI, and to Reimann et al. (2011) for a survey of its advantages and limitations in particular.

2. For a review of the use of behavioral genetics in studying social cognition and a general discussion of various methodologies, readers are referred to Ebstein et al. (2015).
3. Interested readers are referred to Simonson and Sela (2011) for the detailed discussion of the methodology of twin study and how heritability effect could be singled out by comparing monozygotic and dizygotic twins.

References

Aaker, J. L. (1997). Dimensions of Brand Personality. *Journal of Marketing Research, 34*(3), 347–356.

Ariely, D., & Berns, G. S. (2010). Neuromarketing: The Hope and Hype of Neuroimaging in Business. *Nature Reviews Neuroscience, 11*(4), 284–292.

Aspara, J., & Van Den Bergh, B. (2014). Naturally Designed for Masculinity Vs. Femininity? Prenatal Testosterone Predicts Male Consumers' Choices of Gender-Imaged Products. *International Journal of Research in Marketing, 31*(1), 117–121.

Azoulay, A., & Kapferer, J.-N. (2003). Do Brand Personality Scales Really Measure Brand Personality? *Journal of Brand Management, 11*(2), 143–155.

Camerer, C., & Yoon, C. (2015). Introduction to the Journal of Marketing Research Special Issue on Neuroscience and Marketing. *Journal of Marketing Research, 52*(4), 423–426.

Cesarini, D., Johannesson, M., Lichtenstein, P., Sandewall, Ö., & Wallace, B. (2010). Genetic Variation in Financial Decision-Making. *The Journal of Finance, 65*(5), 1725–1754.

Cesarini, D., Johannesson, M., Magnusson, P. K. E., & Wallace, B. (2012). The Behavioral Genetics of Behavioral Anomalies. *Management Science, 58*(1), 21–34.

Chen, Y.-P., Nelson, L. D., & Hsu, M. (2015). From "Where" to "What": Distributed Representations of Brand Associations in the Human Brain. *Journal of Marketing Research, 52*(4), 453–466.

Dietvorst, R. C., Verbeke, W. J. M. I., Bagozzi, R. P., Yoon, C., Smits, M., & Van Der Lugt, A. (2009). A Sales Force-Specific Theory-of-Mind Scale: Tests of Its Validity by Classical Methods and Functional Magnetic Resonance Imaging. *Journal of Marketing Research, 46*(5), 653–668.

Durante, K. M., Griskevicius, V., Hill, S. E., Perilloux, C., & Norman, P. L. (2011). Ovulation, Female Competition, and Product Choice: Hormonal Influences on Consumer Behavior. *Journal of Consumer Research, 37*(6), 921–934.

Ebstein, R. P., Zhong, S., Chark, R., San Lai, P., & Chew, S. H. (2015). Modeling the Genetics of Social Cognition in the Laboratory. In *The Oxford Handbook of Molecular Psychology* (p. 167). New York: Oxford University Press.

Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster Failure: Why fMRI Inferences for Spatial Extent Have Inflated False-Positive Rates. *Proceedings of the National Academy of Sciences, 113*(28), 7900–7905.

Erk, S., Spitzer, M., Wunderlich, A. P., Galley, L., & Walter, H. (2002). Cultural Objects Modulate Reward Circuitry. *Neuroreport, 13*(18), 2499–2503.

Fisher, C. E., Chin, L., & Klitzman, R. (2010). Defining Neuromarketing: Practices and Professional Challenges. *Harvard Review of Psychiatry, 18*(4), 230–237.

Hariri, A. R., & Weinberger, D. R. (2003). Imaging Genomics. *British Medical Bulletin, 65*(1), 259–270.

Hariri, A. R., Mattay, V. S., Tessitore, A., Kolachana, B., Fera, F., Goldman, D., Egan, M. F., & Weinberger, D. R. (2002). Serotonin Transporter Genetic Variation and the Response of the Human Amygdala. *Science, 297*(5580), 400–403.

Hedgcock, W., & Rao, A. R. (2009). Trade-Off Aversion as an Explanation for the Attraction Effect: A Functional Magnetic Resonance Imaging Study. *Journal of Marketing Research, 46*(1), 1–13.

Huber, J., Payne, J. W., & Puto, C. (1982). Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis. *Journal of Consumer Research, 9*(1), 90–98.

Kitayama, S., King, A., Yoon, C., Tompson, S., Huff, S., & Liberzon, I. (2014). The Dopamine D4 Receptor Gene (DRD4) Moderates Cultural Difference in Independent Versus Interdependent Social Orientation. *Psychological Science, 25*(6), 1169–1177.

Knafo, A., & Plomin, R. (2006). Prosocial Behavior from Early to Middle Childhood: Genetic and Environmental Influences on Stability and Change. *Developmental Psychology, 42*(5), 771.

Lee, N., Broderick, A. J., & Chamberlain, L. (2007). What Is 'Neuromarketing'? A Discussion and Agenda for Future Research. *International Journal of Psychophysiology, 63*(2), 199–204.

Lichters, M., Brunnlieb, C., Nave, G., Sarstedt, M., & Vogt, B. (2016). The Influence of Serotonin Deficiency on Choice Deferral and the Compromise Effect. *Journal of Marketing Research, 53*(2), 183–198.

Luce, M. F., Bettman, J. R., & Payne, J. W. (2001). An Integrated Model of Trade-Off Difficulty and Consumer Choice. *Journal of Consumer Research, 1*(11), 11–35.

Mandel, N. (2003). Shifting Selves and Decision Making: The Effects of Self-Construal Priming on Consumer Risk-Taking. *Journal of Consumer Research, 30*(1), 30–40.

McClure, S. M., Li, J., Tomlin, D., Cypert, K. S., Montague, L. M., & Read Montague, P. (2004). Neural Correlates of Behavioral Preference for Culturally Familiar Drinks. *Neuron, 44*(2), 379–387.

Morin, C. (2011). Neuromarketing: The New Science of Consumer Behavior. *Society, 48*(2), 131–135.

O'Connell, L. A., & Hofmann, H. A. (2012). Evolution of a Vertebrate Social Decision-Making Network. *Science, 336*(6085), 1154–1157.

Parducci, A. (1965). Category Judgment: A Range-Frequency Model. *Psychological Review, 72*(6), 407.

Plassmann, H., Ambler, T., Braeutigam, S., & Kenning, P. (2007). What Can Advertisers Learn from Neuroscience? *International Journal of Advertising, 26*(2), 151–175.

Plassmann, H., & Weber, B. (2015). Individual Differences in Marketing Placebo Effects: Evidence from Brain Imaging and Behavioral Experiments. *Journal of Marketing Research, 52*(4), 493–510.

Plassmann, H., O'Doherty, J., Shiv, B., & Rangel, A. (2008). Marketing Actions Can Modulate Neural Representations of Experienced Pleasantness. *Proceedings of the National Academy of Sciences, 105*(3), 1050–1054.

Plassmann, H., Ramsøy, T. Z., & Milosavljevic, M. (2012). Branding the Brain: A Critical Review and Outlook. *Journal of Consumer Psychology, 22*(1), 18–36.

Plassmann, H., Venkatraman, V., Huettel, S., & Yoon, C. (2015). Consumer Neuroscience: Applications, Challenges, and Possible Solutions. *Journal of Marketing Research, 52*(4), 427–435.

Poldrack, R. A. (2006). Can Cognitive Processes Be Inferred from Neuroimaging Data? *Trends in Cognitive Sciences, 10*(2), 59–63.

Rao, A. R., & Monroe, K. B. (1989). The Effect of Price, Brand Name, and Store Name on Buyers' Perceptions of Product Quality: An Integrative Review. *Journal of Marketing Research, 26*(3), 351–357.

Reimann, M., Schilke, O., Weber, B., Neuhaus, C., & Zaichkowsky, J. (2011). Functional Magnetic Resonance Imaging in Consumer Research: A Review and Application. *Psychology and Marketing, 28*(6), 608–637.

Rietveld, C. A., Medland, S. E., Derringer, J., Yang, J., Esko, T., Martin, N. W., Westra, H.-J., et al. (2013). GWAS of 126,559 Individuals Identifies Genetic Variants Associated with Educational Attainment. *Science, 340*(6139), 1467–1471.

Schaefer, M., & Rotte, M. (2007). Favorite Brands as Cultural Objects Modulate Reward Circuit. *Neuroreport, 18*(2), 141–145.

Schaefer, M., Berens, H., Heinze, H.-J., & Rotte, M. (2006). Neural Correlates of Culturally Familiar Brands of Car Manufacturers. *NeuroImage, 31*(2), 861–865.

Shiv, B. (2007). Emotions, Decisions, and the Brain. *Journal of Consumer Psychology, 17*(3), 174–178.

Shiv, B., Bechara, A., Levin, I., Alba, J. W., Bettman, J. R., Dube, L., Isen, A., et al. (2005). Decision Neuroscience. *Marketing Letters, 16*(3–4), 375–386.

Simonson, I. (1989). Choice Based on Reasons: The Case of Attraction and Compromise Effects. *Journal of Consumer Research, 16*(2), 158–174.

Simonson, I., & Sela, A. (2011). On the Heritability of Consumer Decision Making: An Exploratory Approach for Studying Genetic Effects on Judgment and Choice. *Journal of Consumer Research, 37*(6), 951–966.

Smidts, A., Hsu, M., Sanfey, A. G., Boksem, M. A. S., Ebstein, R. B., Huettel, S. A., Kable, J. W., et al. (2014). Advancing Consumer Neuroscience. *Marketing Letters, 25*(3), 257–267.

Tu, Y., & Hsee, C. K. (2016). Consumer Happiness Derived from Inherent Preferences Versus Learned Preferences. *Current Opinion in Psychology, 10*, 83–88.

Venkatraman, V., Clithero, J. A., Fitzsimons, G. J., & Huettel, S. (2012). New Scanner Data for Brand Marketers: How Neuroscience Can Help Better Understand Differences in Brand Preferences. *Journal of Consumer Psychology, 22*, 143–153.

Waber, R. L., Shiv, B., & Carmon, Z. (2008). Commercial Features of Placebo and Therapeutic. *JAMA, 299*(9), 1016–1017.

Wager, T. D., Rilling, J. K., Smith, E. E., Sokolik, A., Casey, K. L., Davidson, R. J., Kosslyn, S. M., Rose, R. M., & Cohen, J. D. (2004). Placebo-Induced Changes in fMRI in the Anticipation and Experience of Pain. *Science, 303*(5661), 1162–1167.

Wedell, D. H. (1991). Distinguishing Among Models of Contextually Induced Preference Reversals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(4), 767.

Wedell, D. H., & Pettibone, J. C. (1996). Using Judgments to Understand Decoy Effects in Choice. *Organizational Behavior and Human Decision Processes, 67*(3), 326–344.

Yoon, C., Gutchess, A. H., Feinberg, F., & Polk, T. A. (2006). A Functional Magnetic Resonance Imaging Study of Neural Dissociations Between Brand and Person Judgments. *Journal of Consumer Research, 33*(1), 31–40.

9

The Futures Polygon Development

Antonio Pacinelli

Subjective Evaluations and Probabilities. Some Reflections

Let us say that the main issue of the method proposed in this chapter derives from the use of subjective probability and, in particular, of the conditional probability.[1] The philosophy of subjective probability is mainly due to the contributions of de Finetti, but the probabilistic evaluations provided by subjects appear in the late 1970s in the famous theorem of Thomas Bayes, in which, although it is defined as "a priori probability", its subjective nature emerges clearly.

In the theorem, the latter is corrected through the experimental evidence defined as "likelihood" (deriving from the empirical evidence of the data) and provides as an output a "posterior probability" containing the evaluations of both objective and subjective sources. The subjectivists believe that there are as many probabilities as the number of subjects, and that the probability is only a measure of the "degree of subjective belief",

A. Pacinelli (✉)
Department of Legal and Social Sciences, "G. d'Annunzio University", Chieti-Pescara, Italy

L. Moutinho, M. Sokele (eds.), *Innovative Research Methodologies in Management*, DOI 10.1007/978-3-319-64400-4_9

in particular, is "the degree of trust that a coherent individual attributes to the occurrence of an event". All bets are legitimate, except those that lead to a sure loss: such wagers are called inconsistent. The requirement of consistency is the only one to be respected, as a requirement of rationality.[2] A "belief" (which in the case of a future event can be assimilated to a prediction) is said to be incorrect when additional information reveals its falsehood. According to the subjectivists, however, this conception is wrong, because when the subjective evaluation is made, the additional information is not yet known. It is the transition from a priori to a posterior probability, which includes changes in "beliefs" (Suppes 1984).

De Finetti (1937) stated:

> Whatever the influence of observation on the future prediction, it does not imply and does not mean at all that we correct the primitive probability evaluation $P(E_{n+1})$ after it was denied by the experience, substituting for another probability $P^*(E_{n+1})$ that conforms to this experience and, therefore, probably closer to the real probability. On the contrary, it manifests itself only in the sense that when we learn from the experience that A is the result of the first *n* trials, our judgment will no longer be expressed by the probability $P(E_{n+1})$, but rather by the probability $P(E_{n+1}|A)$, namely what our initial opinion already attributed to the event E_{n+1}, considered as dependent on the result A. Nothing in this initial opinion is either repudiated or corrected: it is not the function P that has been modified (replaced by another function P^*), but the argument (E_{n+1}), replaced by $(E_{n+1}|A)$, and it is just to be faithful to the initial opinion (as it manifests in the choices of P) and consistent in our judgment, that our predictions vary when a change occurs in known circumstances.

And, he goes on: "It is clear what is required to achieve the convergence of views among observers whose initial opinions differ."

On the question whether the information can confirm or deny an opinion, furthermore, de Finetti affirmed:

> […] observation cannot confirm or deny an opinion, which is and can be nothing more than an opinion, therefore neither true nor false; observation can only give us information that is susceptible to influence our opinion. The meaning of this statement is very precise: It means that to the probability of a fact subordinate to this information – and distinct from that of the same fact not subordinated to anything else – we can actually attribute a different value.

On this, see also Suppes (1984) page 60.

The subjective data, however, is considered "scientifically weak" with respect to the objective data, and this consideration, which can certainly be shared, is justified by the lack of symmetry constraints between past and future in the former, which according to Hempel[3] are necessary conditions for the objective approach.[4]

Suppes, however, reminds us that while Hempel strongly defended his "thesis of symmetry", according to which explanation and prediction are structurally identical, Hume stated that "there is no logic guarantee that the future is similar to the past, or more generically that the unknown is similar to the known" (Suppes 1984).

The Scottish philosopher asserts that if we think of past events as a guide to making predictions about future events, then the future must necessarily retrace the past. But, in his works, Hume demonstrated that it is equally logically correct to assume a universe in which past physical laws do not coincide with the present ones and that are not uniform in any area of the space.[5] The foregoing considerations are reflected in the literature, where it is distinguished between "Forecast" (or "prognosis") and "Foresight". In the first case, it is required the respect of the theory of symmetry which, for clarity in our case, is between past and future. So that the methods for obtaining images of the future can be brought back to analysis and extrapolations of trends and relationships derived from historical objective data. The second case does not require the respect of the symmetry conditions and relies on methods that use subjective judgements (without renouncing to the support of the objective data). Even though subjective probabilities are unanimously considered "scientifically weak", they can be reinforced by methods that, in addition to using several subjective evaluations, use the procedure of "panels" which, by means of subsequent iterations, facilitates the convergence of opinions towards an evaluation interval shorter than the initial one (consensus). Another problem of subjective probability concerns "exchangeability", because a succession of tosses can respect the principle of exchangeability, but the same is harder to imagine if we consider events. Indeed, de Finetti himself states that: "A sequence of events is said, intuitively, exchangeable if the order in which the events occur does not affect their probability." That is a rather difficult situation when dealing with temporal events.

Even in the light of its limitations, the advantages of subjectivism are undeniable; for example, you can use or not the observed frequencies, or you can assign a probability to a completely new event and therefore without having objective data.

The development of social project scenarios, in particular, draws on three subjective sources: experts, representatives, and citizens (or end users). From the last two, we get the "desiderata" that represent the images of the normative scenario, while the first ones provide assessments at different stages both in the explorative scenarios and in the normative ones.[6] The experts are involved especially in the project scenario, which implies the tactical choices needed to approach trend extrapolations (explorative scenarios) towards the desired images (normative scenario).

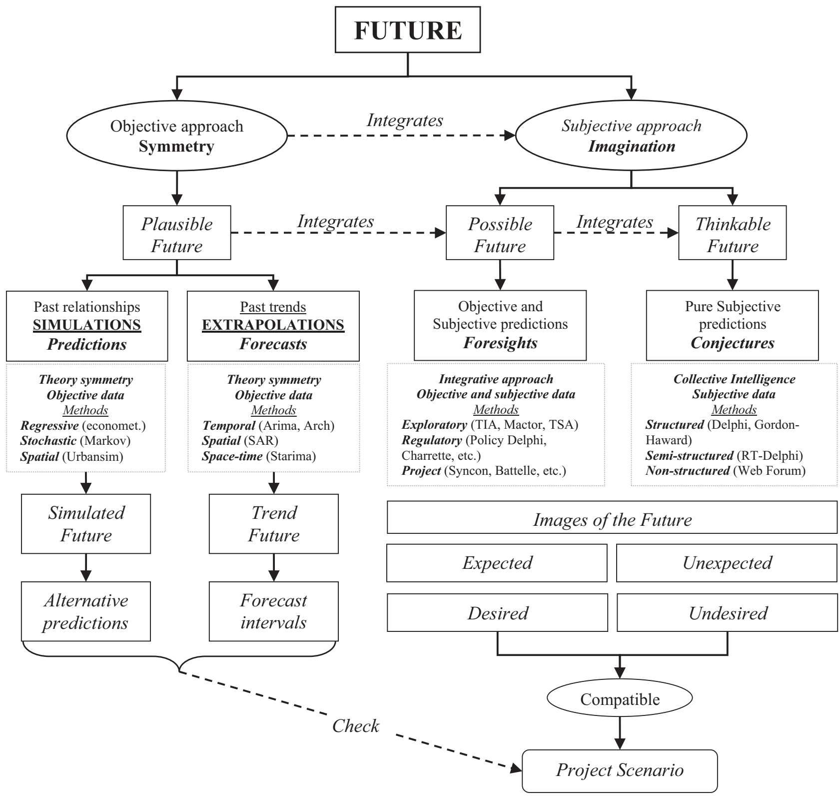

Fig. 9.1 Approaches and methods for the study of the future

A comprehensive scheme, with the objective and subjective methods for studying the future, is reported in Fig. 9.1 (see also Pacinelli 2012).

Subjective Impact Approaches

It is appropriate to precede the description of the Futures Polygon (FP) method by some reflections, which are useful for locating it within the area of subjective impact analysis. Methods for the study of event impacts can be assigned to at least three different areas:

- consequential impact events when one or more events in a system happen: *Event Impact*;
- one or more events impact on a trend: *Trend Impact*;
- impacts among several events: *Cross-impact*.

Event impact analysis treats consequential impacts arising from one occurring event, which can be expected or unavoidable (as for some natural phenomena) or can derive from an intervention, political, or other. It is possible to include the Futures Wheel (FW) (Glenn 1972) and its derivative and complementary FP method (Pacinelli 2006) into this area. The FP has recourse to the Delphi procedure in order to obtain probability evaluations of each Futures "wheel" event and to transform them into a polygon. The aim of the FP is to graphically depict the occurrence probability of the whole set of events obtained by the FW.

Trend impact analysis (TIA) is designed to study the impact of possible future events on the trend of certain quantitative or qualitative variable (Gordon 1994). The approach governing these methods consists of correcting the extrapolated trend of one or more phenomena, on the basis of subjective evaluations which are formulated by experts on the effects that a determined set of future events may produce on that trend. The TIA can be integrated by other forecasting methods based on models.

Cross-Impact Analysis allows consideration of the mutual impacts of a set of events with each other, generally collected into a matrix. Cross-impact, based on the study of interactions among a set of events, is the richest among these methods, each of which has its own goal, depending on the objective for which it was created: correcting the probabilities obtained by a Delphi (Dalkey and Helmer 1963), tracing trajectories towards the future (Kane 1972), or building scenarios (Brauers and Weber 1988).[7] Cross-impact, in turn, can be classified in three different types, based on the approach used by the techniques which are part of it: a *simulative* approach, a *heuristic* approach, or an *optimization* approach.

The idea of the FP stemmed from reading about the FW in Futures Research Methodology 1.0 (Glenn 1994) and realizing that the FW lacked the concept of evaluating the likelihood of the forecasted impacts, an indispensable element in exploring the future. The inventor of the FW regarded *unanimity*[8] as a good indicator of the plausibility of events/impacts on which there is agreement, since the unanimity can be interpreted as a "guarantee" that the impact will happen within a "realistic temporal horizon". The plausibility judgement should precede any forecasting or scenario because if the event is not plausible, it does not fall within the "cone of plausibility" (someone defines it as the forecast cone), which is a fundamental instrument of investigation about the future. The previous observations stimulate some questions:

- What is the probability that the plausible events have to happen within a certain *temporal horizon*?
- How many years does the system of interest require to register a first reaction to the impact?
- How many years does the impact intensity require to get to its maximum? How long does the impact last? What is the impact consolidation level? (as in Gordon's *Trend Impact Analysis*, 1994).

Two complementary problem areas emerge from the FW approach:

- the evaluation of the probability of an "impact scenario" generated by the FW;
- the determination of a "realistic temporal horizon" for the results of the FW.

From the previous, two different FP versions originate from them. To give an answer to the above questions, a first method, complementary to FW, was created with the capacity to give indications about the probability of the "impact scenario" (for each event and for the overall scenario), while a second version was finalized to yield a realistic temporal horizon evaluation. The method is called FP (Pacinelli 2006) because the different probabilities of the events/impacts obtained by a FW and their own different years of occurrence generate a polygon if connected with a broken line.

The Futures Polygon

FP should be applied at the conclusion of the phases related to the scenario-building activity and not after a single event impact evaluation. The principal characteristic of the event impact methods, as for example the FW and its derivatives, is to build the scenario expansion from the events considered in it. In other words, each of the events which constitute a scenario can activate elements in its turn which make the happening of other events most probable. The application of event impact, with a set of *n* events constituting the scenario, generates for each of them an expansion of the scenario itself because other primary-order impacts, deriving from the previous ones, are added in and the expansion can increase again if we also consider the second-order impacts. The process is the following:

- First of all, the "wheel of futures", obtained by the FW, is adopted. The application of the rule of unanimity for the plausibility evaluation of the impacts has eliminated the events on which there is no agreement among participants, so the work continues only on the impacts which are considered plausible. The requirement to work only on plausible events is really important because it excludes situations in which some experts could value impossible (or nearly) certain impacts;
- When the result of the FW is ready, each FW participant (or a new participant panel) is asked to attribute a probability of occurrence to the

recognized impacts, following the well-known Delphi method (Dalkey and Helmer 1963). The group could in fact agree, during *face to face* discussion, on single probabilities (or intervals) to be assigned to the impacts, or subjective evaluations can be reinforced by structured communication procedures, as for example by means of the Delphi. The question, addressed by the facilitator, will be formulated as following: "What probability do you attribute to the occurrence of each plausible event deriving from the Futures Wheel within the defined time horizon?"

- When the assessment of the probability of occurrence within the settled date has been done for all the impacts of the FW for a given event, the graphic representation is constructed, using a polar reference system. An O point, called the "pole", is fixed on the plane (it corresponds, in this case, to the event whose future implications must be studied), and a length, L, is chosen, with some value suitable for the size of the final graphic. The set of the n impacts is then represented as a collection of n radii, departing from the pole, evenly spaced in angle (i.e. the angular spacing between impact radii must be $360^{\circ}/n$). The length of the radius representing an impact is proportional to the probability of the impact, namely set equal to pL, where p is the probability of the impact on which the group reached a consensus (the final interquartile range or the median of the Delphi). Of course, being $p \leq 1$, we have $0 \leq pL \leq 1$. After this, the endpoints of the n radii are connected to form a polygon, contained within the circle of radius L (corresponding to a probability of 1), and having the general shape of Glenn's Wheel.

The higher the probabilities of the impacts are, the more the polygon will approach its maximum area, theoretically reachable when all the events are considered certain. Therefore, the proposed method can be also used to evaluate the result obtained by the FW. The consideration of an impact as plausible by a group of experts does not imply a high probability for it to take place within a certain time horizon, but it only implies that it is plausible if the principal event happens.

Starting from Fig. 9.2, it is easy to build a first measure for the evaluation of the scenario as a whole, for example by making the ratio between the sum of the "degrees of possibility" (from 0 to 100) of the impacting events $\Pr(\mathit{impact})_j$ and the number of impacting events—say K—multiplied by 100. The use of the degree of possibility with values from 0 to

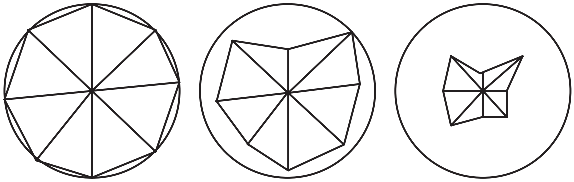

Fig. 9.2 Examples of impacts evaluated as "certain", "very probable", or "poorly probable"

100 instead of the probability (which is from 0 to 1) allows us to avoid the classical constraints of the latter and is more easily understandable for the participants.

So, by considering a FW with radius 1·100 and the impacts k_j (with $j = 1, 2, \dots, K$), each of which is associated with a degree of possibility of occurrence $\Pr(k_j)$—ranging from 0 to 100—within the given time horizon, we have:

$$ESFP = \frac{\sum_{j=1}^{K} \Pr\left(k_j\right)}{K \cdot 100}$$

The measure *ESFP* (Evaluation of the Scenario of the Future Polygon) is a simple composition ratio that provides the usual relative frequency, with a range from 0 (all the impacts of the scenario are impossible, i.e. have zero probability of occurrence) to 1 (all the impacts are certain, i.e. all have probability equal to 1).

In the first scenarios of Fig. 9.3 (left), the ratio *ESFP* is equal to 1, since each impact has probability 1 to occur. In the second case (centre), the value of *ESFP* is greater than zero and less than one, while in the third case (right), the measure is equal to zero.

Among the advantages of the FP, there is (as it is also for the FW) the possibility of expanding the scenarios, so as to show both the direct and indirect consequences of its occurrence. Further, the FP allows the evaluation of the probability that the whole set of impacts happens within a certain time horizon or the evaluation of within which temporal horizon all the impacts, or a part of them, individuated by a FW, will happen. Another strength of the FW retained by the FP is that by hypothesizing

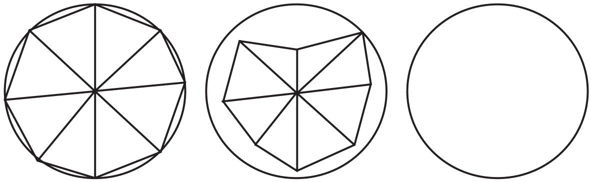

Fig. 9.3 Impacts which are evaluated as "max probable", "very probable", or "min probable"

an intervention of almost any kind, the flexibility of the approach allows the collection of information on possible impacts in each of the communities concerned, using the *mediated participation* (through the representatives) and/or the *direct participation* (through the citizens).[9] However, it is worth noting that the FP is more suitable for the *technical participation*, which is that applied through the experts (Pacinelli 2007), given the difficulties deriving from giving probability evaluations. The principal weaknesses of the FP, not present in the FW because it does not use probability evaluations as in all the other impact methods, are those relating to the use of subjective conditional probabilities. The use of the methods for the convergence of the judgements of the expert about the probabilities (as the Delphi method and its derivatives) reinforces subjective evaluations. In fact, their application generates a collective intelligence (Lévy 1994) with high level of "share", but it is very difficult to obtain subjective conditional probabilities evaluations which are coherent with the Bayesian standards (Nair and Sarin 1979; Moskowitz and Sarin 1983).

The Futures Polygon for the Scenario Expansion

The FP is complementary to the FW method and therefore, having been created in symbiosis with it, their integration is implicit. In any case, the FP and the FW, integrated with other methods, allow the expansion of the scenarios extending the temporal horizon and enlarging the funnel of the considered set of events. Below is an application in which the Focus Group, the Policy Delphi, the FW, and the FP are used in an integrated approach, applied to the development and expansion of a scenario on the job possibilities of disadvantaged workers. In Fig. 9.4, it is possible to observe a hypothesis of integration of methods, in which the output of one method is used as input for another method, which goes from the building of the scenario to its expansion.

Glenn (1972) with the FW and Gordon (1994) with the TIA afforded the opportunity to move from the theme of future "impacts", which had stagnated on the cross-impact for a long time, towards alternative approaches. In fact, the first proposes a method to individuate impacts (unanimously considered as plausible) which can occur as the expression of a certain event and is therefore related to an event impacting on a system, for example on a well-defined social or environmental system. The second, moving attention to the impacts that a set of events has on one or more time trends, allows focusing the attention towards an important concept: the *reaction time*. In fact, the experts of a TIA, among other things, are asked to formulate a subjective evaluation of the "primary impact time (or year)", referring to the "reaction time" that is the time to first noticeable impact (expressed as number of years, months, or what-

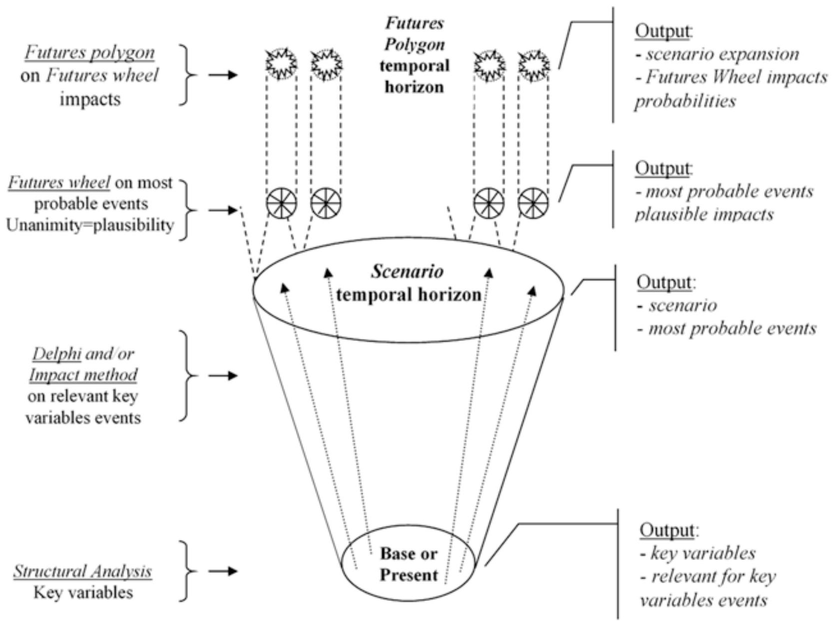

Fig. 9.4 Futures Polygon for scenario expansion

ever necessary for the trend to produce a first signal of reaction). It is obviously possible to go back to the *reaction time* from the *primary impact year* for each of the impacts obtained by the FW. Having the reaction times allows one to promptly programme the corrective interventions, before the negative effects of the event manifest on the social tissue.

Lastly, two FP variants are proposed, both starting from the FW. While the first version, described above, brings the FW towards the scenario expansion, up to a certain time horizon, the second type of FP studies the temporal horizon of each of the impacts individuated by the FW. In the second version, the length of the radius of an impact is set proportional to the reaction time for that impact, producing a reaction time representation for each event, rather than the probability representation. It is then appropriate to calculate the average reaction times and highlight the average year by a circumference. See Fig. 9.5 for an example (Table 9.1).

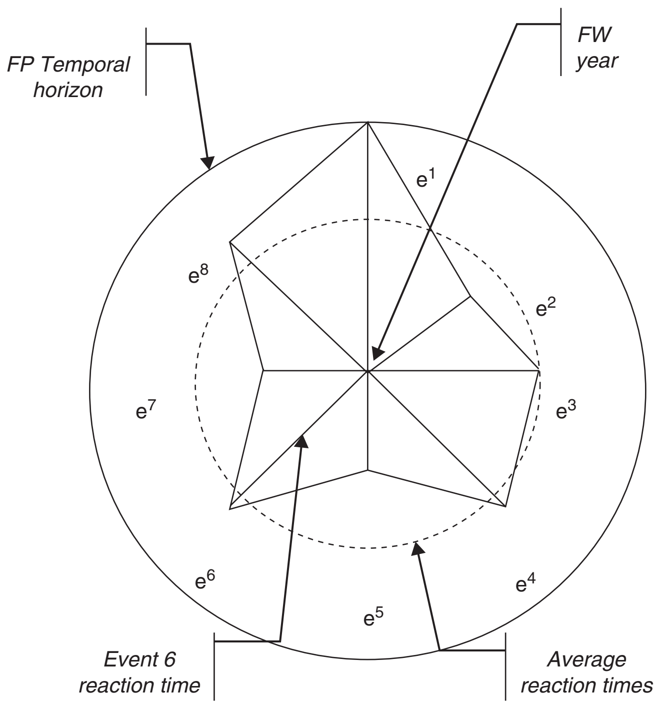

Fig. 9.5 Second version of the Future Polygon

Table 9.1 Synthetic recap of the Futures Polygon method

Phases	Description of Futures Polygon activities
1°	Participant selection and meeting organization
2°	Instructions about the method (probability evaluations) and acquisition of Futures Wheel results
3°	Inclusion of the impacts obtained by Futures Wheel first iteration questionnaire and pre-test
4°	Delphi evaluation on impacting events probability individuated by Futures Wheel
4_1°	Administering and collection of first iteration questionnaire according to Delphi procedure
4_2°	Elaboration of first iteration questionnaire data (interquartile intervals)
4_3°	Drafting of second iteration questionnaire
4_4°	Administering and collection of second iteration questionnaire
4_5°	Elaboration of second iteration questionnaire data (intervals and motivations)
4_6°	Drafting of third iteration questionnaire with intervals and spaces for motivations and counter-motivations
4_7°	Administering and collection of third iteration questionnaire
4_8°	Elaboration of third iteration questionnaire data (intervals, motivation, and counter-motivations)
4_9° [a]	Calculation of interval median events (possibly corrected by the Gordon and Hayward method) Gordon and Hayward 1968.
5°	Use of probability obtained by Delphi for Futures Polygon construction
6°	Group consideration of event probability evaluation obtained by Futures Wheel
7°	Results presentation and comment

[a]It should be noted that Phase 4 could be replaced by a Real-Time Delphi

An Application: The Use of the Futures Polygon in Integration with Other Methods

The experience is derived from a survey on "Identification of niches work demand for poor in the Chieti-Ortona area" (Italy), carried out within the European Project Equal-Linea (years 2002–2003). The action aimed for by the research was:

> to favour social and working integration and re-integration of disadvantaged categories and/or of those population groups with higher exclusion risk from the local work market by removing individuated obstacles and barriers.

The research focus comprises dropped-out young people without useful qualifications at a professional level, jobless adults (over 40 or so), and physically or mentally disabled people. Various methods in integration, such as Focus Group, Policy Delphi, and FW, were used. The results of the FW were subsequently used to apply the FP technique. The study was structured into different phases, on the basis of a research design in which each output of a method was used as input for the following one. A synthetic design of the research plan is illustrated in Fig. 9.6.

The following issues emerged from the Focus Group[10]: territorial unease, work environment awareness, coach work project, social assistant as tutor, job sharing, working integration, and work demand/offer match. A vast part of the information collected by the Focus Group flowed into the Policy Delphi. During the preliminary survey phase, a heterogeneous panel was constructed, formed by representatives from the communities in the study area, entrepreneurial associations, and professional orders; also technical competences (expertise) and local reality knowledge criteria were applied. The age of the participants goes from 30 to 55 years. Further, the involvement of a control group was considered secondary because the subsequent FW implementation was deemed sufficient to analyse the validity of the emerged results.

During the preliminary survey phase, the research group, on the basis of the results of the focus group, produced a first list of items/

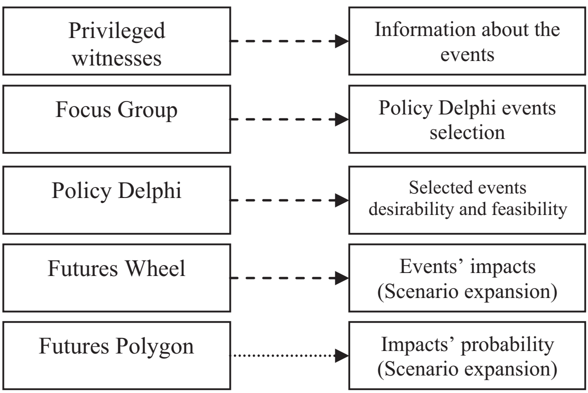

Fig. 9.6 Methods integration

events which, once analysed, selected, and quantitatively reduced, resulted in a definitive drafting. Twenty-eight intervention hypotheses, divided between option items, or rather actions (projects or proposals), and goal items (objectives), were submitted to the panel. As regards the evaluation scale, we adopted the desirability and technical or political feasibility scales, comparing potential developments (feasibility evaluations) and desirable developments (desirability evaluations). In this way, we moved from a "technical" or political forecasting dimension (expected or probable future analysis) to a "political" forecasting dimension (shared hypothesis about the desired future individuation). For brevity, the results of all the considered events are not reported in the present work. As an example, we show the event which prevailed in terms of desirability and feasibility, that is: *Working guidance days in schools*.

This phase of the research proceeded on the basis of the integrative approach, so using the Policy Delphi results as inputs for the FW (Di Giandomenico 2004). The panel contained 12 participants who were invited to project themselves into the future and to imagine consequences (impacts) that the events, if realized, would have generated into the analysed territorial reality. With respect to the original protocol of the method, the research was carried on writing the name of the phenomenon (or event) into the centre of a piece of paper and drawing little radii containing the primary-order impacts (or consequences) at the extremity. Afterwards, secondary impacts forming the second ring of the wheel were reported for each primary-order impact. The procedure went on until a useful and clear representation of the implication related to the event was obtained. All the panel members agreed about the real causality links between the main event and the primary impacts, and between the primary impacts and the successive ones, in respect of the unanimity rule. In Fig. 9.7, for practical reasons (readability of the diagram and aims of the present work), we report only the impacts of the first order (the application went on until the third order of impact).

Finally, by using the FW, we applied the FP, with a time horizon of eight years. The results are reported in Fig. 9.8.

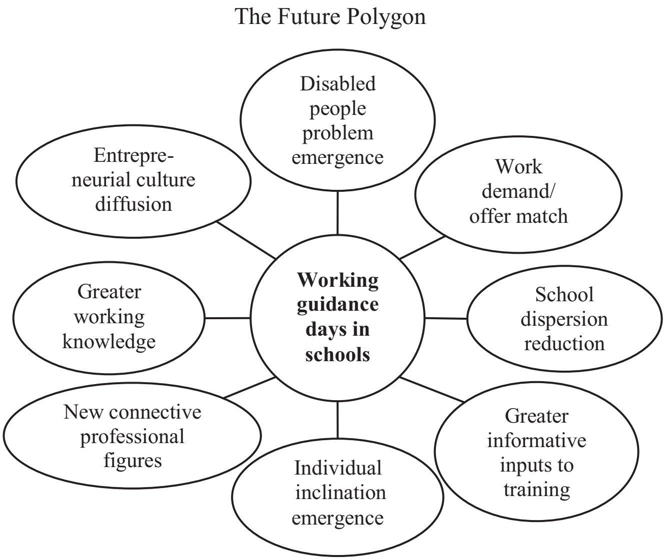

Fig. 9.7 Futures Wheel. Primary order impacts. Event: "Working guidance days in schools"

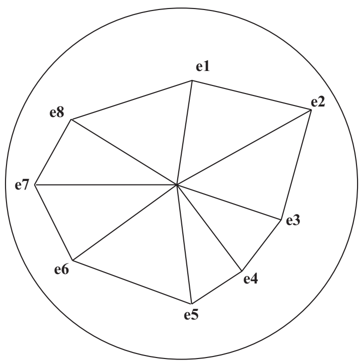

Fig. 9.8 Futures Polygon "Working guidance days in schools"

Notes

1. It is, indeed, difficult to obtain subjective conditional probabilities, which are coherent with the Bayesian assumptions (Nair and Sarin 1979; Moskowitz and Sarin 1983).

2. It can be shown that this conception of probability satisfies the Kolmogorov's axioms (Suppes 1984).
3. The German philosopher Carl Gustav Hempel is known for formulating the nomologic-deductive model, in which the set of knowledge necessary for explaining a phenomenon is said *explanans* and includes one or more cover laws and initial conditions. Whereas the phenomenon that the explanation is to be provided is said *explanandum*. Following Hempel, the explanandum derives from an explanans that consists of a class of laws $L_1,\ldots,L_n$ and some initial conditions $C_1,\ldots,C_k$.
4. In such cases, the use of methods as the Delphi, which facilitates the convergence of probabilistic evaluations, strengthens the subjective judgements, because it reduces the dissent by generating a collective intelligence with high levels of consensus (Pacinelli 2002).
5. In the exploration of the future, Bruno de Finetti distinguished between *prediction* and *foresight*. He believed that "say before" (prediction) is something other than "see before" (foresight), placing the subjective moment upstream in the first case and downstream in the second. In a foresight, all the available information must be taken into account, whether generated by objective or subjective data.
6. For the "Desiderata Stability", see Di Zio and Pacinelli 2009.
7. Particularly, cross-impact simulates different decisions compared to different future situations, aiming to determine optimal and/or preferable strategies.
8. The *unanimity rule* is an important point of reference for people using participatory methods (Pacinelli 2007), but it is not sufficient to make a foresight, which needs an occurrence probability for each event/impact.
9. On this, see among others Pacinelli 2002.
10. The work carried out in a focus group can generate interesting additional information if it is supported, as was the case in the present example, by its complementary technique called "debate evaluation".

References

Brauers, J., & Weber, M. (1988). A New Method of Scenario Analysis for Strategic Planning. *Journal of Forecasting, 7*(1), 31–47.

Dalkey, N. C., & Helmer, O. (1963). An Experimental Application of Delphi Method to the Use of Experts. *Management Science, 9*(3), 458–467.

De Finetti, B. (1937). La Prévision: Ses Lois, logiques, Sef Sources Subjectives. *Annales de l'Institut Henri Poincaré, 7*, 1–68.
Di Giandomenico, M. (2004). "L'applicazione della tecnica Futures Wheel finalizzata alla individuazione degli impatti sulla realtà chietina-ortonese dati dalla realizzazione di due soluzioni: "giornate di orientamento al lavoro presso le scuole" e "progetto coach Work." In *Fabbisogni lavorativi delle imprese dell'area chietino-ortonese*, Equal-linea, edited by Cannarsa, Vasto, 56–72.
Di Zio, S., & Pacinelli, A. (2009). Desiderata Stability. Methodological Considerations. In J. Kultalahti, I. Karppi, O. kuktalahti, & E. Todisco (Eds.), *Globalisation* (pp. 99–118). Finland: East-West Books Helsinki.
Glenn, J. C. (1972). Futurizing Teaching vs Futures Course. *Social Science Record, 9*(3), 26–29.
Glenn, C. J. (1994). The Participatory Methodology. In J. C. Glenn & T. J. Gordon (Eds.), *Futures Research Methodology*. Washington, DC: The Millennium Project, American Council for United Nations University.
Gordon, T. J. (1994). The Trend Impact Analysis. In J. C. Glenn & T. J. Gordon (Eds.), *Futures Research Methodology*. Washington, DC: The Millennium Project, American Council for United Nations University.
Gordon, T. J., & Hayward, H. (1968). Initial Experiments with the Cross-Impact Matrix Method of Forecasting. *Futures, 1*(2), 100–116.
Kane, J. (1972). A Primer for a New Cross Impact Language- KSIM. *Technological Forecasting and Social Change, 4*(2), 129–142.
Lévy, P. (1994). *L'intelligence collective. Pour anthropologie du cyberspace*. Paris: La Découverte.
Moskowitz, H., & Sarin, R. K. (1983). Improving the Consistency of Conditional Probability Assessments for Forecasting and Decision Making. *Management Science, 29*(6), 735–749.
Nair, K., & Sarin, R. K. (1979). Generating Future Scenarios—Their Use in Strategic Planning. *Long Range Planning, 12*(3), 57–61.
Pacinelli, A. (2002). Sull'uso di metodi soggettivi nella Pianificazione Sociale Partecipata: verso la Democrazia Continua. *Statistica & Società, 1*(2), 23–28.
Pacinelli, A. (2006). A Complementary Method to Future Wheel: The Future Polygon. *Futures Research Quarterly, 22*(1), 71–78.
Pacinelli, A. (2007). *Metodi per la ricerca sociale partecipata*. Milano: Franco Angeli.
Pacinelli, A. (2012). I metodi della previsione. In R. Poli & S. Arnaldi (Eds.), *La previsione sociale. Introduzione allo studio dei futuri* (pp. 149–163). Roma: Carocci editrice.
Suppes, P. (1984). *La logica del probabile: un approccio bayesiano alla razionalità*. Bologna: Clueb.

Index[1]

[1]Note: Page numbers followed by 'n' refers to note.

L. Moutinho, M. Sokele (eds.), *Innovative Research Methodologies in Management*,
DOI 10.1007/978-3-319-64400-4