Towards Global Experimental Design using Bayesian networks

Case studies on modeling sensory satiation

PHAN Vân Anh

Thesis committee

Promotor

Prof. dr. ir. Martinus A.J.S. van Boekel
Professor of Product Design and Quality Management
Wageningen University

Co-promotors

Dr.ir. Matthijs Dekker
Associate professor, Food Quality and Design
Wageningen University

Dr. Ursula Garczarek
Data scientist
Unilever Food and Health Research Institute, Vlaardingen, The Netherlands

Other members

Prof.dr. Hal MacFie, Reading University, England
Prof.dr.ir. Hans C.M. van Trijp, Wageningen University
Prof.dr. Hendriek C. Boshuizen, Wageningen University
Prof.dr.ir. Kees de Graaf, Wageningen University

This research was conducted under the auspices of the Graduate School VLAG (Advanced studies in Food Technology, Agrobiotechnology, Nutrition and Health Sciences).

Towards Global Experimental Design using Bayesian networks

Case studies on modeling sensory satiation

Phan Vân Anh (V.A. Phan)

Thesis

submitted in fulfillment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof.dr. M.J. Kropff,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Friday 26 April 2013
at 4 p.m. in the Aula.

Phan Vân Anh (V.A. Phan)
Towards Global Experimental Design using Bayesian networks. Case studies on modeling sensory satiation
156 pages

Thesis, Wageningen University, Wageningen, NL (2013)
With abstract, summary in English, and summary in Dutch

ISBN: 978-94-6173-537-9

Abstract

Food science problems are complex. Scientists may be able to capture more of the complexity of an investigated theme if they were able to integrate related studies. Unfortunately, individual studies are usually not designed to allow such integration, and the common statistical methods cannot be used for analyzing integrated data. The modeling technique of Bayesian networks has gained popularity in many fields of application due to its ability to deal with complexity, but has emerged only recently in food science. This thesis used data from experiments on sensory satiation as case studies. The objective was to explore the use of Bayesian networks to combine raw data of independently performed but related experiments to build a quantitative model of sensory satiation.

Methods

This thesis started with introducing the theoretical background of Bayesian networks to food science. The available data from various independent experiments on sensory satiation were then examined for their potential to be combined. Finally, the outcomes obtained using Bayesian networks on a single dataset were compared with the published outcomes of the respective study, in which classical statistical procedures were used to analyze the data.

Results

Two hurdles were identified when combining the data of related studies that were performed independently and without the intention of combining their data. The first hurdle was a lack of essential information for reliable estimations of parameters of the combined model network. This information could be obtained by deriving it from existing information in the individual studies or by performing extra experiments; these practices are, however, not always possible. The second hurdle was a possible conflict in causal relationships underlying the individual experimental designs, which can cause misleading analyses of the combined dataset. This was the case for some experiments that involved the control of secondary explanatory variables. As such,

an approach termed as Global Experimental Design was proposed in this thesis as a solution to overcome these hurdles. This approach emphasizes the building of an overall network prior to designing individual studies.

In comparison to using the classical statistical procedures, more information can be extracted using Bayesian networks. This technique could make use of the domain knowledge in a transparent manner as well as empirical data with missing values.

Conclusions

It is possible to combine raw data from related studies for a meaningful analysis if effort is made in the phase of experimental design. The approach of Global Experimental Design outlines this phase with the building of an overall network. By using Bayesian networks as a tool for exploratory analysis, scientists are able to gain more insights into a research domain.

Table of contents

CHAPTER 1

General introduction

This thesis explored the use of Bayesian networks, a modern modeling technique, in the field of food science. The exploration was performed with the data on food satiation that were already available. The thesis was part of a larger project entitled "Sensory specific satiation: linking product properties to obesity prevention". Various controlled experiments were independently designed and conducted to understand **sensory satiation**, i.e. how different sensory aspects influence satiation. In these experiments, researchers manipulated the composition of some sensory stimuli or oral/ nasal exposure to sensory stimuli during food consumption. Their designs involved information on sensory perception (e.g. taste and aroma) and oral processing characteristics (e.g. bite size and bite frequency).

This introduction starts with the definition of satiation and satiety. It is followed by an overview of the complexity of satiation to demonstrate the need of modeling to understand this process. Bayesian networks are then briefly presented as a potential tool for modeling food-related problems. The chapter concludes with the objective and outline of this thesis.

1.1 Food intake: satiation and satiety

There are two processes involved in the consumption of food: satiation and satiety (Blundell et al., 1988). Satiation is the process that develops during a course of eating (meals or snacks) and brings this course to an end (meal termination). Satiety is the process that takes place after an eating course and inhibits the start of the next eating course (meal initiation). As such, the feeling of hunger is reduced with the development of satiation and is suppressed by satiety. It is thus expected that satiation determines the meal size (how much food is eaten in a meal), and satiety determines the meal frequency (how many meals are eaten a day).

Blundell et al. (1988) have illustrated the processes of satiation and satiety by the "satiety cascade" (Figure 1.1). These authors have also identified four mediating processes that have control over satiation and satiety: sensory, cognitive, post-ingestive, and post-absorptive.

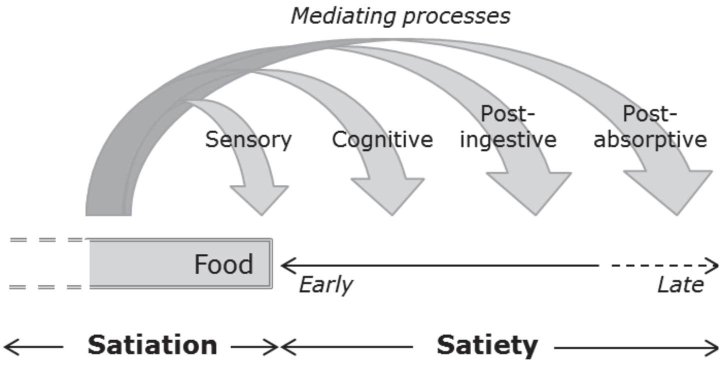


Figure 1.1: The satiety cascade of Blundell et al. (1988).

1.2 Satiation: complexity and the need of modeling

Satiation or meal size results from the choices of what to eat and drink, and of how much to consume (Booth, 1990). So, what influences these choices? We can view the influencing factors belonging to three groups: the Actual, the Inner, and the Outer (Figure 1.2).

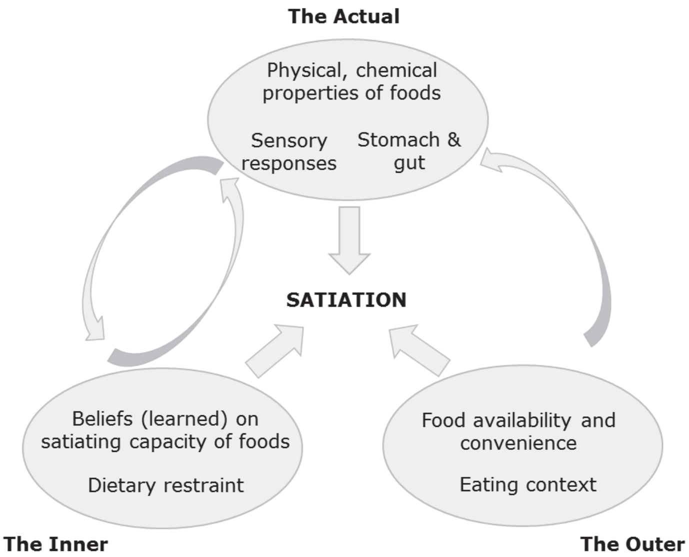

The Actual
Physical, chemical properties of foods
Sensory responses
Stomach & gut
SATIATION
Beliefs (learned) on satiating capacity of foods
Dietary restraint
Food availability and convenience
Eating context
The Inner
The Outer

Figure 1.2: The various influences that contribute to the complexity of satiation.

The "Actual" factors concern the responses of human senses to the food, the food itself, and also the stomach and gut signals during the consumption. A food presents various stimuli to different human senses: vision, hearing, touch, smell, and taste. The overall sensory perception strongly affects the liking of the food (palatability or pleasure); and the liking in turn can influence how much of the food is eaten (Sorensen et al., 2003). When a food is consumed until satiation, the perceived pleasantness decreases specifically for this food; it does not change however, or decreases much less, for other (uneaten) foods. This phenomenon is called "Sensory specific satiation/satiety, SSS" (Rolls, 1986). In addition, the chemical and physical properties of the food can directly influence how the food is processed in the oral cavity. For example, different food textures ranging from liquid to solid determine the level of mastication needed (or not at all). This difference can lead to a short or rather long oral residence time, or different eating rates (Viskaal-van Dongen et al., 2011). A high rate of eating is strongly correlated with a high intake, as shown in various studies (Spiegel et al., 1993; Andrade et al., 2008; Zijlstra et al., 2010; Viskaal-van Dongen et al., 2011). It is also believed that a longer residence in the oral cavity enhances the oro-exposure to the sensory signals, hence contributing to the development of an earlier satiation (de Graaf, 2012). To decide on whether to continue or stop eating a food, the brain uses not only the sensory signals (sensory processes) but also the signals from the stomach and the gut hormones (metabolic processes). The state of hunger prior to a meal influences the amount to be consumed (Decastro, 1988); this hunger state is controlled by some gastrointestinal hormones, e.g. ghrelin, leptin, and glucose. During the meal, the degree of stomach distention and the release of some other hormones, such as cholecystokinin (CCK) and glucagon-like peptide 1 (GLP-1), trigger brain-signaling of satiation (Woods et al., 1998; Blom et al., 2004). Liddle et al. (1985) observed that human plasma CCK levels increase seven-fold during meals, peaking between 10 and 30 min after meal initiation and gradually falling when the meal ends.

The “Inner” factors account for the contribution of human cognition to the development of satiation. The sensory signals during eating are linked to the metabolic consequences. These learning processes shape the eating pattern of each individual (de Graaf & Kok, 2010). It is believed that sensory attributes of a food (e.g. taste, smell, and texture) are associated with its quality and energy content, and thus guide food intake behavior (Woods, 2009). In other words, humans have unconsciously learned about the satiating capacity of different foods. These learned associations (beliefs) are built-in and automatically affect the food choice and the amount to be eaten. Another cognitive aspect that plays an important role in determining the meal size is dietary restraint, i.e. controlling body weight by limiting food intake (Vanstrien et al., 1986). These cognitive factors (“Inner” factors) interactively give direct feedback to the sensory and metabolic processes (“Actual” factors).

The “Outer” factors encompass the eating environment. It could be the availability of foods or the ambiance of the meal. For example, portion size has a robust, positive effect on food intake (Kral & Rolls, 2004; Piernas & Popkin, 2011). Stubbs et al. (2001) showed that increasing the variety of foods that are identical in composition but differ in sensory perception can increase food and energy intake. This is explained by the sensory specific satiation/satiety phenomenon. The unchanged (or less changed) pleasantness towards uneaten foods (or not yet exposed flavor) encourages us to eat more when presented with greater variety. The amount of food eaten can increase with the presence of distracting factors, such as friends or family (Hetherington et al., 2006), or television or music (Bellisle et al., 2004; Stroebele & de Castro, 2006; Temple et al., 2007). The eating environment (“Outer” factors) itself also possibly affects the sensory and metabolic processes (“Actual factors”) due to distraction.

As described above, the development of satiation is a **highly complex process.** It involves a large number of variables and many of these are interrelated. The capacity of human beings for causal reasoning with severable interrelated influencing factors in their head is limited. That is why we need a mathematical model to extend our capability in that respect. As our problem also possesses a high degree of variability (natural variation) and uncertainty (lack of knowledge) finally a statistical model is needed to capture and communicate the insights in satiation.

1.3 Modeling with Bayesian networks

Machine learning techniques are known as the convergence of artificial intelligence and statistics. Unlike classical statistical analysis, with which researchers must formulate and test each hypothesis individually, these modern techniques can automate both hypothesis generation and testing process (Cunningham, 1995). A Bayesian network model has two components: graphical (model structure) and probabilistic (model parameters) (Heckerman, 1995). The graphical nature makes it easy to grasp the overall picture as the causal relationships among variables are visualized. The probabilistic nature makes it transparent to reason through the problem as the relationships are quantified by conditional probabilities. Therefore, this modeling technique can deal with complexity and facilitates an easy communication among model users of different scientific backgrounds.

Owing to its practical features, Bayesian networks have been increasingly applied in many fields, such as finance, medical diagnosis, and genetics (Pourret et al., 2008). Figure 1.3 shows an indication of this growth in popularity by the number of publications over time recorded in the online database "Web of Science". Two search criteria were used: i) the **topic** must include "Bayesian network", or "Bayes net", or "belief network", and ii) the **research area** excluded "Computer science" and "Mathematics". The second criterion assured that only applications of Bayesian networks in other fields were counted. Only records until 2011 were used, taking into account a possible delay in document indexing.

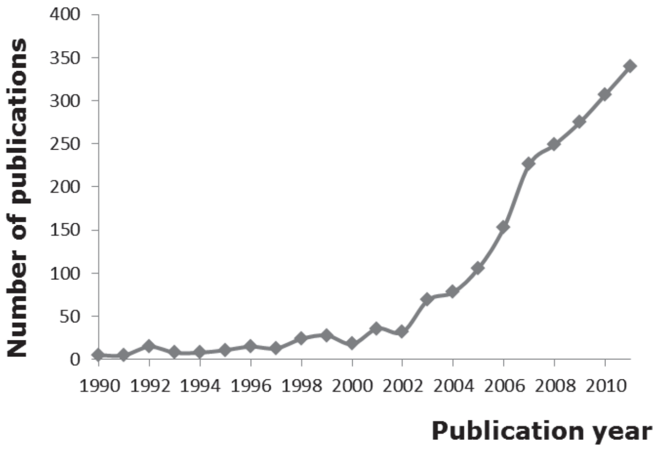


Figure 1.3: The increasing growth of Bayesian network applications in various domains as measured by the number of publications in the Web of Science (see text for more details).

Yet, this modeling technique is not so much applied in food research. An active application of Bayesian networks can be found only in the area of microbial risk assessment (Barker et al., 2002; Barker et al., 2005; Smid et al., 2012), where modeling as such has had a long tradition. Bayesian networks have recently also been used in the field of human nutrition by Mioche et al. (2011a; 2011b). In those papers, it is shown how to apply Bayesian networks for predicting fat-free mass through easily available information on sex, age, weight, and height. Food product design also appears to be an area that can potentially benefit from this modeling technique as Corney (2000) discussed with sensory and consumer data.

1.4 Objective and outline

This thesis was part of a larger project on satiation with its own specific objective. The objective of this thesis was to **explore** the use of **Bayesian networks** to **combine raw data** of independently performed but related experiments to **build a quantitative model** of sensory satiation.

The outline of the thesis is as follows. Chapter 2 introduces the theoretical background of the Bayesian network technique and its potential applications in food science. A food example was used as the basis to present the main features of Bayesian networks. Chapter 3 describes the first hurdle encountered when combining data: a lack of Structural Linking Information. Chapter 4 describes the second hurdle encountered when analyzing the combined data: a possible conflict in causal relationships underlying the experimental designs. These two hurdles need to be overcome when intending to combine data for a meaningful pooled analysis of the data. While Chapter 3 and Chapter 4 focus at a theoretical level, Chapter 5 illustrates some practical benefits of using Bayesian networks as a modeling method. Based on the same dataset obtained from a single study, this latter chapter describes what kind of extra information scientists can obtain with Bayesian network analysis as compared to with common statistical procedures. The general discussion (Chapter 6) closes this thesis with two parts. The first part communicates the approach of Global Experimental Design by synthesizing the results obtained from Chapter 3 and Chapter 4. This approach provides guidance on how to design individual related studies that allows their data to be integrated. The second part gives a conclusion on the whole of the thesis and discusses the outlook.

1.5 References

Andrade, A. M., Greene, G. W., & Melanson, K. J. (2008). Eating slowly led to decreases in energy intake within meals in healthy women. J*ournal of the American Dietetic Association, 108(7)*, 1186-1191.

Barker, G. C., Malakar, P. K., Del Torre, M., Stecchini, M. L., & Peck, M. W. (2005). Probabilistic representation of the exposure of consumers to Clostridium botulinum neurotoxin in a minimally processed potato product. *Int J Food Microbiol, 100(1-3)*, 345-357.

Barker, G. C., Talbot, N. L. C., & Peck, M. W. (2002). Risk assessment for Clostridium botulinum: a network approach. *International Biodeterioration & Biodegradation(50)*, 167-175.

Bellisle, F., Dalix, A. M., & Slama, G. (2004). Non food-related environmental stimuli induce increased meal intake in healthy women: comparison of television viewing versus listening to a recorded story in laboratory settings. *Appetite, 43(2)*, 175-180.

Blom, W., de Graaf, C., Smeets, P., Stafleu, A., & Hendriks, H. (2004). Biomarkers of satiation and satiety: A review. *International Journal of Obesity, 28*, S214-S214.

Blundell, J. E., Hill, A. J., & Rogers, P. J. (1988). Hunger and the Satiety Cascade - Their

Importance for Food Acceptance in the Late 20th-Century. *Food Acceptability*, 233-250.

Booth, D. A. (1990). Sensory Influences on Food-Intake. *Nutrition Reviews, 48(2)*, 71-77.

Corney, D. P. A. (2000). Designing food with Bayesian Belief Networks. Parmee, I.ed. , *Adaptive computing in design and manufacture*, 83-94.

Cunningham, S. J. (1995). Machine learning and statistics: a matter of perspective. In. Hamilton, New Zealand: University of Waikato, Department of Computer Science.

de Graaf, C. (2012). Texture and satiation: The role of oro-sensory exposure time. *Physiol Behav*, *107*(4):496-501.

de Graaf, C., & Kok, F. J. (2010). Slow food, fast food and the control of food intake. *Nature Reviews Endocrinology, 6(5)*, 290-293.

Decastro, J. M. (1988). Physiological, Environmental, and Subjective Determinants of Food-Intake in Humans - a Meal Pattern-Analysis. *Physiol Behav, 44*(4-5), 651-659.

Heckerman, D. (1995). A tutorial on learning with Bayesian networks. In, Technical report MSR-TR-95-06, *Microsoft Research.*

Hetherington, M. M., Anderson, A. S., Norton, G. N. M., & Newson, L. (2006). Situational effects on meal intake: A comparison of eating alone and eating with others. *Physiol Behav, 88*(4-5), 498-505.

Kral, T. V., & Rolls, B. J. (2004). Energy density and portion size: their independent and combined effects on energy intake. *Physiol Behav, 82*(1), 131-138.

Liddle, R. A., Goldfine, I. D., Rosen, M. S., Taplitz, R. A., & Williams, J. A. (1985). Cholecystokinin Bioactivity in Human-Plasma - Molecular-Forms, Responses to Feeding, and Relationship to Gallbladder Contraction. J*ournal of Clinical Investigation, 75*(4), 1144-1152.

Mioche, L., Bidot, C., & Denis, J. B. (2011). Body composition predicted with a Bayesian network from simple variables. *British Journal of Nutrition, 105*(8), 1265-1271.

Mioche, L., Brigand, A., Bidot, C., & Denis, J. B. (2011). Fat-Free Mass Predictions through a Bayesian Network Enable Body Composition Comparisons in Various Populations. J*ournal of Nutrition, 141*(8), 1573-1580.

Piernas, C., & Popkin, B. M. (2011). Increased portion sizes from energy-dense foods affect total energy intake at eating occasions in US children and adolescents: patterns and trends by age group and sociodemographic characteristics, 1977-2006. *American Journal of Clinical Nutrition, 94*(5), 1324-1332.

Pourret, O., Naïm, P., & Marcot, B. (2008). Bayesian Networks: A *Practical Guide to Applications*: Wiley.

Rolls, B. J. (1986). Sensory-Specific Satiety. *Nutrition Reviews, 44*(3), 93-101.

Smid, J. H., Heres, L., Havelaar, A. H., & Pielaat, A. (2012). A Biotracing Model of Salmonella in the Pork Production Chain. *Journal of Food Protection, 75*(2), 270-280.

Sorensen, L. B., Moller, P., Flint, A., Martens, M., & Raben, A. (2003). Effect of sensory perception of foods on appetite and food intake: a review of studies on humans. *International Journal of Obesity, 27*(10), 1152-1166.

Spiegel, T. A., Kaplan, J. M., Tomassini, A., & Stellar, E. (1993). Bite Size, Ingestion Rate, and Meal Size in Lean and Obese Women. *Appetite, 21*(2), 131-145.

Stroebele, N., & de Castro, J. M. (2006). Listening to music while eating is related to increases in people's food intake and meal duration. *Appetite, 47*(3), 285-289.

Stubbs, R. J., Johnstone, A. M., Mazlan, N., Mbaiwa, S. E., & Ferris, S. (2001). Effect of altering the variety of sensorially distinct foods, of the same macronutrient content, on food intake and body weight in men. *European Journal of Clinical Nutrition, 55*(1), 19-28.

Temple, J. L., Giacomelli, A. M., Kent, K. M., Roemmich, J. N., & Epstein, L. H. (2007). Television watching increases motivated responding for food and energy intake in children. *American Journal of Clinical Nutrition, 85*(2), 355-361.

Vanstrien, T., Frijters, J. E. R., Vanstaveren, W. A., Defares, P. B., & Deurenberg, P. (1986). The Predictive-Validity of the Dutch Restrained Eating Scale. *International Journal of Eating Disorders, 5*(4), 747-755.

Viskaal-van Dongen, M., Kok, F. J., & de Graaf, C. (2011). Eating rate of commonly consumed foods promotes food and energy intake. *Appetite, 56*(1), 25-31.

Woods, S. C. (2009). The Control of Food Intake: Behavioral versus Molecular Perspectives. *Cell Metabolism, 9*(6), 489-498.

Woods, S. C., Seeley, R. J., Porte, D., & Schwartz, M. W. (1998). Signals that regulate food intake and energy homeostasis. *Science, 280*(5368), 1378-1383.

Zijlstra, N., Mars, M., Stafleu, A., & de Graaf, C. (2010). The effect of texture differences on satiation in 3 pairs of solid foods. *Appetite, 55*(3), 490-497.

CHAPTER 2

Bayesian networks for food science: theoretical background and potential applications

Published as: V.A. Phan, M. Dekker, U. Garczarek, M.A.J.S. van Boekel. **2010**. In S. R. Jaeger & H. Mac Fie (Eds.). Consumer-driven innovation in food and personal products. Cambridge: Woodhead Publishing Limited, pages 487-513.

Abstract

Although Bayesian networks have gained popularity in many fields, they have just recently emerged in food-related problems. This technique can be used as a tool for prediction, explanation, exploration, or decision-making under uncertainty. This chapter mainly provides a theoretical background of Bayesian networks through a food example. It also discusses the advantages and challenges, as well as potential applications of Bayesian networks in food area.

2.1 Introduction

Food research is highly complex. Food technologists and researchers need to take into account not only physical and chemical interactions between food ingredients under processing, but also biological interactions between food and microorganism and those between food and the human body. Owing to its nature, we need to consider the variability and uncertainty of the system. Variability reflects natural variation whereas uncertainty represents the lack of human knowledge (van Boekel, 2008, pages 2-5). For instance, perception responses to the same odorant can vary between human subjects, or even within one subject at different psychological and physiological states (variability). Besides this, the mechanism of how odorants trigger olfactory receptors has not yet been fully understood (uncertainty). Therefore, we humans build models to simplify and approximate the real world as a way to handle complex problems.

One of the challenges of the food industry in the 21st century is to reformulate commonly eaten foods. This task has been defined in response to the dietary recommendations for lower intake of saturated fat, *trans* fat, sugar and salt (van Raaij et al., 2008). The reduction of these components requires huge research efforts to recreate the conventional flavor and texture that is desirable to consumers. As such, prediction of sensory attributes and consumer acceptance while modifying physical chemical properties of foods is a valuable tool. Deterministic models essentially ignore uncertainty and variability of complex problems. Stochastic or probabilistic approaches, however, suggest possible solutions by expressing uncertainty and variability through probability distributions (Fearn, 2004).

Recent food research has witnessed an increasing application of modern measurement techniques. Hence, more and more data are generated and food scientists need to work with large datasets. The capability of data analysis techniques to provide efficient explanations of data and explorations of implicit information is thus of importance. Cunningham (1995) has discussed this point while bringing together classical and modern statistical approaches. In classical statistical analysis, researchers must formulate and test each hypothesis individually. The information discovery process becomes time-consuming and difficult to manage. In response, machine-learning techniques, which are the convergence of artificial intelligence and statistics, have been intensively developed over the last decades. These techniques can automate both hypothesis generation and testing processes.

Bayesian networks, also referred to as *Bayesian belief networks, belief networks, Bayes nets, or causal probabilistic networks*, are one machine learning technique based on a probabilistic approach. This technique can be used as a tool for prediction, explanation, exploration or decision-making under uncertainty (Heckerman, 1995, Kjaerulff and Madsen, 2008). Bayesian networks are growing in popularity with numerous applications covering a variety of areas, such as finance, medical diagnosis, robotics, genetics, and ecology. General introductions to Bayesian networks as well as real-life case studies in these domains are presented by Pourret, Naïm, and Marcot (2008). An early application of Bayesian logic can be found in medical diagnosis (Barnett et al., 1998). A model system was developed from a database of thousand clinical findings such as symptoms, laboratory data and associated diseases. This model can predict the most likely diseases when provided with a description of new patients' data.

Despite the wide use of Bayesian networks in various fields, its presence in food-related problems has emerged very recently (van Boekel, 2004). Modeling with Bayesian networks has mostly focused on microbial risk assessment in the food production chain (Barker et al., 2005, Barker et al., 2002, Carlin et al., 2000). This kind of models was shown to add new information in a structured and simple manner (Barker et al., 2005). To the authors' knowledge, the first published effort in designing food was to build Bayesian network models relating sensory features to consumer preference (Corney, 2000). It was shown that Bayesian networks could be a valuable addition to food design and could be built from small data sets.

In short, Bayesian networks are able to handle variability and uncertainty in explaining, exploring information and particularly in predicting behaviors of systems. Although it is promising in solving problems in food research, Bayesian networks have not yet garnered enough attention within the food science community. This is probably because available tutorials on this technique often require an advanced mathematical background that few food experts have. The present paper aims to make ideas and techniques of Bayesian networks accessible to food scientists by describing a Bayesian network model using a food example (2.2); showing benefits of the model once it has been built (2.3); and explaining the theories behind Bayesian networks (2.4, 2.5, and

2.6). We discuss then the advantages and challenges, as well as potential applications of this technique in food area (2.7), and finally provide sources for further reading (2.8).

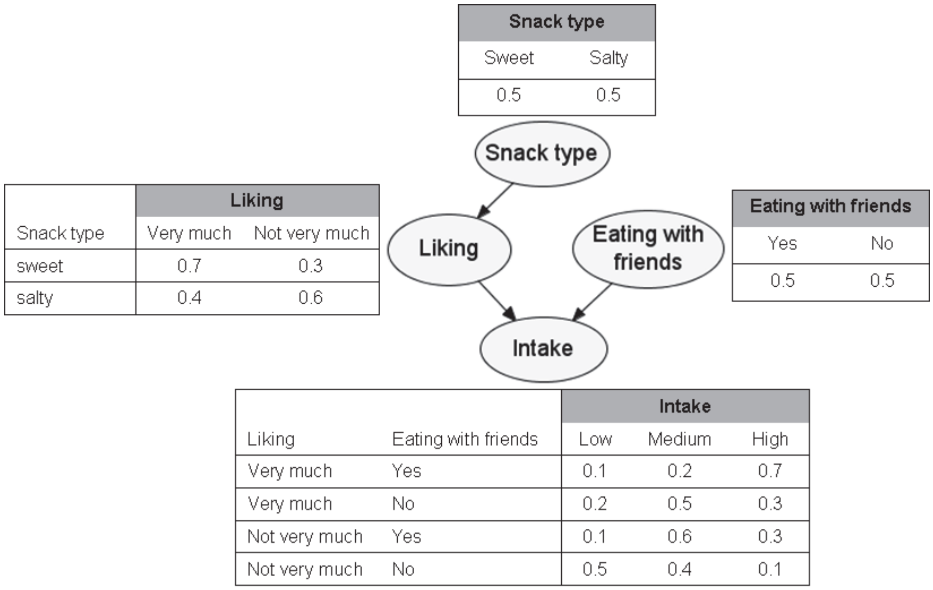
Hypothetical examples of Bayesian networks are used throughout the paper. Variables, probability values and dependent relations were suggested based on the knowledge and beliefs of the authors. Terminologies and concepts (formatted **bold**) concerning Bayesian networks are gently introduced while the paper focuses on the examples.

2.2 Concepts of Bayesian networks

Suppose we conducted a consumer test on snack consumption among teenagers (N = 200). There were four treatment conditions of two levels of snack types: sweet and salty, and two levels of eating environments: with friends and without friends. In each condition, teenagers first tasted snack samples and scored their liking on a continuous hedonic line scale ranging from 0 (not at all) to 100 (very much). They were then allowed to eat as much as they wanted. The total amount of snack consumed (intake) by each teenager was recorded. Data were generated by HUGIN software (HUGIN Researcher 7.2, http://www.hugin.com/), and a sample of 20 cases is shown in Appendix 2.A. We were interested in four variables: 'Snack type', 'Liking', 'Eating with friends' and 'Intake', and wanted to examine their relationships using the technique of Bayesian networks.

A Bayesian network has two aspects: **qualitative** and **quantitative** (Figure 2.1). The qualitative aspect is a **graph** formed by a set of labeled **nodes** (labeled ellipses, implying respective variables) linked to each other by a set of **arrows** (implying dependence relations among variables). Each node in the graph is associated with a table called Conditional Probability Table (**CPT**). The set of these CPTs represents the quantitative aspect of the model. They allow the quantification of relations among variables through **probability** expressions.

Definition 2.1: *Probability of an event A is the likelihood or chance that A will occur, denoted as P(A).*



Snack type	
Sweet	Salty
0.5	0.5

	Liking	
Snack type	Very much	Not very much
sweet	0.7	0.3
salty	0.4	0.6

Eating with friends	
Yes	No
0.5	0.5

		Intake		
Liking	Eating with friends	Low	Medium	High
Very much	Yes	0.1	0.2	0.7
Very much	No	0.2	0.5	0.3
Not very much	Yes	0.1	0.6	0.3
Not very much	No	0.5	0.4	0.1

Figure 2.1: A hypothetical Bayesian network of snack consumption. Labeled ellipses (nodes) represent respective variables of interest. Arrows indicate dependent relations between the two linked variables. The table associated to each node identifies different states that the variable can take, and the probability that the variable takes a specific state (given or not certain conditions). The probabilities associated with 'Snack type' and 'Eating with friends' were fixed by the experimental design. The probabilities associated with 'Liking' and 'Intake' resulted from the hypothetical data.

The arrow pointing from **parent** node to its **child** node suggests a possible cause-effect relationship. For instance, in Figure 2.1, the node 'Snack type' is a parent of 'Liking', i.e. the type of snack could influence liking scores. The node 'Intake' has two parents: 'Liking' and 'Eating with friends', i.e. snack consumption is supposedly affected by these two variables. These interactions (placement of the arrows) were suggested by the present authors.

In Bayesian networks, the graph is directed and acyclic. It means that the nodes must be connected by arrows, and there is no way from one node back to itself if following the arrows. This Directed Acyclic Graph (**DAG**) is considered as the **structure** of the Bayesian network model.

In snack consumption data, the values of two variables 'Liking' and 'Intake' are typically treated **continuous** because they can be given by any real number (between 0 and 100 for 'Liking' and any record for 'Intake'). In principle, Bayesian networks can handle both continuous and **discrete variables**. Many general-purpose algorithms, however, only deal with models containing discrete variables. Therefore, continuous data used in Bayesian networks are often discretized, i.e. creating a countable set of values.

Continuous variables can be converted into discrete variables by setting categories (referred to as **states**). In this case, two states of ‘Liking’ could be ‘Very much’, which was used to label liking scores greater or equal to 70; and ‘Not very much’ to label the rest (Appendix 2.A). The intervals and respective names of the states are generally suggested by domain experts, and preferably based on earlier empirical findings. The values of ‘Intake’ in our hypothetical network were also set into three states in the same manner: ‘Low’, ‘Medium’, and ‘High’. The data of ‘Snack type’ and ‘Eating with friends’ were categorical themselves (set by the experimental design). ‘Snack type’ had two states: ‘Sweet’ and ‘Salty’, and ‘Eating with friends’ had two states: ‘Yes’ and ‘No’. When one variable takes a specific state, its value is defined, and is treated as an **event**. For example, (‘Liking’ = ‘Very much’) and (‘Snack type’ = ‘Sweet’) are two events.

In a DAG, if a node has no parent, each value in its associated CPT represents the probability of the respective variable taking a specific state. For instance, the CPT of ‘Snack type’ says P(‘Snack type’ = ‘Sweet’) = 0.5 and P(‘Snack type’ = ‘Salty’) = 0.5; and that of ‘Eating with friends’ says P(‘Eating with friend’ = ‘Yes’) = 0.5 and P (‘Eating with friend’ = ‘No’) = 0.5. These probabilities reflect the randomization process of the experiment: ‘the chance of a teenager receiving a sweet or salty snack is equal, and his/her chance for eating snacks alone or with friends is also the same’. If a node has one or more parents, the associated CPT indicates the probability of the respective variable taking a specific state, given that the state of its parent variable(s) has been specified. For instance, having ‘Snack type’ as the unique parent, the CPT of the node ‘Liking’ is read as follows:

P(‘Liking’ = ‘Very much’ | ‘Snack type’ = ‘Sweet’) = 0.7, or in words: ‘given that a snack is sweet, the probability of this snack being liked very much is 0.7’

P(‘Liking’ = ‘Not very much’ | ‘Snack type’ = ‘Sweet’) = 0.3

P(‘Liking’ = ‘Very much’ | ‘Snack type’ = ‘Salty’) = 0.4

P(‘Liking’ = ‘Not very much’ | ‘Snack type’ = ‘Salty’) = 0.6

The probabilities above were obtained by counting the frequency of liking score values labeled as ‘Very much’ or ‘Not very much’ given by teenagers when (‘Snack type' = 'Sweet') and when ('Snack type' = 'Salty').

The node 'Intake' has two parents: 'Liking' and 'Eating with friends'. The probabilities in its CPT were determined as the frequency of intake values being labeled as 'Low', 'Medium' or 'High' for each of 3 x 2 state combinations of the two parent variables. For instance, we can say that teenagers consume a lot of snack if they like it very much and while eating with friends from the probabilities below:

P('Intake' = 'Low' | 'Liking' = 'Very much', 'Eating with friends' = 'Yes') = 0.1
P('Intake' = 'Medium' | 'Liking' = 'Very much', 'Eating with friends' = 'Yes') = 0.2
P('Intake' = 'High' | 'Liking' = 'Very much', 'Eating with friends' = 'Yes') = 0.7

The probabilities in the CPTs of 'Liking' and 'Intake' are called **conditional probabilities**, because they are conditioned to the state(s) of their parent(s). All values of the set of CPTs of a Bayesian network are recognized as **parameters** of the model.

Definition 2.2: *Conditional probability is the probability of an event A given that another event B has occurred, denoted P(A|B).*

2.3 Use of Bayesian networks

Suppose that we have obtained a Bayesian network comprising of its structure (a set of nodes linked to each other by a set of arrows, known as the qualitative aspect), and its parameters (a set of conditional probability tables CPTs, known as the quantitative aspect). What we can do then is to perform **inference**. The probabilistic inference is the computation of probabilities of interest given the model (Heckerman, 1995). For example, from the network of snack consumption (Figure 2.1), we wanted to compute the probability that teenagers eat a low (or medium, or high) amount of a snack, given that they are eating sweet snacks with friends. This computation is equivalent to predicting the snack consumption when certain information is available.

We used HUGIN software to illustrate the inference procedure within Bayesian networks. On the HUGIN interface, the probabilities are represented in percentage and visualized using horizontal bars.

2.3.1 Initial probability distribution

The initial probability distribution of the snack consumption network is presented in Figure 2.2. Compared with the network in Figure 2.1, the DAG stays the same; whereas, the **overall marginal probability** values are shown instead of conditional probability tables.

Definition 2.3: *Overall marginal probability is the probability of one variable taking a specific state while not knowing the values of all other variables in the network.*

The overall marginal probabilities of one variable are automatically calculated based on the CPT of that variable and the CPT(s) of its parent node(s). For instance, from the conditional probability values of the CPTs associated with 'Intake', 'Liking' and 'Eating with friends', we obtained the following overall marginal probabilities: P('Intake' = 'Low') = 0.21, P('Intake' = 'Medium') = 0.41, and P('Intake' = 'High') = 0.36. This set of overall marginal probabilities specifies the **overall marginal probability distribution** the variable 'Intake', denoted as P('Intake'). Similarly, P('Eating with friends') includes ('Yes' = 0.50; 'No' = 0.50), and P('Liking') includes ('Very much' = 0.55; 'Not very much' = 0.45).

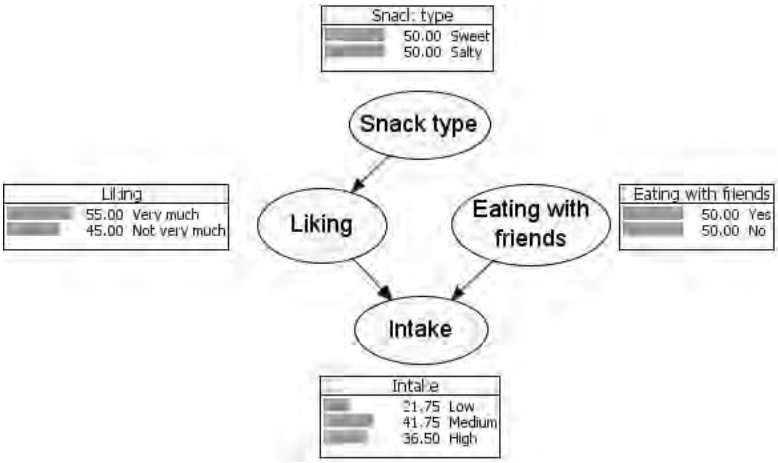


Figure 2.2: Initial probability distribution (HUGIN interface). Overall marginal probabilities of each variable are represented by horizontal bars and by percentages. These probability values were calculated by the software from the associated CPT of the variable, and the CPT(s) of its parent node(s).

2.3.2 Reasoning from cause to effect

We wanted to know how 'Snack type' influences 'Intake'. Once the initial probability distribution of the network was given (Figure 2.2), **evidences** should be set on the variable 'Snack type' to answer this question. An evidence could be the information observed or potential evidence about one hypothesis to be tested.

It was hypothesized that snacks are sweet, so P('Snack type' = 'Sweet') was set equal to 1.0 (Figure 2.3a). The software computes the probability distributions of other variables and all probabilities were conditioned by the event ('Snack type' = 'Sweet'). Overall marginal probability distribution for each variable was replaced by its **conditional marginal probability distribution**.

Definition 2.4: *Conditional marginal probability is the probability of one variable taking a specific state while knowing the value of at least one other variable in the network.*

When no information about the type of snack was given, the (overall) marginal probability P('Liking' = 'Very much') = 0.55, but when it was certain that the eaten snack is sweet, its (conditional) marginal probability P('Liking' = 'Very much' | 'Snack type' = 'Sweet') increased to 0.7. The marginal probability P('Intake' = 'High') = 0.36 also increased when being conditioned with ('Snack type' = 'Sweet'): P('Intake' = 'High' | 'Snack type' = 'Sweet') = 0.41. This shift in probability distribution gave us more "confidence" to say that teenagers would like a snack very much and consume more when they are given sweet snacks. In addition, the probability distribution of the variable 'Eating with friends' did not change under the evidence P('Snack type' = 'Sweet') = 1.0, which means 'Snack type' had no influence on the consumption environment. This observation is obvious because these two variables were independently manipulated in the experimental design. When it was certain that the eaten snack is salty (P('Snack type' = 'Salty') = 1.0), the resulting probability distributions are presented in Figure 2.3b.

Getting back to the question how 'Snack type' influences 'Intake', it is enough to compare the probability distribution of the variable 'Intake' when the evidence was P('Snack type' = 'Sweet') = 1.0 (Figure 2.3a) and when P('Snack type' = 'Salty') = 1.0 (Figure 2.3b). The distribution of 'Intake' had more "weight" on 'High' state and less "weight" on 'Low' state when P('Snack type' = 'Sweet') = 1.0 than when the other evidence was set. We can conclude from this hypothetical network that teenagers are more likely to have a higher intake when the snacks are sweet rather than salty. In the same manner, the influence of 'Liking' and 'Eating with friends' on 'Intake' could be tested by setting new evidences on these variables.

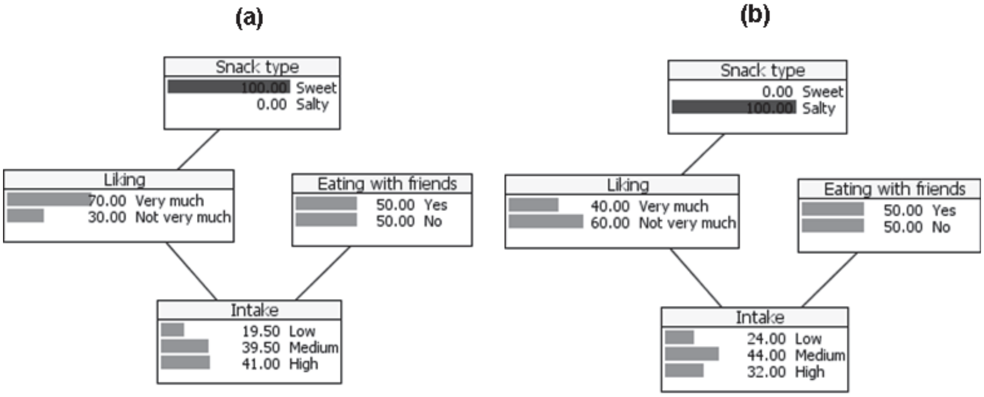


Figure 2.3: Inference: influence of snack type. When it is certain that 'Snack type' = 'Sweet' (a), it is very likely that teenagers like it 'Very much', and that the amount of snack consumed is higher than when 'Snack type' = 'Salty' (b). This results from the greater probability of 'Intake' taking 'High' state and the lower probability of 'Intake' taking 'Low' state when the evidence is 'Snack type' = 'Sweet'. The type of snack has no influence on whether or not teenagers are having snacks with their friends.

2.3.3 Combined influence of variables

Bayesian network models allow a clear visualization of the combined effect of two variables. Figure 2.4 shows the probability distributions of 'Liking' and 'Intake' when evidences were set for 'Snack type' and 'Eating with friends'.

Let us consider Figure 2.4a (eating sweet snacks) as the baseline of Figure 2.4b (eating sweet snacks with friends) and Figure 2.4c (eating sweet snacks without friends). Adding the information of the eating environment (with friends or alone) either increased (Figure 2.4b) or decreased (Figure 2.4c) the probability of consuming a high amount of sweet snacks. The same trend was observed when comparing the probability distribution of 'Intake' given three input evidences: i) eating salty snacks, ii) eating salty snacks with friends, and iii) eating salty snacks alone (illustration not shown). Hence, the combined effect of 'Snack type' and 'Eating with friends' was present: 'Eating with friends' enhanced the influence of 'Snack type' on 'Intake'.

In this section, we wanted to predict the snack consumption when it is known that teenagers are eating sweet snacks with friends. The answer is indeed the probability distribution of the node 'Intake' when it was set that P('Snack type' = 'Sweet') = 1.0 and P('Eating with friends' = 'Yes') = 1.0 (Figure 2.4b).

In the same manner, we can set evidence for more variables. For example, we can predict the intake when teenagers are eating salty snacks with friends, with the added knowledge that they all like salty snacks very much.

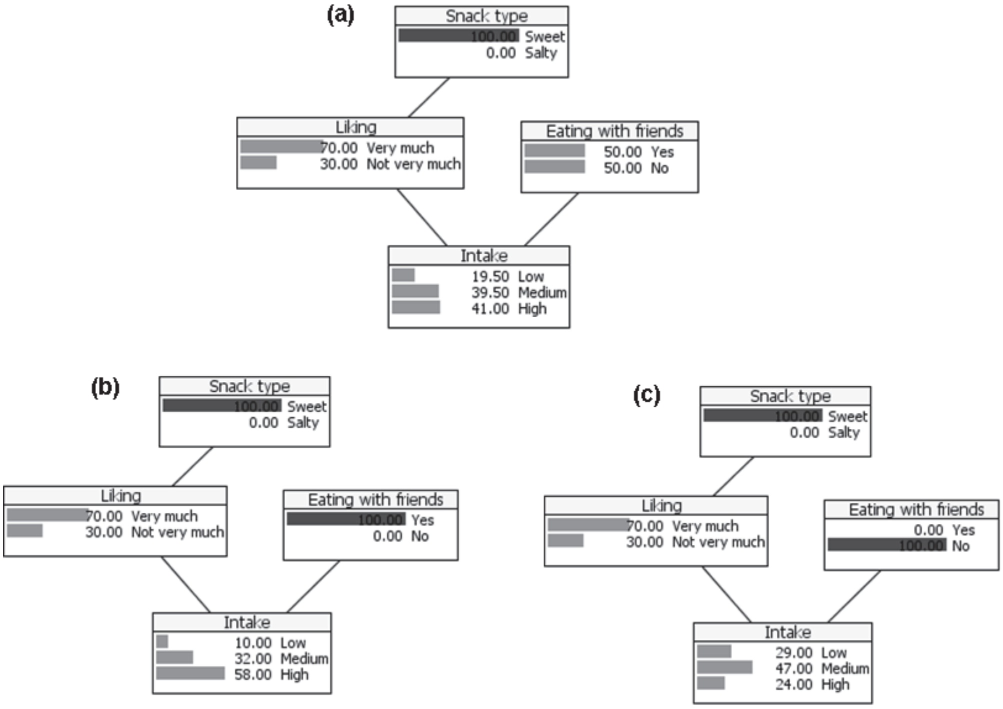


Figure 2.4: Inference: combined evidences. When it is known that teenagers are eating sweet snacks (a), the probability of consuming a high intake increases if they are eating with friends (b), as compared to when they are eating alone (c).

2.3.4 Reasoning from effect to cause

The inferences performed in Figure 2.3 and Figure 2.4 are forward reasoning, i.e. from cause to effect. Bayesian network models also allow backward reasoning, i.e. from effect to cause.

Suppose the only information we know was the amount of snacks consumed. If this amount was low (Figure 2.5a), the eaten snacks were more likely to be salty than sweet (P = 0.55 vs. P = 0.45) and it were very likely (P = 0.77) that teenagers ate snacks alone. The opposite trends were found when the intake was high (Figure 2.5b).

In short, predictions can be made with Bayesian networks through the inference procedure. The backward reasoning is a particular strength of these models.

It could be useful in product design. For instance, a model relating input attributes to output attributes can deduce the most likely states of how input attributes should be in order to obtain the desired output attributes.

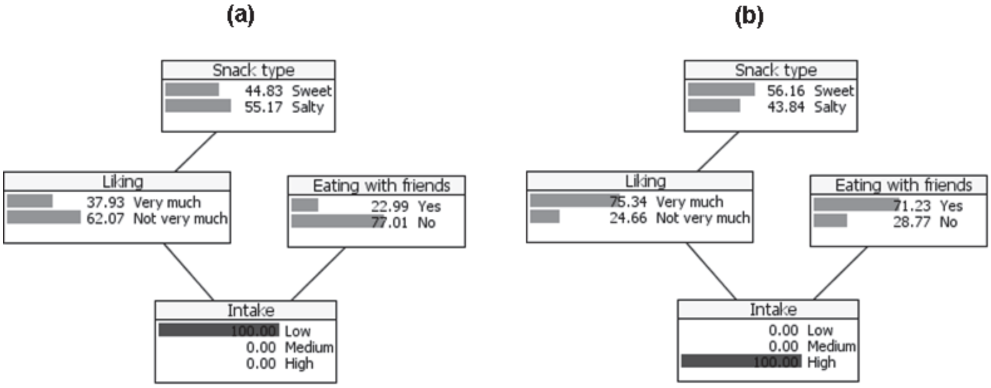
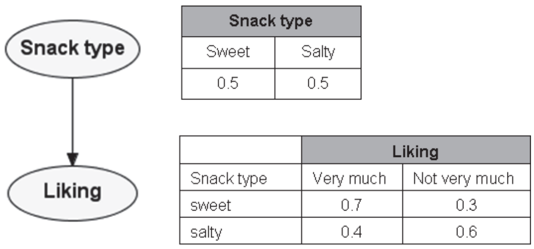


Figure 2.5: Inference: backward reasoning. If the 'Intake' is known to be 'Low', the type of snack is deduced to be more likely salty than sweet and more likely to be eaten in the absence of friends (a). If the 'Intake' is known to be 'High', it is very likely that teenagers consumed sweet snacks together with friends (b).

2.4 Inference in simple models

This section explains how the probabilities are calculated in a simple network model. Suppose we work on a network relating 'Liking' to 'Snack type' (Figure 2.6). This network was extracted from the hypothetical network on snack consumption among teenagers (Figure 2.1). We wanted to know how likely a snack is to be 'Sweet' (or 'Salty') if it was observed that teenagers like the given snack 'Very much'. Thus, it was needed to compute two conditional probabilities: P('Snack type' = 'Sweet' | 'Liking' = 'Very much') and P('Snack type' = 'Salty' | 'Liking' = 'Very much').

The variable 'Snack type' had two states: 'Sweet' and 'Salty'. Its overall marginal probability distribution P('Snack type') was quantified to be (0.5, 0.5) by the experimental design (Figure 2.6). We did not know yet the overall marginal probability distribution of the variable 'Liking'. However, the relationship between 'Snack type' and 'Liking' was quantified through the **conditional probability distribution** P ('Liking' | 'Snack type') = (0.7, 0.3, 0.4, 0.6) (Figure 2.6).



Snack type	
Sweet	Salty
0.5	0.5

	Liking	
Snack type	Very much	Not very much
sweet	0.7	0.3
salty	0.4	0.6

Figure 2.6: A Bayesian network model relating snack type to liking. The overall marginal probability distribution of 'Snack type' ('Sweet' or 'Salty') resulted directly from the experimental design. The conditional probabilities associated with 'Liking' ('Very much' or 'Not very much') were the frequency of the liking scores being either 'Very much' or 'Not very much' for each type of snack.

In order to compute the overall marginal probabilities of 'Liking', we need to know the **joint probability distribution** of the given Bayesian network.

Definition 2.5: *Joint probability of two events and is the probability that both events occur together, denoted as P (A,B).*

For instance, the joint probability of two events ('Snack type' = 'Sweet') and ('Liking' = 'Very much') is P('Snack type' = 'Sweet', 'Liking' = 'Very much'), which represents the probability that one snack is found to be both sweet and liked very much.

Definition 2.6: *Joint probability distribution of two discrete variables X and Y , denoted as P(X,Y) , is the set of joint probabilities P(X=x, Y=y) , where and are any state of X and Y , respectively.*

For instance, the joint probability distribution of two variables 'Snack type' and 'Liking' P('Snack type', 'Liking') consists of four following joint probabilities:

P('Snack type' = 'Sweet', 'Liking' = 'Very much'),
P('Snack type' = 'Sweet', 'Liking' = 'Not very much'),
P('Snack type' = 'Salty', 'Liking' = 'Very much'),
P('Snack type' = 'Salty', 'Liking' = 'Not very much').

2.4.1 Calculation of joint probabilities and overall marginal probabilities

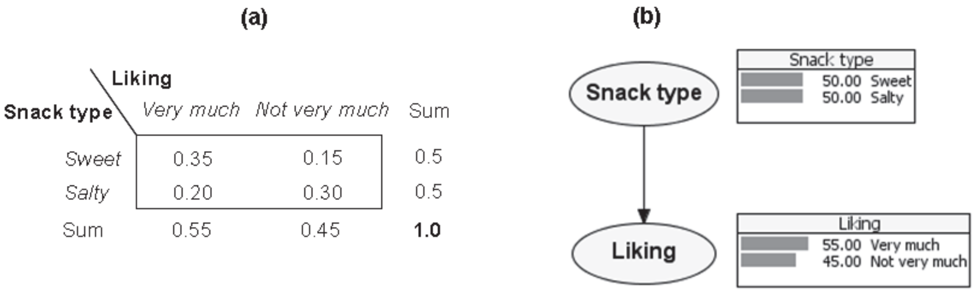
The **fundamental rules** of probability allow the calculation of the joint probability from the marginal probability and conditional probability:

$$P(A,B) = P(A|B) * P(B) = P(B|A) * P(A) \qquad \textbf{Equation 2.1}$$

Applying directly Equation 2.1, the joint probability distribution of P('Snack type', 'Liking') could be obtained from P('Snack type') and P('Liking' | 'Snack type'), for example:

P('Snack type' = 'Sweet', 'Liking' = 'Very much') = P('Liking' = 'Very much' | 'Snack type' = 'Sweet') * P('Snack type' = 'Sweet') = 0.7 * 0.5 = 0.35

Four joint probabilities of the distribution P('Snack type', 'Liking') are shown in the joint probability table in Figure 2.7a.


(a)

Snack type \ Liking	Very much	Not very much	Sum
Sweet	0.35	0.15	0.5
Salty	0.20	0.30	0.5
Sum	0.55	0.45	**1.0**

(b)

Figure 2.7: Calculation of marginal probability distributions. This example was done on the model in Figure 2.6. The marginal probabilities of ' Liking' were obtained by summing up all rows of the **joint probability table** (a) and those of 'Snack type' found by summing up all its columns. The same results calculated by HUGIN software are shown in (b).

The **law of total probability** says, for any event *A* , that if there is a set of *n* mutually exclusive and exhaustive events $E_i (i=1,\dots,n)$ [1], then:

$$P(A) = \sum_{i=1}^{n} P(A,E_i) \qquad \textbf{Equation 2.2}$$

This enables us to calculate the marginal probability distribution P('Liking') based on the two mutual exclusive and exhaustive events of 'Snack type':

P('Liking' = 'Very much') = P('Liking' = 'Very much', 'Snack type' = 'Sweet') + P

[1] *n* events $E_1, E_2,\dots,E_n$ are said to be mutually exclusive and exhaustive if no two of them do occur at the same time ($E_i \cap E_j \neq 0$ with *i,j* ε *n* and *i* ≠ *j*) and their individual probabilities sum up to 1 $\sum_{i=1}^{n} P(E_i)=1$

('Liking' = 'Very much', 'Snack type' = 'Salty') = 0.35 + 0.20 = 0.55

P('Liking' = 'Not very much') = P('Liking' = 'Not very much', 'Snack type' = 'Sweet') + P('Liking' = 'Not very much', 'Snack type' = 'Salty') = 0.15 + 0.30 = 0.45

The rule of this calculation is to sum up all rows of the joint probability table (Figure 2.7a). If summing up all columns, the marginal probability distribution P('Snack type') is again found. The same results given by HUGIN software are shown in Figure 2.7b.

2.4.2 Calculation of conditional probabilities of interest

At this stage, our conditional (marginal) probabilities of interest could be computed using the derived form of Equation 2.1:

P('Snack type' = 'Sweet' | 'Liking' = 'Very much') = P('Snack type' = 'Sweet', 'Liking' = 'Very much') / P('Liking' = 'Very much') = 0.35 / 0.55 = 0.6364

P('Snack type' = 'Salty' | 'Liking' = 'Very much') = P('Snack type' = 'Salty', 'Liking' = 'Very much') / P('Liking' = 'Very much') = 0.20 / 0.55 = 0.3636

Note that the probability P('Snack type' = 'Salty' | 'Liking' = 'Very much') can also be derived from P('Snack type' = 'Sweet' | 'Liking' = 'Very much') because all marginal probabilities of one variable sum up to 1.

These outcomes were also given automatically by HUGIN software when setting evidence ('Liking' = 'Very much') (Figure 2.8a). Similar steps allowed us to obtain the probability distribution of 'Snack type' when the evidence 'Liking' = 'Not very much' was set (Figure 2.8b). In short, the joint distribution of a Bayesian network is the key to do inference.

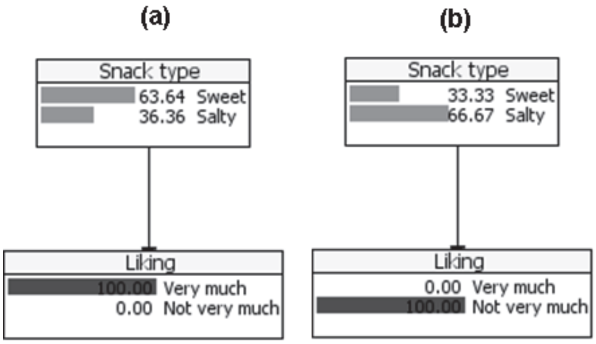


Figure 2.8: Inference in the model relating snack type to liking. The distribution of 'Snack type' given evidence on the variable 'Liking' was found by calculating the joint probability P ('Snack type', 'Liking').

2.5 Inference in complex models

So far, we have considered only the inference procedure in the network containing two variables ('Snack type' and 'Liking') and each variable had only two states. The joint probability distribution of this network consisted of four joint probability values and only three of those needed to be specified (the last one is dependent on the rest). In real world problems, however, we are typically interested in looking for relationships among a large number of variables (Heckerman, 1995).

Consider, for example, a network connecting *n* variables $(X_1,X_2,\ldots,X_n)$.Assuming that each variable of this network takes only two states, its joint probability distribution $P(X_1,X_2,\ldots,X_n)$ is specified by $(2^n - 1)$[1] joint probability values. This exponential relationship results in an enormous number when is large. If the variables have more than two states, this number grows even more rapidly. To simplify the calculation of the joint probabilities, assumptions on probabilistic relations are used in Bayesian networks, such as dependence and conditional independence.

2.5.1 Problem example

We used again the network on snack consumption, except that the variable 'Purchase intention' was included, denoted as "Extended snack consumption network" (Figure 2.9). When teenagers tasted and gave liking scores for snack samples, they also stated whether or not they have the intention to purchase the product. Values of 'Purchase intention', given as either 'Yes' or 'No', were assumed to be influenced only by the variable 'Liking'.

The structure of Bayesian networks can be read by three typical connections linking a group of three nodes. These typical connections are **serial** ($X \rightarrow Y \rightarrow Z$), ($X \leftarrow Y \leftarrow Z$), **diverging** ($X \leftarrow Y \rightarrow Z$), and **converging** ($X \rightarrow Y \leftarrow Z$). In the extended snack consumption network (Figure 2.9), for instance, ('Snack type' → 'Liking' → 'Intake') and ('Snack type' → 'Liking' → 'Purchase intention') are two serial connections, ('Purchase intention' ← 'Liking' → 'Intake') is a diverging connection, and ('Liking' → 'Intake' ← 'Eating with friends') is a converging connection. These kinds of connections will be referred to while examining network probabilistic relations in this chapter.

[1] The number of joint probabilities of the network is 2^n. However, as all these probabilities have to sum up to 1, the last one is dependent on the other values, which results in the number $(2^n – 1)$.

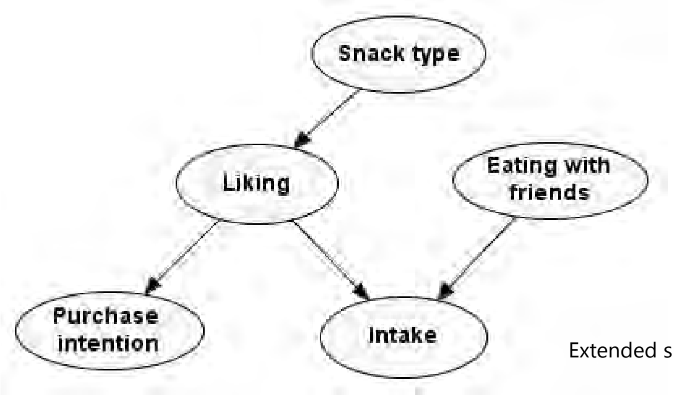


Figure 2.9:
Extended snack consumption network.

To perform inference on this network, the joint probability distribution over the network needs to be specified, i.e. P('Intake', 'Purchase intention', 'Liking', 'Snack type', 'Eating with friends'), or abbreviated as P('Int', 'Pur', 'Lik', 'Sna', 'Eat').

Applying the fundamental rules of probability of Equation 2.1, the joint probability distribution of the network of variables can be decomposed into the product of conditional and marginal probability distributions:

$$\begin{aligned} &P(X_1,X_2,\ldots,X_n) = P(X_1 \,|\, X_2,\ldots,X_n) * P(X_2,\ldots,X_n) \\ &= P(X_1 \,|\, X_2,\ldots,X_n) * P(X_2,X_3,\ldots,X_n) * P(X_3,X_4,\ldots,X_n) \\ &= P(X_1 \,|\, X_2,\ldots,X_n) * P(X_2 \,|\, X_3,\ldots,X_n) * .. * P(X_{n\text{-}1} \,|\, X_n) * P(X_n) \end{aligned}$$ **Equation 2.3**

This allows us to rewrite the joint probability distribution of the extended snack consumption network as follows:

P('Int', 'Pur', 'Lik', 'Sna', 'Eat') = P('Int' | 'Pur', 'Lik', 'Sna', 'Eat') * P('Pur' | 'Lik', 'Sna', 'Eat') * P('Lik' | 'Sna', 'Eat') * P('Sna' | 'Eat') * P('Eat')

The joint probability distribution can be thus calculated through the conditional probability distributions. These conditional probability distributions can be simplified when specific assumptions about probabilistic relations among the five variables are defined: assumptions about their dependencies.

2.5.2 Independence and Conditional dependence

Let us consider the variable 'Eating with friends'. It is linked directly to 'Intake', indirectly to 'Liking' through a converging connection, indirectly to 'Snack type' and 'Purchase intention' through one converging connection and one serial connection (Figure 2.9).

On the one hand, when no information in the network was given, changing the marginal probability distribution of 'Eating with friends' did not affect those of 'Snack type', 'Liking' and 'Purchase intention' (Figure 2.10a,b). In turn, different evidences on these three variables did not lead to any modification in values of 'Eating with friends' (illustrations not shown). It is said that information cannot be transmitted through a converging connection.

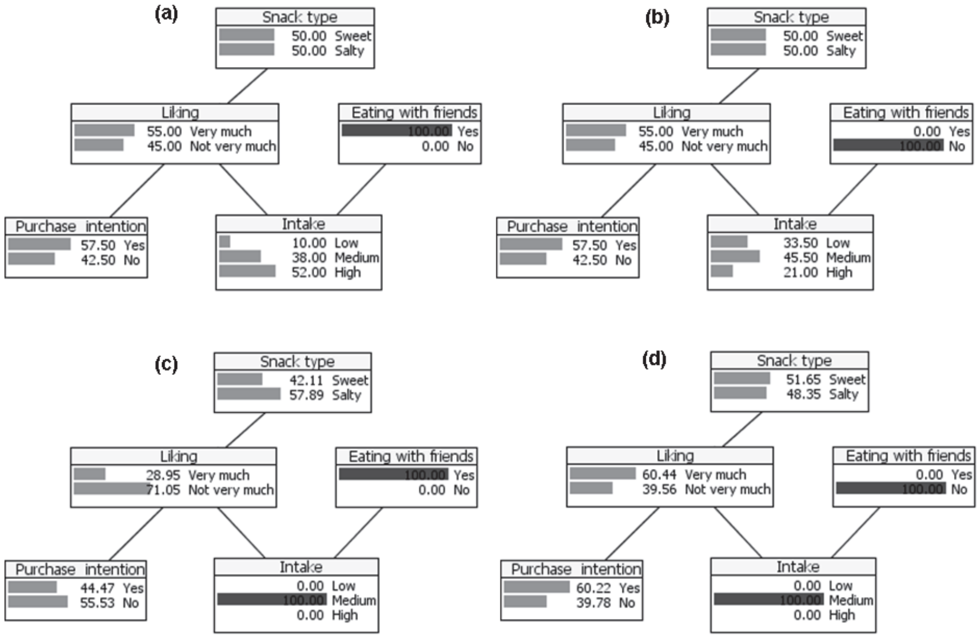


Figure 2.10: Independence and conditional dependence. The variable 'Eating with friends' is independent of 'Snack type', 'Liking' and 'Purchase intention' because modifying values of 'Eating with friends' (a, b) does not lead to any changes on the marginal probability distributions of the other three variables. However, when prior information on 'Intake' is provided (for example, 'Intake' = 'Medium'), modifications of 'Eating with friends' (c, d) affect marginal probability distributions 'Snack type', 'Liking' and 'Purchase intention'. Thus, 'Eating with friends' becomes conditional dependent to 'Snack type', 'Liking' and 'Purchase intention' given values of Intake.

In probability theory, two events (or variables) are said to be **independent** if the probability (distribution) of one event (or variable) does not change whether or not provided with information about the other:

Definition 2.7: *Two events A and B (P(A) ≠ 0 and P(B) ≠ 0) are independent if P(A|B) = P(A).*

Definition 2.8: *Two discrete variables X and Y are independent if P(X=x | Y=y) = P(X=x) for any state x, y of X and Y, respectively; or simply expressed by probability distribution if P(X|Y) = P(X).*

From the definitions of probabilistic independence, it can be interpreted that 'Eating with friends' is independent of the three variables 'Snack type', 'Liking' and 'Purchase intention'. Consequently, P('Sna' | 'Eat') = P('Sna'); P('Lik' | 'Eat') = P('Lik'); P('Pur' | 'Eat') = P('Pur').

On the other hand, when knowing the value of the middle node of the converging connection ('Liking' → 'Intake' ← 'Eating with friends'), changing the marginal probability distribution of 'Eating with friends' appears to affect those of 'Purchase intention', 'Liking' and 'Snack type' (Figure 2.10c,d). In this situation, 'Eating with friends' became **conditional dependent** to 'Purchase intention', 'Liking' and 'Snack type' (given values of 'Intake'). Thus, it is said that information can be transmitted through a converging connection only if information about the middle node is provided.

2.5.3 Dependence and conditional independence

Consider now the variable 'Purchase intention'. It is linked directly to 'Liking', indirectly to 'Snack type' through a serial connection, indirectly to 'Intake' through a diverging connection, and indirectly to 'Eating with friends' through one diverging connection and one converging connection.

When no information in the network was given, changing the marginal probability distribution of 'Purchase intention' affected those of 'Snack type', 'Liking', 'Intake' (Figure 2.11a,b). However, when evidence was set for 'Liking', e.g. 'Liking' = 'Very much', added information on 'Purchase intention' had no more effect on (conditional) marginal probability distributions of 'Intake' and 'Snack type' (Figure

2.11c,d). Similarly, given ‘Liking’ = ‘Very much’, added information on ‘Intake’ (or ‘Snack type’) did not influence neither the probability distributions of the other two nodes (illustrations not shown). It is thus said that information can be transmitted through serial and diverging connections. This flow of information, however, can be blocked by providing evidence on the middle node of these two connections.

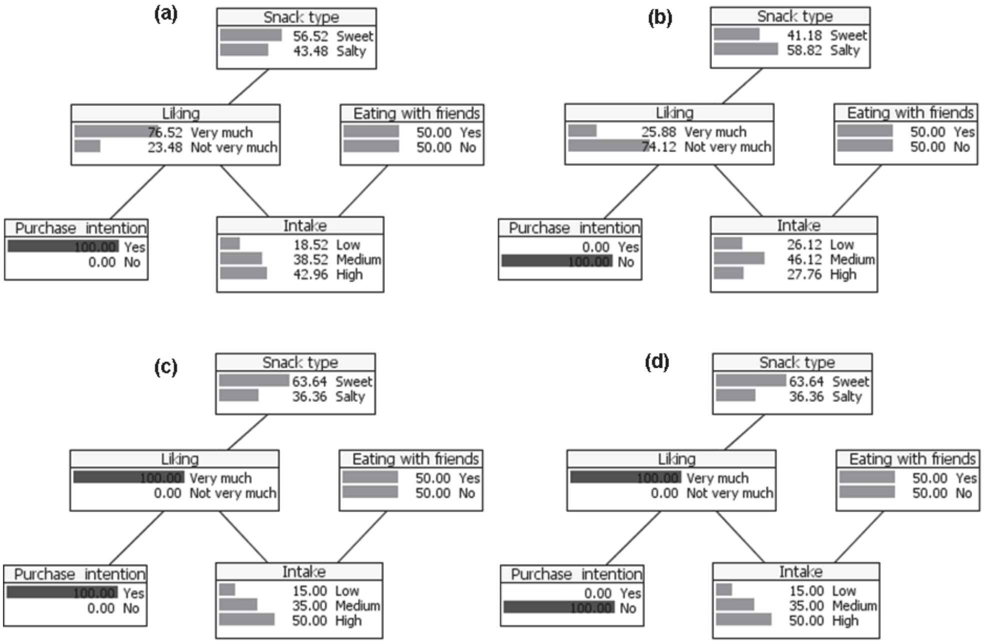


Figure 2.11: Dependence and conditional independence. Information can be transmitted from ‘Purchase intention’ to ‘Snack type’ through a serial connection (X→Y→Z) and to ‘Intake’ through a diverging connection (X←Y→Z) (a,b). However, this flow of information is blocked when evidence is set for ‘Liking’, the middle node in serial and diverging connections (c,d).

Briefly, although three variables ‘Snack type’, ‘Purchase intention’ and ‘Intake’ do not link directly to each other, they are not independent. New information about one variable can lead to changes in values of the other two variables through the updated information on the middle variable ‘Liking’. However, when the value of ‘Liking’ is known, new information about one of the three variables ‘Snack type’, ‘Purchase intention’ and ‘Intake’ does not change the values of the other two. This observation is an example of the concept of conditional independence in probability theory:

Definition 2.9: *Two events A and B are conditionally independent given event C if P(C) ≠ 0 and P(A|B,C) = P(A|C).*

Definition 2.10: *Two discrete variables X and Y are conditionally independent given another random variable Z if P(X=x |Y=y, Z=z) = P(X=x |Z=z) for any state x, y, z of X, Y and Z, respectively; or simply expressed by probability distribution P(X|Y,Z) = P(X|Z) or P(X|Y,Z) = P(Y|Z).*

According to the definition of conditional independence, three variables 'Snack type', 'Purchase intention' and 'Intake' are conditional independent to each other given 'Liking'. We can thus simplify some conditional probabilities, such as P('Int' | 'Pur', 'Lik') = P('Int' | 'Lik'), P('Int' | 'Sna', 'Lik') = P('Int' | 'Lik').

To summarize, two dependent variables X and Y can become conditionally independent if there is a third variable Z forming a serial connection (X→Z→Y or X← Z←Y) or a diverging connection (X←Z→Y). Two independent variables X and Y can become conditionally dependent if there is a third variable Z forming a converging connection (X→Z←Y).

2.5.4 Joint probability distribution in Bayesian networks

Having defined probabilistic relations in the network, let us come back to the calculation of the joint probability distribution as proposed in Section 2.5.1:

P('Int', 'Pur', 'Lik', 'Sna', 'Eat') = P('Int' | 'Pur', 'Lik', 'Sna', 'Eat') * P('Pur' | 'Lik', 'Sna', 'Eat') * P('Lik' | 'Sna', 'Eat') * P('Sna' | 'Eat') * P('Eat')

Given that 'Intake', 'Snack type' and 'Purchase intention' are conditional independent given 'Liking', and 'Eating with friends' is independent to 'Snack type', 'Purchase intention' and 'Liking', the following relationships were established:

P('Int' | 'Pur', 'Lik', 'Sna', 'Eat') = P('Int' | 'Lik', 'Eat')
P('Pur' | 'Lik', 'Sna', 'Eat') = P('Pur' | 'Lik', 'Eat') = P('Pur' | 'Lik')
P('Lik' | 'Sna', 'Eat') = P('Lik' | 'Sna')
P('Sna' | 'Eat') = P('Sna')

resulting in:

P('Int', 'Pur', 'Lik', 'Sna', 'Eat') = P('Int' | 'Lik', 'Eat') * P('Pur' | 'Lik')* P('Lik' | 'Sna') * P('Sna')* P('Eat')

The joint probability distribution P('Int', 'Pur', 'Lik', 'Sna', 'Eat') can therefore be calculated from the product of conditional probability distributions of each variable given its parent(s) and marginal probability distributions of variables that have no parents. To generalize, the joint probability distribution of the network having variables in Equation 2.3 can be computed as the product of conditional probability distributions of each node given its parent(s):

$$P(X_1,X_2,\ldots,X_n) = \prod_{i=1}^{n} P\,(\,X_i\,|\,parents(X_i))$$ **Equation 2.4**

If the node has no parent, its conditional probability distribution is actually its marginal probability distribution.

In short, identifying independence and conditional independence relations among the set of variables of interest is essential to compute the joint probability distribution, which in turn enables us to perform inference on the network.

In this section, the inferences in the network were performed to illustrate the probabilistic relations among the variables. In practice, however, if the structure is defined by domain experts, it also implies probabilistic relations through the identification of serial, diverging, and converging connections. If the structure is not known yet, these probabilistic relations could be examined based on the data, and the structure is then built from these relations. This learning process will be briefly discussed in the next section.

2.6 Learning Bayesian Networks

2.6.1 Definition of Bayesian networks

Most papers on Bayesian networks begin with stating the definition of a Bayesian network model, which is difficult to relate to real world problems in food science. We hope that after having introduced basic terminologies and concepts, the definition below can now be more easily connected to the content:

Definition 2.11: *A Bayesian network is a graphical model for probabilistic relationships over a set of variables. It consists of a qualitative aspect, encoding (conditional) dependence and independence among variables; and a quantitative aspect, encoding* *the joint probability distribution over these variables*.

2.6.2 Learning Bayesian networks

To construct (or to learn) a Bayesian network model, we need to specify its **structure** (a set of nodes linked by a set of arrows or a DAG) and its **parameters** (all conditional probabilities forming the Conditional Probability Table for each node). The input that can be used to learn a Bayesian network are the so-called **domain knowledge** and **empirical data** (new observations). The domain knowledge can be the common knowledge of the domain (collected from published scientific papers) or the beliefs of domain experts (hypotheses). The empirical data involved may be complete or incomplete (containing missing values).

The network structure can be elicited from domain knowledge, as in the cases where domain experts are able to specify relevant variables and interactions among them (Corney, 2000). The structure would be then considered as known. Theoretically and practically, domain knowledge also allows the specification of the network parameters (probability values) as in the case of expert systems (Heckerman et al., 1995). These probabilities are to be otherwise estimated from the data.

In some cases, the network structure is not known or incomplete. Empirical data is therefore the only input for inducing structure and estimating parameters. Learning the structure of a Bayesian network from data is a challenge pursued within the machine-learning domain. The task is even harder with incomplete data. The underlying computational issues, mathematical challenges, as well as the general problems related to such a board inductive learning task go beyond the scope of this small introduction. The interested reader is pointed to a current comprehensive review on that subject in Daly et al. (2011).

2.6.3 Known structure, complete data

Let us consider the network on snack consumption among teenagers (Figure 2.1). Conclusions from various studies (domain knowledge) were used to define the structure of this network: i) flavor of a food product is an important factor determining the liking for it, (ii) the more we like a product, the more we eat it, iii) the social interaction during a meal also influences the amount of food we eat.

The data of snack consumption study was assumed to be complete (a sample of the dataset is shown in Appendix 2.A). The conditional probabilities of the CPT for each variable were simply the frequency of its specified state given the state of its parents. Having obtained the complete structure and all the parameters, inference can be performed.

2.6.4 Known structure, incomplete data

In reality, data is often not complete, due to some variables not being observed for all cases. The frequency cannot be accessed in such cases. To solve this problem, the missing data could be assigned to certain expected values based on available data using EM- (Expectation-Maximization) algorithm (Lauritzen, 1995). This algorithm uses an iterative method to maximize the probability of the observed data given the (estimated) parameters of the network.

2.7 Discussions

Bayesian networks, as well as other machine learning techniques, are rather complementary than contradictory to classical statistical approaches in analyzing data (Cunningham, 1995). At the present time, not many applications of Bayesian networks in food area have been published. In this section, we discuss the general advantages and disadvantages of this approach in view of using food data, as well as the potential applications of Bayesian networks in food areas.

2.7.1 End-user friendly communicator

Bayesian networks provide a good visual communication tool of mathematical relations to end-users through graphical representation. They can give fast responses to queries (inferences) once the model is completed.

2.7.2 Handle complex problems

Assumptions on probabilistic dependence and independence allow scientists to model complex problems using Bayesian networks. In a large network, it would be enough to examine relations of each variable with its parent variables. Reasoning on learned causal relationships can then be done to predict behavior of the whole system.

First, we can estimate and visualize how the "cause" influences its "effect" (forward reasoning). This feature serves to explain as well as to explore information from our system. Second, backward reasoning reveals how to manipulate the "causes" to obtain certain desired values of its "effect". This feature of Bayesian network is valuable in designing food products driven by any desired characteristics or consumer demands.

2.7.3 Use of prior knowledge

Learning the structure of a network is the most difficult task, especially from small datasets. Fortunately, Bayesian networks enable us to combine domain knowledge with empirical data. In food-related problems, existing knowledge could provide information to define (at least partly) dependence relations between variables of interest.

2.7.4 Handle incomplete datasets

Gathering food data, particularly concerning human responses, is very expensive and time-consuming. Thus, typical features of food datasets are small and often incomplete (Corney, 2000). The EM-algorithm, which is one among several possible solutions, allows the approximation of the missing observations of one variable through the state of other variables (Heckerman, 1995).

Generally, larger data sets yield more reliable estimations of probabilities. However, there is no such criterion describing "enough data" to perform the analysis. The performance of the networks is best validated when testing with new data.

2.7.5 Discretization of continuous variables

While food data often have continuous values, Bayesian network software can deal with continuous variables in only a limited manner. Hence, it is necessary to convert continuous variables into discrete variables. This is a disadvantage of Bayesian networks due to a huge information loss, especially in linear relationships (Myllymäki et al., 2002). Furthermore, finding "the appropriate" way to discretize data is another issue. The number of intervals and the division points can lead to different results (Myllymäki et al., 2002). On the one hand, the bigger the number of intervals, the better the real relationships of variables can be captured. On the other hand, the increase of this number requires larger amounts of data to estimate all the probabilities. Generally, domain experts perform this step based on specific goals of the modeling or on other relevant information.

It should be noted that, however, research on Bayesian networks is evolving very fast and promises more flexible uses of continuous data.

2.7.6 Potential applications of Bayesian networks in the food area

In the food area, most published models are related to chemical kinetics and microbial growth and they are based on deterministic approaches. The application of Bayesian networks in modeling is at the early stages, and mostly concerns microbial risk assessment. van Boekel (2004) has discussed Bayesian solutions with respect to the inherent variability and uncertainty in food-science problems, from food quality–safety management to food design aspects.

Food quality and safety management often involves a large number of variables, and these variables are not always observed or measured due to economic or technological constraints. Bayesian networks are suitable to handle these problems, and could be applied in building models to control different dimensions of quality, as well as to detect potential risk factors along the food chain.

Food design is driven by consumer preference, which can be generally accessed by sensory attributes of a product. Conventional flavor and texture are widely accepted and constitute the so-called "balance" of a food. Recent efforts of the food industry, however, are to remove a large portion of saturated and *trans* fats, and to reduce the amount of salt and sugar from food products without losing the balance in flavor and texture. These efforts interfere not only physical and chemical interactions of different ingredients at the food level, but also multi-modal perceptions at the brain level. We can practically handle interactions at the food level. Huge uncertainty due to the lack of knowledge at the brain level, however, does not allow us to control the perception integration. Therefore, deterministic food design limits itself within various isolated contexts. Bayesian networks might be valuable in product design. First, this technique is capable to deal with uncertainty. Second, it provides a possibility to combine different related studies, which enable us to consider a complex problem as a whole.

Particularly, consumer and marketing research is giving more and more attention to Bayesian networks beside Structural Equation Modeling as a conventional technique (Blodgett and Anderson, 2000, Gupta and Kim, 2007, Repères research). These two techniques have been shown to complement each other (Gupta and Kim, 2007). The number of observations in consumer and marketing research is rather large, which enables the parameter learning and possibly structure learning in Bayesian networks. From this point of view, sensory studies may encounter challenges when using Bayesian networks due to the limited sample size. However, the possibility to use domain knowledge may be of help in these cases. More modeling work with sensory data is expected in future to examine the potential application of Bayesian networks in this field.

2.8 Sources for further reading

“Learning Bayesian Networks” written by Neapolitan (2003) is highly recommended to readers who want to get an in-depth understanding on Bayesian networks. Besides, Heckerman (1995) wrote “A Tutorial on Learning with Bayesian networks” which highlighted well main features and discussed technical problems.

For readers whose interest lies in applications, a short and gentle introduction “Bayesian networks without Tears” given by Charniak (1991), or a more detail introduction written by Murphy (1998) are advisable. Technical approaches are described in detail in ”Bayesian Networks and Influence Diagrams: A Guide to Construction and Analysis” (Kjaerulff and Madsen, 2008). Particularly, the textbook “Bayesian networks: A Practical Guide to Applications” (Pourret, Naïm and Marcot, 2008) brings in many applications in various fields.

There are a considerable number of software packages available in open source or commercially to build Bayesian networks. They were listed and given a detailed description in (Korb and Nicholson, 2004). Here are some examples:

HUGIN (http://www.hugin.com/, Hugin Expert A/S) is a commercial product that supports an easy use by click-and-point procedures. HUGIN can learn structure and parameters from discrete data, and also support inference on Bayesian networks having continuous variables. HUGIN version 7.2, however, cannot learn parameters from continuous data. Besides, decision and utility nodes can be added to Bayesian networks, resulting in the so-called “Influence diagrams”, to support the decision-making process.

Netica (http://www.norsys.com, Norsys Software Corp.) is also a widely used commercial software that supports Bayesian networks and Influence diagrams. Netica can learn only parameters and work only with discrete nodes.

BayesiaLab (http://www.bayesia.com/, Bayesia Ltd) is commercially available to learn Bayesian networks, both parameters and structure. However, discretization of continuous variables is also required. This tool does not support utility and decision nodes.

Bayes Net Toolbox (http://people.cs.ubc.ca/~murphyk/Software/BNT/bnt.html, Murphy K) is a widely used and powerful mathematical software package, and runs only on Matlab. This free software supports both parameter and structure learning.

gR (http://www.ci.tuwien.ac.at/gR/), a language and environment for statistical computing and graphics, provide free packages to learn Bayesian networks. Package deal (Bøttcher and Dethlefsen, 2003) can deal with both discrete and continuous variables in learning structure and parameters. This package also allows transferring information to HUGIN interface.

2.9 References

Barker, G.C., Malakar, P.K., Del Torre. M., Stecchini, M.L., & Peck, M.W. (2005). Probabilistic representation of the exposure of consumers to Clostridium botulinum neurotoxin in a minimally processed potato product. *Int J Food Microbiol*, 100, 345-57.

Barker, G.C., Talbot, N.L.C., & Peck, M.W. (2002). Risk assessment for Clostridium botulinum: a network approach. *International Biodeterioration and Biodegradation*, 50, 167-75.

Barnett, G.O., Famiglietti, K.T., Kim, R.J., Hoffer, E.P., & Feldman, M.J. (1998). DXplain on the Internet. *Proc AMIA Symp*, 607-11.

Blodgett, J.G. & Anderson, R.D. (2000). A Bayesian network Model of the Consumer Complaint Process. *Journal of Service Research*, 2, 321-38.

Carlin, F., Girardin, H., Peck, M.W., Stringer, S.C., Barker, G.C., Martinez, A., Fernandez, A., Fernandez, P., Waites, W.M., Movahedi, S., Van Leusden, F., Nauta, M., Moezelaar, R., Torre, M.D., & Litman, S. (2000).'Research on factors allowing a risk assessment of spore-forming pathogenic bacteria in cooked chilled foods containing vegetables: a FAIR collaborative project. *Int J Food Microbiol*, 60, 117-35.

Charniak, E. (1991). Bayesian Networks without Tears. AI Magazine, 12, 50-63.

Corney, D.P.A. (2000), 'Designing food with Bayesian Belief Networks', In Parmee IC, *Evolutionary Design and Manufacture* ACDM2000, London, Springer-Verlag, 83-94.

Cunningham, S.J. (1995). Machine learning and statistics: a matter of perspective. Hamilton, New Zealand, University of Waikato, Department of Computer Science. Available from: http://waikato.researchgateway.ac.nz/handle/10289/1089 [Accessed 10 September 2009].

Daly, R., Qiang, C., & Aitken, S. (2011). Learning Bayesian networks: Approaches and issues. *The Knowledge Engineering Review*. 26(2):99-157.

Fearn, T. (2004). Bayesian statistics and the agro-food production chain. In van Boekel, M.A.J.S., Stein, A. & van Bruggen, A.H.C. (Eds.). *Bayesian Statistics and Quality Modeling in the Agro-Food Production Chain*, Dordrecht Kluwer Academic Publishers., 11-16.

Gupta, S. & Kim, H.W. (2007). Linking structural equation modeling to Bayesian networks: Decision support for customer retention in virtual communities. *European Journal of Operational Research*, 190, 818-33.

Heckerman, D. (1995). A tutorial on learning with Bayesian networks. Technical report MSR-TR-95-06, Microsoft Research. Available from: http://research.microsoft.com/apps/pubs/default.aspx?id=69588 [Accessed 10 September 2009].

Kjaerulff, U.B. & Madsen, A.L. (2008). *Bayesian Networks and Influence Diagrams: A Guide to Construction and Analysis*. Springer.

Lauritzen, S.L. (1995). The EM algorithm for graphical association models with missing data'. *Computational Statistics & Data Analysis*, 19, 191-201.

Murphy, K.P. (1998). A Brief Introduction to Graphical Models and Bayesian Networks. Berkeley, CA, Department of Computer Science, University of California. Available from: http://people.cs.ubc.ca/~murphyk/Bayes/bnintro.html [Accessed 10 September 2009].

Korb, K.B. & Nicholson, A.E. (2004). Appendix B - Software packages. *Bayesian*

Artificial Intelligence, Chapman & Hall/CRC. Available from: http://www.csse.monash.edu.au/bai/book/appendix_b.pdf [Accessed 10 September 2009].

Myllymäki, P., Silander, T., Tirri, H., & Uronen, P. (2002). B-Course: a web-based tool for Bayesian and causal data analysis. *International Journal on Artificial Intelligence Tools*, 11, 369-308.

Neapolitan, R.E. (2003). *Learning Bayesian Networks. Englewood Cliffs*, NJ, Prentice Hall.

Pourret, O., Naïm, P., & Marcot, B. 2008. *Bayesian Networks: A Practical Guide to Applications*, Wiley.

Repères research. Available from: http://reperes.eu/index.php?n=32&p=chap5 [Accessed 10 September 2009].

van Boekel, M.A.J.S. (2004). Bayesian solutions for food-science problems?. In van Boekel, M.A.J.S., Stein, A., & van Bruggen, A.H.C. (Eds). *Bayesian Statistics and Quality Modeling in the Agro-Food Production Chain*, Dordrecht Kluwer Academic Publishers., 17-27.

van Boekel, M.A.J.S. (2008). Kinetics modeling of reations in food. CRC Press.

van Raaij, J., Hendriksen, M., & Verhagen, H. (2008). Potential for improvement of population diet through reformulation of commonly eaten foods. *Public Health Nutrition*, 8, 1-6.

2.10 Appendix 2.A

A consumer test (n = 200) was hypothetically performed. A sample (20 cases) of the hypothetical data is shown in Table 2.1. The test was designed by 2 x 2 treatment combinations, which comprised of two snack types: sweet and salty, and two eating environments: alone and with friends In each treatment condition, teenagers scored their liking for the snack, and their *ad libitum* intake was recorded. Data in Table 2.1 were generated by HUGIN software and cases (1 case = results of one subject per treatment) are listed randomly, i.e. not necessarily in order of subject, test product or eating environment.

Table 2.1: Sample of snack consumption data

Case	Snack type	Eating with friends	Liking	Liking[1] (discretized)	Intake (g)	Intake[2] (discretized)
1	Salty	No	90	Very much	65.5	High
2	Sweet	No	62	Not very much	47.0	Medium
3	Sweet	No	50	Not very much	70.3	High
4	Sweet	No	56	Not very much	39.5	Low
5	Salty	No	75	Very much	69.6	High
6	Sweet	Yes	82	Very much	72.0	High
7	Salty	Yes	88	Very much	80.0	High
8	Sweet	Yes	72	Very much	65.4	High
9	Salty	No	81	Very much	30.2	Low
10	Sweet	Yes	49	Not very much	74.0	High
11	Sweet	Yes	69	Not very much	67.6	High
12	Sweet	No	73	Very much	56.3	Medium
13	Sweet	No	91	Very much	54.0	Medium
14	Sweet	No	78	Very much	18.0	Low
15	Salty	No	55	Not very much	40.8	Low
16	Salty	Yes	54	Not very much	86.1	High
17	Salty	No	83	Very much	69.0	High
18	Sweet	No	92	Very much	73.5	High
19	Sweet	No	71	Very much	90.3	High
20	Salty	No	80	Very much	82.1	High
..	..	..	..	..	..	..

[1] Liking scores were obtained by subjective ratings on a continuous line hedonic scale ranging from 0 (Not at all) to 100 (Very much). These continuous data were converted into two categories: 'Not very much' (value < 70), and 'Very much' (value >= 70).

[2] Intake data are also continuous and were discretized into three categories: 'Low' (value < 45), 'Medium' (45 <= value < 65), and 'High' (value >= 65).

CHAPTER 3

On the use of Bayesian networks to combine raw data from related studies on sensory satiation

Published as: V.A. Phan, M. P. Ramaekers, D.P. Bolhuis, U. Garczarek, M.A.J.S. van Boekel, M. Dekker. 2012. Food Quality and Preference, 26 (1), 119 - 127.

Abstract

Bayesian networks were used to combine raw datasets from two independently performed but related studies. Both studies investigated how different sensory aspects influence *ad libitum* intake of a tomato soup. The Aroma study varied aroma concentration and aroma duration as the explanatory variables, and the Taste study varied salt intensity. To enable data integration, the Aroma study needed information on salt aspects for all of its observations. Likewise, the Taste study needed information on aroma aspects. This information was used to link the two single networks, each representing one study, into a combined network. It was therefore referred to as Structural Linking Information. The approach taken was seen as an example to communicate a potential benefit as well as the challenges when combining raw datasets from independent studies. The combined network was able to generate additional insights into complex relationships encountered with research on satiation. The main challenge resulted from the missing of Structural Linking Information. In this chapter, we suggested different strategies to obtaining the structural linking information, and also proposed the approach of Global Experimental Design to avoid this problem. The nature of the chapter is theoretical rather than analytical due to the limitations caused by the small size of datasets.

3.1 Introduction

Food and nutrition researchers conduct controlled experiments to investigate causal relationships between explanatory variables and outcome variables. This type of experiments usually yields useful information to better understand mechanisms of the system behavior. However, as many variables are artificially kept constant, these experiments do not reflect the complexity of real-life situations. It is therefore of interest to understand the combined effects of independently manipulated variables on common outcome variables. Combining information from related studies can make this possible, and thereby provides more insights into a specific domain. Therefore, a practical tool supporting this combination is needed.

Meta-analysis is a popular statistical procedure that assists the combination of results obtained from related studies concerning a single theme (Charlton, 1996; Egger et al., 1997). This procedure has been mainly used in medial field and typically based on the summary characteristics that are available in published papers such as effect size, sample size, mean, and variance (Sutton & Higgin, 2008). Additional assumptions and statistical modeling approaches need to be carefully chosen to reduce bias and uncertainties. The goal of most meta-analyses is comparatively simple: estimating the effect of one explanatory variable on one outcome variable. All other variables (e.g. age, gender) are seen as noise factors that have to be taken into account appropriately. In food and nutrition research, however, the combined effects and the interactions among many influencing factors are of high interest. One needs to look beyond the published summary statistics for individual variables. For example, the correlational structure among variables should be taken into account. This information can be derived from raw datasets. Analyses of combined datasets can be superior to meta-analysis if done with the same amount of care because fewer assumptions are required. Despite its potential, examples and appropriate methodology for this approach have hardly been published.

When addressing complex relationships, domain or expert knowledge plays an important role in specifying causal relationships in model-building. Although this approach is used in the medical field, such as in health economy (Le and Doctor, 2011) and economics, by structure equation modeling (Hoyle, 1995), it is relatively unexplored by the food science community.

Bayesian networks are probabilistic graphical models consisting of two components: graphical (network structure) and probabilistic (network parameters) (Heckerman, 1995). The structure is a graph formed by a set of variables linked to each other by a set of arrows. These arrows imply possible cause-effect relationships. The network parameters are the set of conditional probability values that quantify these relationships. These two network components can be inferred and estimated based on the combination of empirical data and domain knowledge (Heckerman et al., 1995). Owing to its probabilistic and graphical nature, this modeling technique can handle complexity and uncertainty. When related studies yield different Bayesian network models that partly overlap, these networks can, under certain restrictions, be combined to build a larger single network. An example of this approach in biology has been shown by combining heterogeneous biological data sources to predict gene function (Troyanskaya et al., 2003).

Bayesian networks have been rarely applied in food-related problems despite the popularity of this technique in various fields. Published applications mostly deal with microbial risk assessment (Barker et al., 2005, Barker et al., 2002, Carlin et al., 2000, Smid et al., 2011). Corney (2000) has also discussed Bayesian networks as a valuable tool for food design by linking sensory attributes with consumer preference. More research is needed to further explore the potential of Bayesian networks in food design applications.

This chapter explore the potential use of Bayesian networks to combine raw data from related studies and the formal incorporation of domain knowledge in model-building. The exploration was based on two studies that were independently performed but closely related. The first study investigated the effect of retro-nasal aroma release profile on the *ad libitum* intake of a tomato soup (Ramaekers et al., submitted for publication). The second study investigated the effect of perceived intensity of saltiness on the *ad libitum* intake of two equally palatable tomato soups (Bolhuis et al., 2010). The combination of the datasets of these two studies was expected to result in a single model relating *ad libitum* intake of tomato soups to the combined effect of salt intensity and aroma release profile. The objective of this chapter was to use this practical example to communicate the approach and its potential through a general theoretical discussion. It has to remain theoretical because the studies were small in set up, and the available data was not sufficient to validate the predictive accuracy of the model.

This chapter deals with three issues: i) the requirements for combining raw data, ii) the strategies to obtaining the missing data needed for the combination, and iii) the recommendations for designing future related experiments such that their data can be combined.

3.2 Description of the case studies

Sensory perception has been suggested to contribute to satiation and thus to meal termination (Hetherington, 1996). The meal termination process can be assessed by *ad libitum* intake, i.e. the amount of food eaten by individual subjects till they are pleasantly satiated. Two experimental studies investigated how aroma and taste aspects influence *ad libitum* intake. It was hypothesized that increased sensory stimulation leads to lower food intake, which is referred to as sensory satiation. In this chapter, the *ad libitum* intake is expressed in weight (gram), and variable names and their states are put in single quotation marks.

3.2.1 Aroma study

The Aroma study (Ramaekers et al., submitted for publication) worked with four aroma release profiles combined with the same tomato soup base (Figure 3.1a illustrates one profile). These profiles resulted from a 2 x 2 crossover design with two variables: 'Aroma concentration' and 'Aroma duration'. The two states of 'Aroma concentration' were 'High' and 'Low', and those of 'Aroma duration' were 'Long' and 'Short'. The aroma profiles were determined based on some release profiles recorded in-vivo during natural consumption of a real tomato soup. They were then regenerated using an olfactometer in the actual experiment. The reference aroma profiles were referred to as 'Normal concentration' and 'Normal duration'. As compared to the reference profiles, the state 'High' of 'Aroma concentration' was higher, and the state 'Low' was lower, than the 'Normal concentration'. The state 'Short' of 'Aroma duration' was equal to the 'Normal duration'.

In the defined test conditions, different tomato aroma profiles were introduced into the nose of the subjects as they consumed the same soup base. The *ad libitum* intake of the soup was recorded for each aroma profile (Figure 3.1b).

The subjects were also asked to rate the 'Pleasantness' and 'Flavor intensity' after consuming the first 30 g of the soup. Data from 38 subjects were used for statistical analysis in the original paper.

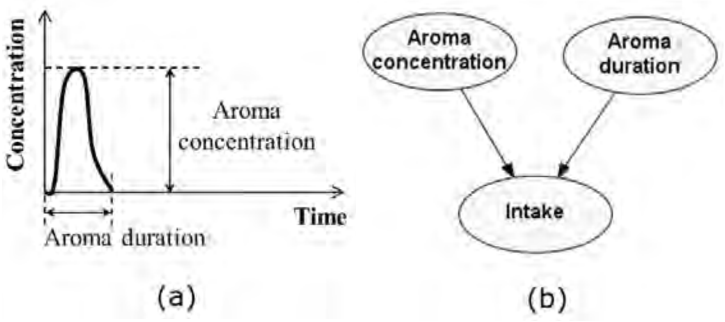


Figure 3.1: Illustrations of the Aroma study. 'Aroma concentration' and 'Aroma duration' were two derived variables representing an aroma release profile (a). The network (b) represents the investigated effects.

3.2.2 Taste study

The Taste study (Bolhuis et al., 2010) worked with two tomato soups that differed in salt concentration, namely 'Low' and 'High', but had similar rated pleasantness (Figure 3.2a). These two soups were first selected, in the pilot experiment, for each subject based on their individual pleasantness ratings for 5 soups varying in salt concentration. In the main experiment, the *ad libitum* intake was measured as the subjects consumed *in doublicate* their two soups (Figure 3.2b). Before each replicate, either 'Salt intensity' or 'Flavor intensity' and 'Pleasantness' were rated by tasting a soup sample. Data from 47 subjects were used for statistical analysis in the original paper.

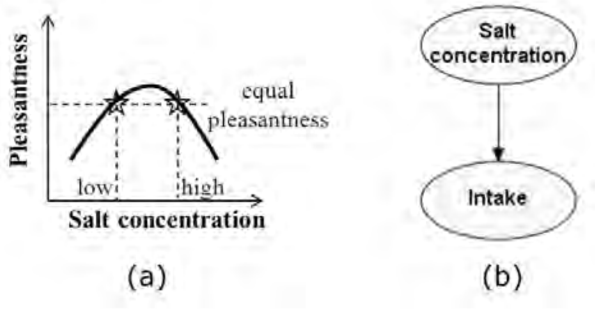


Figure 3.2: Illustrations of the Taste study. Two salt concentrations were chosen based on a pilot experiment, being 'low' and 'high' with similar pleasantness ratings (a). Network (b) represents the investigated relationship in the main experiment.

3.2.3 Compatibility for data combination

The tomato soups and experimental settings of the two studies were not identical. To check the validity of pooling the intake values in the combined model predicting 'Intake', the distributions of the 'Intake' values obtained from both studies were inspected (Figure 3.3). The two probability density distributions show a similar right-skewed shape. The median value of 'Intake' was 335 g for the Aroma study, and 353 g for the Taste study. Their difference of around 5% justifies combining these two *ad libitum* intake studies and generalizing the model for these types of soups.

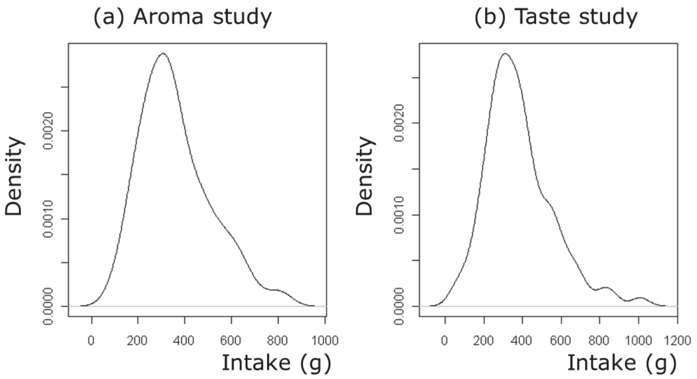


Figure 3.3: Probability density distribution of 'Intake' from the Aroma study (a) and Taste study (b).

In addition to 'Intake', the Aroma and Taste studies had two other common variables: 'Pleasantness' and 'Flavor intensity'. They were both rated on the same scale (Visual Analogue Scale 100 mm) with similar questions. These ratings were given after the subjects consumed 30 g of soups in the Aroma study or tasted about 15 g of soups in the Taste study. Thus, in both studies, 'Pleasantness' could be referred to as the initial pleasantness of the soup, and 'Flavor intensity' as the perceived overall flavor intensity. In other words, 'Pleasantness' obtained from the two studies was meant for the same concept and measured in a comparable manner. This compatibility allows us to treat them as one variable in the combined database. The same argument was also applied for 'Flavor intensity'.

3.3 Combining data

The Aroma and Taste studies were not initially designed for data integration using Bayesian networks. Instead, the technique of ANOVA was foreseen to separately analyze the treatment effects in each study. Therefore, this section looks first at the data situation and identifies the necessary requirement for combining data (3.3.1). We propose then some possible practices to meet this requirement (3.3.2), and introduce the combined database (3.3.3) for later use in the Bayesian modeling section.

3.3.1 Necessary requirement for combining data

Table 3.1 gives an overview of the raw combined database when pooling available data from the Aroma and Taste studies together. This database contains many missing values. These missing data were systematic because their pattern was not random, i.e. not being spread across the table but concentrated in certain columns for certain rows. This resulted from the experimental designs.

Theoretically, the Expectation-Maximization (EM) algorithm (Lauritzen, 1995), adapted in many Bayesian network software, can estimate missing values based on the available data. In the current situation, however, such estimation relies only on the information about dependencies in one study to fill in the missing values in another study. Because this process does not model the systematic differences, it would result in biased estimates. Therefore, one should use background knowledge of the original studies or other measures to fill in as much missing information as possible. In Table 3.1, part of the missing information is essential to integrate the two datasets and is referred to as "Structural Linking Information". Obtaining the Structural Linking Information was seen as the necessary requirement for combining raw data from related studies.

Table 3.1 Illustration of the raw combined database. The first four rows represent the observations from the Aroma study (38 subjects x 4 sessions); the last four rows represent the observations from the Taste study (47 subjects x 4 sessions).

Aroma concentration	Aroma duration	Salt concentration	Salt intensity	Flavor intensity	Pleasant-ness	Intake
High	Long	NA	NA	Avail	Avail	Avail
Low	Long	NA	NA	Avail	Avail	Avail
High	Short	NA	NA	Avail	Avail	Avail
Low	Short	NA	NA	Avail	Avail	Avail
NA	NA	High	Avail	NA	NA	Avail
NA	NA	High	NA	Avail	Avail	Avail
NA	NA	Low	Avail	NA	NA	Avail
NA	NA	Low	NA	Avail	Avail	Avail

'NA': data not being available; 'Avail': data being available.

3.3.2 Obtaining systematic missing information

In the Aroma study, the salt concentration of the soup base was not reported. However, the information of the soup's ingredients could be tracked, and salt concentration was calculated to be 208 mg Na/100 g soup (same unit as in the Taste study). This information was then filled in for 'Salt concentration' of all the Aroma's observations. Although 'Salt concentration' was unchanged, 'Salt intensity' ratings could vary across subjects and across measurements. To obtain information on 'Salt intensity', we could recruit, in principle, the same subjects to rate the perceived saltiness of the soup for each test condition. For practical reasons, however, this action could not be taken. Therefore, 'Salt intensity' values for the Aroma's observations were left missing in the final combined database.

In the Taste study, there was no indication about 'Aroma concentration' and 'Aroma duration'. It is known, however, that the aroma aspects of the soups were not altered, and the subjects consumed the soups in a natural setting. Additionally, the reference aroma profiles of the Aroma study were measured during the consumption of a real tomato soup, which is comparable to the situation in the Taste study. The aroma profiles in the Taste study can be thus approximated by the reference aroma profiles. In the Aroma study, the reference aroma profiles were identified with 'Normal' for 'Aroma concentration' and 'Short' for 'Aroma duration'. As a result, we assigned 'Normal' for 'Aroma concentration' of the Taste's observations, and 'Short' for 'Aroma duration'.

Furthermore, in the Taste study, the *ad libitum* intake was measured in duplicate for each salt concentration. There were missing values for 'Flavor intensity' and 'Pleasantness' in one replicate and missing values for 'Salt intensity' in the other replicate. The researchers intended to do so to obtain 'Salt intensity' ratings that do not interfere with the other two ratings. As these missing values were inevitable, they were left missing in the combined database. Nevertheless, by conducting two replicates, the information on all these three variables was available for both states 'Low' and 'High' of 'Salt concentration'. This would lead to a more accurate estimation by the EM learning process for these missing values.

3.3.2 Combined database

Table 3.2 summarizes the combined database to be used to learn the combined network model in the following section. Some missing values in the raw combined database (Table 3.1) were either calculated or assigned a state based on available information from the experimental studies. The current combined database possessed far less missing data (Table 3.2). The EM algorithm will estimate the missing values based on the available information in the modeling process.

In the Aroma study, the mean intake of the first testing session was found to be significantly lower than the three following sessions (Ramaekers et al., submitted for publication). To avoid this bias possibly due to the experimental set-up, data from the first session were excluded. The final combined database (Table 3.2, N = 306) consisted of 118 observations from the Aroma study and 188 observations from the Taste study. Although the two studies had a within-subject design, this combined database did not include information about subjects. A new variable 'd_Intake' (difference in 'Intake') was then introduced into the combined database to assess the within-subject variation. This variable was calculated from the 'Intake' values: 'd_Intake' = 'Intake' – 'individual mean intake'.

Table 3.2 Illustration of the combined database (N = 306). The new variable 'd_Intake' was calculated from 'Intake' to capture within-subject variation (see text).

Aroma concentration	Aroma duration	Salt concentration	Salt intensity	Flavor intensity	Pleasant-ness	Intake	d_Intake
High	Long	calculated	NA	Avail	Avail	Avail	Avail
Low	Long	calculated	NA	Avail	Avail	Avail	Avail
High	Short	calculated	NA	Avail	Avail	Avail	Avail
Low	Short	calculated	NA	Avail	Avail	Avail	Avail
Normal	Short	Avail	Avail	NA	NA	Avail	Avail
Normal	Short	Avail	NA	Avail	Avail	Avail	Avail
Normal	Short	Avail	Avail	NA	NA	Avail	Avail
Normal	Short	Avail	NA	Avail	Avail	Avail	Avail

'NA': data not being available; 'Avail': data being available.

3.4 Bayesian network modeling

The Bayesian network modeling involves learning, i.e. inferring or estimating, two components of a model network: structure and parameters. Having obtained these components, model users can perform inferences on the network to examine the relationships among variables and to make predictions. This section presents first how domain knowledge can be used to define the causal relationships (structure) for the single and combined networks (3.4.1), then explains the automatic parameter learning based on data (3.4.2). Due to the lack of data, not all parameters could be estimated reliably. Nevertheless, the inference in these networks is presented (3.4.3) to illustrate the Bayesian network modeling in general, and to explore and discuss the concept of combining raw data. HUGIN Bayesian networks software (HUGIN Researcher 7.2, tutorial available at http://www.hugin.com/) was used for both learning parameters and inference.

3.4.1 Defining network structures

Two single structures, Aroma network and Taste network (Figure 3.4), were formed by including 'Pleasantness', 'Flavor intensity' and 'Salt intensity' into the initial network of each study (Figure 3.1b, Figure 3.2b).

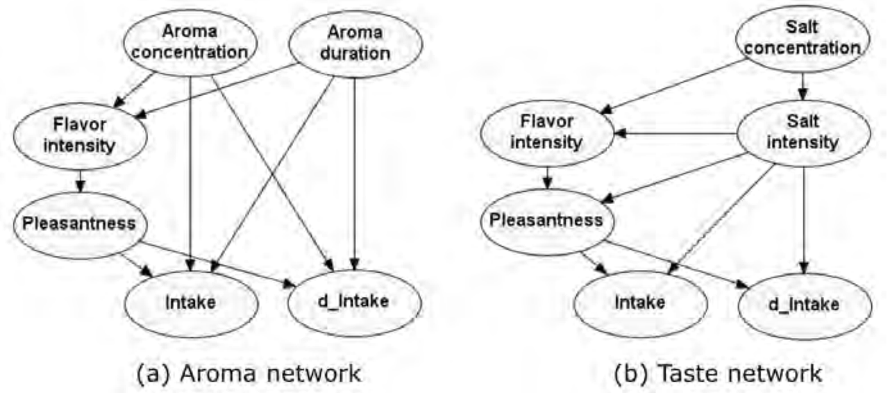


Figure 3.4: Structure of two single networks.

The arrows linking these variables to the existing ones were defined based on domain knowledge. Aroma and taste aspects contribute to the overall flavor perception of a food product (Auvray & Spence, 2008). Hence, four arrows were set from 'Aroma duration', 'Aroma concentration', 'Salt concentration' and 'Salt intensity' to 'Flavor intensity'. The flavor intensity in turn determines largely consumer liking for the food (Auvray & Spence, 2008), which was expressed by the arrow from 'Flavor intensity' to 'Pleasantness'. Being part of the overall flavor perception, 'Salt intensity' was also taken into account as a direct contributor to 'Pleasantness'. The amount of food eaten is influenced by how much a subject finds it pleasant (Zandstra et al., 1999; Zandstra et al., 2000; Vickers et al., 2001), resulting in 'Pleasantness' → 'Intake'.

Moreover, the investigated relationship 'Salt concentration' → 'Intake' (Figure 3.2b) was absent in the Taste network as we assumed in this work that the influence of 'Salt concentration' on 'Intake' goes via 'Salt intensity'.

The structure of the combined network (Figure 3.5) was simply formed by piling up the two single structures. It inherited all variables and arrows of the Aroma and Taste networks. One should be aware that more relationships in these networks could be learned if the data support this.

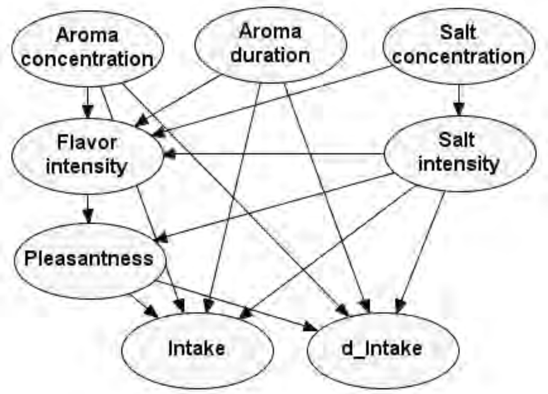


Figure 3.5: Structure of the combined network.

3.4.2 Parameter learning

HUGIN software version 7.2 supports both parameter learning and structure learning from data only with discrete variables (variables with finite number of states). The continuous variables were thus discretized, including 'Flavor intensity', 'Intake', 'd_ Intake' and 'Pleasantness'. The discretization boundaries for the first three variables were chosen such that all the states had an almost equal number of observations. The boundaries applied for 'Pleasantness' were technically meaningful. For instance, ratings for 'Pleasantness' below 50 are in practice considered as not good for a commercial product; ratings between 50 and 70 are acceptable; and ratings higher than 70 are good. Though discrete, 'Salt concentration' was re-set with fewer states. Appendix 3.A provides information on the final states of all variables.

Provided with a network structure and a discrete database, the software built a Conditional Probability Table (CPT) for each variable. This table contains conditional probabilities of the variable taking a specific state given the states of its parents. When having no parents, these probabilities are simply the relative counts of the number of observations differentiated by that state. This rule applies to 'Aroma concentration', 'Aroma duration' and 'Salt concentration' in the combined network. When having one or more parents, these relative counts are estimated for each combination of states of the parents. This is illustrated in Table 3.3 that shows the CPT of 'Pleasantness' under nine state combinations of its two parents 'Salt intensity' and 'Flavor intensity'.

Table 3.3: Conditional Probability Table (CPT) of 'Pleasantness' in the combined network. The CPT shows the probabilities of 'Pleasantness' (in italic) taking one of its states given 9 state combinations of two parents 'Salt intensity' and 'Flavor intensity'. 'Experience' indicates the number of observations for each combination of states of the parents.

Salt intensity	0-33			33-66			66-100		
Flavor intensity	0-50	50-65	65-100	0-50	50-65	65-100	0-50	50-65	65-100
Pleasantness = 0-50	0.79	0.09	0.12	0.25	0[1]	0.19	0[1]	0.41	0.69
Pleasantness = 50-70	0.01	0.22	0.22	0.75	0.71	0[1]	0.55	0.45	0.27
Pleasantness = 70-100	0.20	0.69	0.66	0[1]	0.29	0.81	0.45	0.14	0.04
Experience[2]	44	43	25	40	12	37	19	34	52

[1] values below 0.005 were replaced by 0 for simplification

[2] all numbers were rounded. The original numbers were not integer due to the automatic estimation of missing values on 'Salt intensity' for the Aroma's observations

The records on 'Experience' for each combination indicated the total number of the available observations that matched the information regarding parents' states. The probabilities of 'Pleasantness' taking each of its three states were estimated based on these observations. For example, the conditional probability distribution of 'Pleasantness' given 'Salt intensity = 0-33' and 'Flavor intensity = 0-50' was (0.79, 0.01, 0.20), and its 'Experience' = 44 (Table 3.3). This information means there were 44 observations in the combined database satisfying both 'Salt intensity = 0-33' and 'Flavor intensity = 0-50'. Among these, 79% observations had 'Pleasantness = 0-50', 1% had 'Pleasantness = 50-70', and 20% had 'Pleasantness = 70-100'.

The parameters of a variable are defined as the set of conditional probability values in its CPT. Hence, the number of parameters of a variable or the size of its CPT is the product of the number of its states, number of parents, and number of parent's states. The amount of data required for parameter learning in a Bayesian network depends on the number of parameters of the largest CPT. As a rule of thumb, the minimum number of observations required is five to ten fold the number of parameters of the largest CPT (Spirtes et al., 2000). In the combined network, 'Flavor intensity', 'Intake', and 'd_Intake' had the largest CPT size containing 162 parameters. The combined database (N = 306) is smaller than two fold of this number. This lack of data was reflected in many zero records on 'Experience' for the CPT of these variables (data not shown). This reveals that a number of conditional probability values were estimated based on no data at all. Consequently, these parameters were not reliable with the current available data. It should thus be kept in mind that the predictions made in the presented networks should be interpreted as hypotheses-generating rather than hypotheses-testing.

3.4.3 Inference

In Bayesian networks, the inference procedure is an automatic calculation of probabilities of interest given certain information on one or more variables of the model network. This procedure is illustrated with the Aroma network and then with the combined network. The Aroma network resulted from the Aroma data (N=118) and Aroma network structure (Figure 3.4a), the combined network resulted from the combined database (Table 3.2) and combined network structure (Figure 3.5). For convenience, three states of the discretized variables are later on referred to as 'Low', 'Medium' and 'High', respectively.

Aroma network

Prior to examining the effect of the aroma release profile, let us see how the pleasantness of the tested soup influences intake (Figure 3.6). Overall, a larger effect of 'Pleasantness' was seen on 'Intake' than on 'd_Intake'. When 'Pleasantness' was shifted from 'Low' to 'Medium' and then 'High' (c), the probability of 'Intake' being 'High' increased markedly, from about 5% to 36%, and then to 48%. For 'd_Intake', an increase was visible when 'Pleasantness' shifted from 'Low' to 'Medium', yet little change was observed with 'Medium'-to-'High' shift. Thus, it can be said that how much an individual consumes of a soup differs considerably if he disliked (0-50) or liked (50-70) the soup, but not much if he liked or liked very much (70-100) the soup. These inferences showed that the influence of pleasantness on food intake is more important when assessing a population effect than an individual effect.

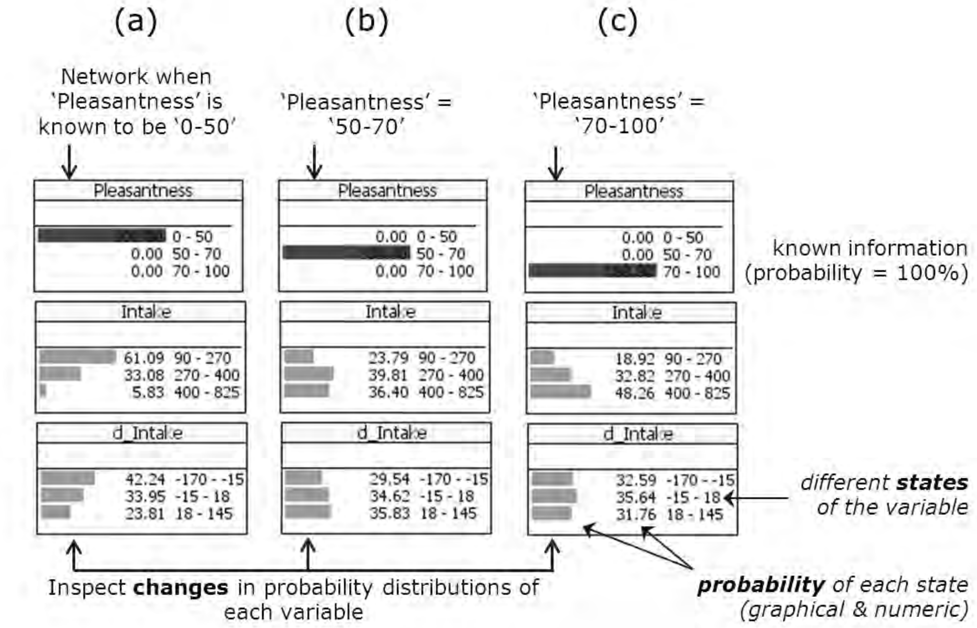


Figure 3.6: Influence of 'Pleasantness' on 'Intake' and 'd_Intake'. When 'Pleasantness' was known (assumed) to be '0-50', the probability of this event was set equal to 100% (a). The probability distributions of all other variables in the network were automatically calculated given this information (only shown for 'Intake' and 'd_Intake', variables being arranged in column). These probability distributions were different when 'Pleasantness' = '50-70' (b) and when 'Pleasantness' = '70-100' (c). Changes in these distributions show how the pleasantness of the tested soup influences intake.

With classical statistical analysis, the aroma profile 'high+long' (high concentration and long duration) was found to produce a significant lower *ad libitum* intake than the three other profiles (Ramaekers et al., submitted for publication). This finding could also be visualized via inferences on the Aroma network (Figure 3.7). Indeed, the probability of having a low 'd_Intake' was considerably higher for 'high+low' profile (about 51%) than for the other three aroma profiles (ranging from 25% to 31%). The effect of aroma profile was much less pronounced for 'Intake'. This result was expected because the large between-subject variation is only reflected in 'Intake', not in 'd_Intake', which resulted in more noise disguising the effect.

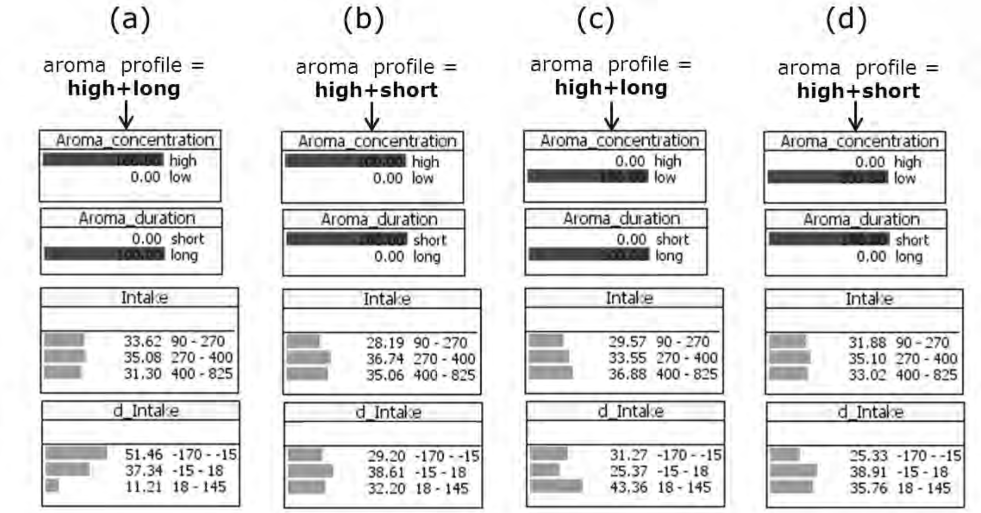


Figure 3.7: Influence of aroma release profile on 'Intake' and 'd_Intake'. Each aroma profile is identified by combining the information on 'Aroma concentration' and 'Aroma duration'. The probability distributions of 'Intake' and 'd_Intake' under different aroma profiles (a, b, c, d) are subject to comparison.

Combined network

Figure 3.8 shows one example of the combined effects of aroma and taste aspects on soup intake. In this figure, the effect of ' Salt concentration' on 'Intake' and 'd_ Intake' at the state 'Long' of 'Aroma duration' was examined. The 'Pleasantness' is fixed ('50-70') to partly rule out the indirect influence of 'Salt concentration' through the 'Pleasantness' pathway. Changes in probability distributions indicate that both 'Intake' and 'd_Intake' seemed to decrease when 'Salt concentration' increased. This pattern was, however, not seen when 'Aroma duration' = 'Short' (inferences not shown). This interaction effect between 'Aroma duration' and 'Salt concentration' can be considered as hypotheses in future studies.

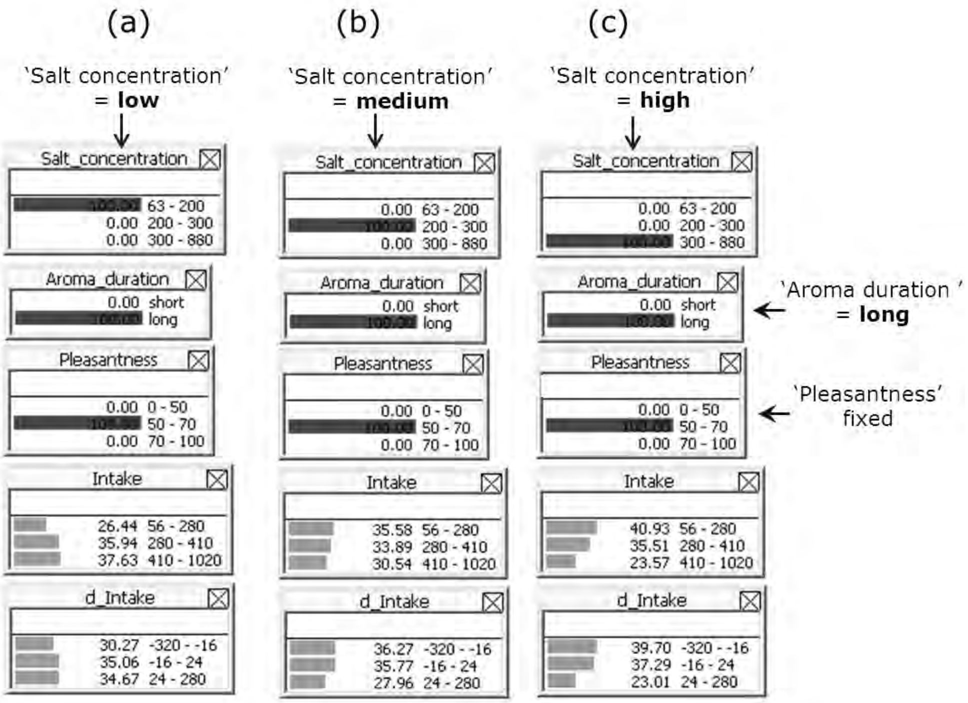


Figure 3.8: Influence of 'Salt concentration' on 'Intake' and 'd_Intake' at the state 'Long' of 'Aroma duration'. The effect of 'Pleasantness' is partly ruled out by fixing it at any state (shown at '50-70'). The probability distributions of 'Intake' and 'd_Intake' under different salt concentrations (a, b, c) are subject to comparison.

3.5 Discussion

This chapter explored the use of Bayesian networks to combine raw data from related studies and to ultimately build a combined model network. First, the discussion focuses on how the Bayesian network modeling technique can cope with data from controlled studies in food research (3.5.1). Second, it takes on the possibility to make use of related databases that have been available (3.5.2). Finally, recommendations are given on how to design future studies such that their results can be combined later on with the Bayesian network framework (3.5.3).

3.5.1 Modeling with Bayesian networks

Bayesian networks formalize the use of domain (expert) knowledge in building the network structure as explained in section 3.4.1. The network structure can also be learned automatically from data. This learning normally requires a vast amount of data, which are unfortunately often scarce in controlled experiments. Yet, controlled experiments allow us to test our hypotheses on cause-effect relationships. Published literature is therefore a reliable source to be referred to as domain knowledge. The specification of the structure can also involve beliefs of domain experts. Since this process is subjective, different groups of experts may not yield the same network structure for the same set of variables. This subjectivity is partly the nature of modeling. Bayesian networks make those assumptions transparent, open for discussion. Furthermore, validation with new data is always the best measure to judge which model is more useful.

Concerning parameter learning from data, section 3.4.2 has shown that a larger amount of data is required than classical statistical analysis. Bayesian networks involve inferences or predictions that always need much more data than a hypothesis testing procedure, such as ANOVA. To reduce the required amount of data, modelers should limit the number of parameters to be estimated. This can be controlled by limiting the number of parents of the variable with largest CPT. Too-many-parents problem can be solved by introducing a hidden variable that captures the influence of two or more parents (Kjaerulff & Madsen, 2008). The number of parameters can considerably decrease as well if continuous data are not discretized. Some Bayesian network software can estimate parameters only for discretized data, e.g. Netica (http://www.norsys.com/, Norsys Software Corp.) and BayesiaLab (http://www.bayesia.com/, Bayesia Ltd.). In this case, a fewer number of states per variable is technically favorable. The latest version of HUGIN Bayesian network software (version 7.5) do support parameter learning for continuous variables, but not yet structure learning. However, there are some other available Bayesian networks software that are able to deal with both tasks for continuous variables (Murphy, 2005).

Inferences in Bayesian networks have been illustrated in section 3.4.3. This procedure allows the influence of any variable on the rest of the network to be easily examined and communicated. It is of particular value when dealing with complex models, which generally cause great difficulty to conventional statistical models. Moreover, a combined network could generate new hypotheses for the research field. For example, new knowledge can come from the prediction of interaction effects of separately controlled variables. The combined network also provides a global view of the model, which actively supports the reasoning process when examining the problem.

3.5.2 Combining related databases

If a number of related databases are already available, three steps can be followed to build a combined network from these databases. The first step is to build a network for each database and then build a combined network based on these single networks. The second step is to construct the raw combined database and to identify the Structural Linking Information among the systematic missing data. Structural Linking Information is the missing data that are essential for the combined network. The third step is to obtain this Structural Linking Information by other means.

More case studies applying this approach are needed to give a general guidance on how to judge which systematic missing data are essential (i.e. Structural Linking Information). Examples of essential and non-essential missing data in the current combined database are discussed as follows. The information on 'Salt intensity' for Aroma's observations and that on 'Aroma concentration' and 'Aroma duration' for Taste's observations were essential (section 3.3.1). The availability of this information would allow for new variables to be added to the individual networks. The Aroma network could be then extended with 'Salt intensity', and the Taste network with 'Aroma concentration' and 'Aroma duration'. As a result, two single networks would share five common variables predicting 'Intake' instead of two, namely 'Pleasantness' and 'Flavor intensity'. These extra common variables strengthen the link between the single networks. More importantly, they carry the hypothesis of the original studies, as they were experimentally controlled variables. Conversely, the missing information on 'Salt concentration' in 'Aroma' can be considered as not essential. The reason is that 'Salt concentration' has no direct arrow to 'Intake', and its influence on 'Intake' is captured by 'Flavor intensity' and 'Salt intensity'. However, when for example 'Sweetness' is included into this network, 'Salt concentration' might have an effect on this variable, thus information on 'Salt concentration' could become essential. For this reason, it would still be better to have such primary information collected.

Having identified the Structural Linking Information, researchers can work on different strategies to obtain this. Some information is easy to extract from the materials and methods of the individual studies. It was the case for salt concentration of the soup used in the Aroma study. Some information can be obtained based on their experimental designs. Assigning states of 'Aroma concentration' and 'Aroma duration' for the Taste's observations was one example. Some information requires extra experiments, individual salt intensity ratings for the Aroma study for instance. Arguably, such measures might not be satisfactory. It might thus be necessary to consult with and obtain consensus from domain experts to justify the use of the obtained data.

3.5.3 Designing future studies to enable the combination of their data

As discussed in the earlier sections, difficulties may arise when combining data of independently performed experiments if they were not originally designed for integration, despite being closely related. However, many problems can be avoided if researchers envision an overall network before designing small experiments that cover part of it.

Envisioning an overall network, designated as Global Experimental Design, means i) identifying all variables of interest of a research theme, ii) standardizing the measurement method for each variable, as well as iii) defining states (or all possible values) that each variable can get. From this overall network, independent studies involving a smaller number of variables can be designed to support the future combination of all the datasets. The first important message is that the states of all variables are judged not only within a single study, but also on a global scale (overall network). The second important message is that each study might have to gather more information than needed for its own scope. In the case studies of this chapter, subjects participating in the Aroma study should have been asked to rate the perceived salt intensity in each testing session. This information might not be of direct value for the hypothesis testing procedure in the Aroma study; yet, it plays an essential role in the combined network as explained in section 3.5.2.

3.6 Conclusions

Bayesian networks act as a complementary modeling technique to classical statistical analysis. This modeling technique can be a potential tool to combine raw data from related studies, resulting in a combined model network. If these studies are not initially designed to be integrated, the systematic missing data in the combined database can cause unreliability in estimating model parameters. Some of these systematic missing data could be essential for the combined network and are referred to as Structural Linking Information. A general guidance on this judgment requires more applied works. Obtaining Structural Linking Information is identified as the necessary requirement for the presented approach. This information can be derived from the background information of the original studies or obtained by extra experiments. These strategies are subject to careful consideration and agreement among the researchers. To prevent the lack of Structural Linking Information, the proposed approach of Global Experimental Design can be used before conducting small and independent studies concerning a specific research theme. The technique of Bayesian networks is a potential tool to combine of different sources of data and to formally incorporate domain knowledge in the model-building process. Such features would allow scientists to gain more information and a more holistic view of the research theme. This chapter is a unique contribution to the Bayesian network modeling field as a data-driven approach rather than a traditional simulation-driven approach.

3.7 References

Auvray, M., & Spence, C. (2008). The multisensory perception of flavor. *Consciousness and Cognition*, 17(3), 1016-1031.

Barker, G. C., Malakar, P. K., Del Torre, M., Stecchini, M. L., & Peck, M. W. (2005). Probabilistic representation of the exposure of consumers to Clostridium botulinum neurotoxin in a minimally processed potato product. *Int J Food Microbiol*, 100(1-3), 345-357.

Barker, G. C., Talbot, N. L. C., & Peck, M. W. (2002). Risk assessment for Clostridium botulinum: a network approach. *International Biodeterioration & Biodegradation*(50), 167-175.

Bolhuis, D. P., Lakemond, C. M. M., de Wijk, R. A., Luning, P. A., & de Graaf, C. (2010). Effect of Salt Intensity on *Ad libitum* Intake of Tomato Soup Similar in Palatability and on Salt Preference after Consumption. *Chemical Senses*, doi: 10.1093/chemse/bjq077.

Carlin, F., Girardin, H., Peck, M. W., Stringer, S. C., Barker, G. C., Martinez, A., et al. (2000).

Research on factors allowing a risk assessment of spore-forming pathogenic bacteria in cooked chilled foods containing vegetables: a FAIR collaborative project. *Int J Food Microbiol*, *60*(2-3), 117-135.

Charlton, B. G. (1996). The uses and abuses of meta-analysis. *Family Practice, 13*(4), 397-401.

Corney, D. P. A. (2000). Designing food with Bayesian Belief Networks. *Parmee, I.ed. , Adaptive computing in design and manufacture*, 83-94.

Egger, M., Smith, G. D., & Phillips, A. N. (1997). Meta-analysis: principles and procedures. *British Medical Journal, 315*(6), 1533-1537.

Heckerman, D. (1995). A tutorial on learning with Bayesian networks. *Technical report MSR-TR-95-06, Microsoft Research*.

Heckerman, D., Geiger, D., & Chickering, D. M. (1995). Learning Bayesian networks: The combination of knowledge and statistical data. *Machine Learning, 20*(3), 197-243.

Hetherington, M. M. (1996). Sensory-specific satiety and its importance in meal termination. *Neuroscience & Biobehavioral Reviews, 20*, 113-117.

Hoyle, R.H. (1995). *Structural Equation Modeling: Concepts, Issues, and Applications*: SAGE Publications.

Kjaerulff, U. B., & Madsen, A. L. (2008). *Bayesian Networks and Influence Diagrams: A Guide to Construction and Analysis*: Springer.

Lauritzen, S. L. (1995). The EM algorithm for graphical association models with missing data. *Computational Statistics & Data Analysis*(19), 191-201.

Le, Q.A., & Doctor, J.N. (2011). Probabilistic mapping of descriptive health status responses onto health state utilities using Bayesian networks: an empirical analysis converting SF-12 into EQ-5D utility index in a national US sample. *Medical Care 49*(5), 451-60.

Murphy, K. (2005). Software Packages for Graphical Models / Bayesian Networks. http://www.cs.ubc.ca/~murphyk/Bayes/bnsoft.html

Smid, J.H., Swart, A.N., Havelaar, A.H., Pielaat, A. (2011). A Practical Framework for the Construction of a Biotracing Model: Application to Salmonella in the Pork Slaughter Chain. Risk Analysis. doi: 10.1111/j.1539-6924.2011.01591.x.

Sutton, A. J., & HigginS, J. P. I. (2008). Recent developments in meta-analysis. *Statistics in Medicine, 27*(5), 625-650.

Troyanskaya, O. G., Dolinski, K., Owen, A. B., Altman, R. B., & Botstein, D. (2003). A Bayesian framework for combining heterogeneous data sources for gene function prediction (in Saccharomyces cerevisiae). *Proceedings of the National Academy of Sciences of the United States of Amercia, 100*(14), 8348-8353.

Vickers, Z., Holtonb , E., & Wanga, J. (2001). Effect of ideal–relative sweetness on yogurt consumption *Food Quality and Preference, 12*(8), 521-526.

Zandstra, E. H., de Graaf, C., Mela, D. A., & van Staveren, W. A. (2000). Short and long-term effects of changes in pleasantness on food intake. *Appetite, 34*, 253-260.

Zandstra, E. H., de Graaf, C., van Trijp, H. C. M., & van Stavenren, W. A. (1999). Laboratory hedonic ratings as predictors of consumption. *Food Quality and Preference, 10*, 411-418.

3.8 Appendix 3.A

Discretization of continuous variables in the combined network. Three intervals for 'Intake' and 'd_Intake' set for the Aroma network are not necessary the same as those set for the combined network.

Salt concentration (mg Na/100 g)	Salt intensity	Flavor intensity	Pleasantness	Intake (g)	d_Intake
63-200	0-33	0-50	0-50	56-280	(-320)- (-16)
200-300	33-66	50-65	50-70	280-410	(-16) – 24
300-880	66-100	65-100	70-100	410-1020	24-280
Low	Short	NA	NA	Avail	Avail

CHAPTER 4

Bayesian networks as a tool to analyze causal relationships in experimental designs: a case study on oro-sensory exposure studies

Submitted for publication by:

V.A. Phan, U. Garczarek, D.P. Bolhuis, M.A.J.S. van Boekel, M. Dekker.

Abstract

This chapter investigated how the causal relationships underlying the experimental design affect the possibility to combine raw data of related studies. Bayesian networks were used to re-examine the experimental design of two published studies on oro-sensory exposure. The role of each explanatory variable in the design was categorized as either directly manipulated (primary) or indirectly manipulated (secondary) through primary variables. It was shown that when a secondary variable is manipulated, causal relationships are reversed. Consequently, it can become impossible to meaningfully analyze the obtained data in combination with those from other related studies that do not follow the same causal structure. Using a secondary variable as the explanatory variable also makes it difficult to translate the findings to real-life situations. The current work has provided additional arguments and insights into using Global Experimental Design as a method to design related controlled experiments for data integration.

4.1 Introduction

Oro-sensory exposure to food leads to earlier meal termination as reviewed by de Graaf (2012). Oro-sensory exposure has been explained as the factor that mediates the observed differences in ad libitum intake due to viscosity (Zijlstra et al., 2008) and texture differences (Zijlstra et al., 2010). To investigate this mediating role further, Weijzen et al. (2009) have explicitly altered the oro-sensory exposure through a number of variables while controlling the eating rate of orangeade. Similar studies have been done with tomato soups (Bolhuis et al., 2011; Bolhuis et al., submitted for publication). Eating rate is known to largely affect the ad libitum intake (Zijlstra et al., 2008; Kellen, 2010). Therefore, the studies on oro-sensory exposure have fixed eating rate to rule out its effect (Weijzen et al., 2009; Bolhuis et al., 2011; Bolhuis et al., submitted for publication). These types of experimental designs involve the manipulation of multiple variables, it is consequently not easy to communicate the causal relationships in a transparent manner. Understanding the underlying causal relationships is important when combing data from independent studies because the causal structures affect the possibility to do a meaningful analysis of the combined data.

Bayesian networks are graphical probabilistic models consisting of two components: structure (graph) and parameters (probabilities) (Heckerman, 1995). The network structure represents the causal relationships among the variables; the network parameters quantify these relationships through probability expressions (Phan et al., 2010). The graphical nature of Bayesian networks is said to make it easy to communicate and comprehend the overall picture of a research domain; even if a large number of variables is involved. Phan et al. (2012) have previously demonstrated the potential of Bayesian networks to combine raw data from independently performed but related studies.

This chapter investigated how the causal relationships underlying the experimental design of related studies determine whether sets of data from separate studies can be combined. We re-examined closely the experimental design of two published studies on oro-sensory exposure (Bolhuis et al., 2011; Bolhuis et al., submitted for publication). Data from another related study (Bolhuis et al., 2012) were also analyzed in combination with data from those two studies. Bayesian networks were used as a tool for visualizing causal relationships, and evaluated for its effectiveness in establishing a clear view of the problem.

4.2 Materials and method

4.2.1 Description of the original studies

Study I investigated how the oro - sensory exposure time ('Food exposure time') and the intensity of saltiness ('Salt intensity') influence the ad libitum intake of tomato soup ('Intake') (Bolhuis et al., 2011). 'Food exposure time' was defined as the average time that food resides in the oral cavity calculated for one gram of food (s/g). A full crossover design was used. Each subject received all six combinations of three states of 'Food exposure time' being 'short', 'long', and 'free' and two states of 'Salt intensity' being 'low' and 'high'. In the current work, we did not take into account the free condition of the food exposure time in order to focus on the general aspects of the experimental design. A number of variables were manipulated to create the short and long exposure times, and the salt concentrations that gave perceptions of low and high intensity were chosen such that they were rated at similar levels of pleasantness.

The study consisted of two separate parts: a primary tasting and the actual crossover design on intake. First, the preliminary tasting was performed to select two soups having different salt concentrations but eliciting the same degree of pleasantness. These soups were assumed to represent the two states of 'Salt intensity'. Subjects tasted five soups varying in salt concentration, and rated pleasantness and relative-to-ideal salt intensity as described in (Bolhuis et al., 2010). The soup with ideal salt intensity was the most pleasant soup, whereas, the soups with low and high salt intensity were less pleasant than the 'ideal' and similar in pleasantness ratings (Figure 4.1). Only the soups with low and high salt intensity were chosen for the

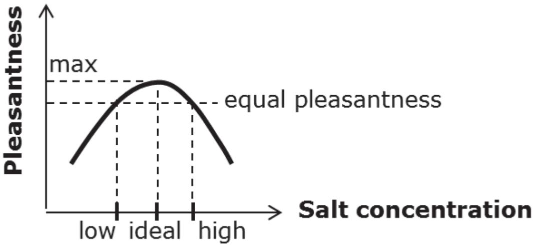


Figure 4.1: Illustration of the soups having low, ideal, and high salt intensity. The soup with ideal salt intensity is the most pleasant soup. The other two soups are similarly pleasant, though they differ in salt concentration. Adapted from Phan and Bolhuis et al. (submitted for publication).

In the intake experiment, subjects ate a fixed preload of raisin buns (calculated as 50% of the average energy intake for a lunch meal) before consuming ad libitum either the soup with low or high salt intensity. Subjects received the soups directly into their mouth via a food-grade tube connected with a peristaltic pump. Under this setting, the states of 'Food exposure time' were specified by controlling three variables: time interval between the start of two subsequent bites ('Bite interval', s), residence time of each bite in the oral cavity ('Bite residence time', s), and amount of each bite ('Bite size', g). The definitions of 'Bite interval' and 'Bite residence time' are illustrated in Figure 4.2. The bite residence time was the total time that the subjects received one bite of soup and kept this bite in their mouth before swallowing it.

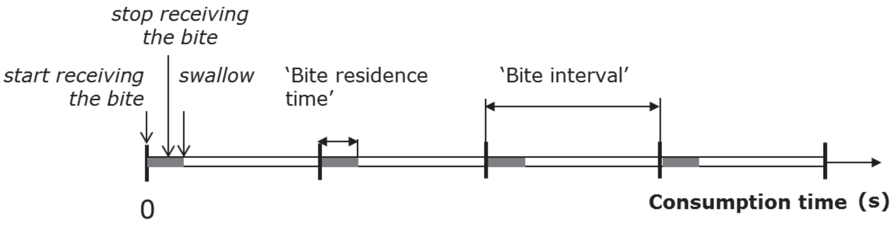


Figure 4.2: Illustration of bite residence time and bite interval.

Table 4.1 summarizes the experimental design of Study I. In the short condition of 'Food exposure time', 'Bite size' was set equal to 15 g, 'Bite interval' = 15 s, and 'Bite residence time' = 3 s. In the long condition of 'Food exposure time', these values were 5 g, 5 s, and 2 s, respectively. 'Bite size' and 'Bite interval' were chosen such that the amount of soup eaten per minute ('Eating rate', g/min) was equal to 60 g/min in both conditions of the exposure time. The ad libitum intake was recorded. Data from 55 subjects were taken into account in the statistical analysis of the original study.

Table 4.1: Experimental design of Study I. In a crossover design, 'Salt intensity' (low, high) and 'Food exposure time' (short, long) were used to explain 'Intake' under a constant 'Eating rate'. 'Bite size', 'Bite interval', and 'Bite residence time' were manipulated to obtain the desired conditions. The data on 'Intake' were observed (represented as 'Avail').

Salt intensity	Food exposure time (s/g)	Bite interval (s)	Bite size (g)	Bite residence time (s)	Eating rate (g/min)	Intake (g)
Low	Short	15	15	3	60	Avail
High	Short	15	15	3	60	Avail
Low	Long	5	5	2	60	Avail
High	Long	5	5	2	60	Avail

Study II investigated how the oro-sensory exposure time ('Food exposure time') and the number of bites per unit of food weight ('Food bite number', bites/g) influence the ad libitum intake ('Intake') of a tomato soup (Bolhuis et al., submitted for publication). A full crossover design was used. Each subject received all four combinations of two states of 'Food exposure time' (short or long) and two states of 'Food bite number' (low or high) as summarized in Table 4.2.

Table 4.2: Experimental design of Study II. In a crossover design, 'Food bite number' (low, high) and 'Food exposure time' (short, long) were used to explain 'Intake' under a constant 'Eating rate'. 'Bite size', 'Bite interval', and 'Bite residence time' were manipulated to obtain the desired conditions. The data on 'Intake' were observed (symbolized as 'Avail').

Food bite number (bites/g)	Food exposure time (s/g)	Bite interval (s)	Bite size (g)	Bite residence time (s)	Eating rate (g/min)	Intake (g)
Low	Short	15	15	3	60	Avail
Low	Long	15	15	9	60	Avail
High	Short	5	5	1	60	Avail
High	Long	5	5	3	60	Avail

One single tomato soup was served after the subjects ate a fixed preload of raisin buns (calculated as 50% of the average energy intake for a lunch meal). Like Study I, the food-grade tube and pump system supported the ad libitum intake experiment. 'Bite size' was set equal to 15 g to obtain the state 'low' and to 5 g to obtain the state 'high' of 'Food bite number'. At low food bite number, two states 'short' and 'long' of 'Food exposure time' were specified by setting 'Bite residence time' at 3 s and 9 s, respectively. At high food bite number, two states 'short' and 'long' of 'Food exposure time' were specified by setting 'Bite residence time' at 1 s and 3 s, respectively. 'Bite interval' was manipulated along with 'Bite size' to keep 'Eating rate' constant at 60 g/min in all conditions: 15 s when 'Bite size' = 15 g, and 5 s when 'Bite size' = 5 g. The ad libitum intake was recorded. Data from 57 subjects were taken into account in the statistical analysis of this study.

4.2.2 Analyzing causal relationships underlying experimental designs using Bayesian networks

In applied statistics, explanatory variables are defined as those that are used in a statistical model to explain the variation of the outcome variable, i.e. dependent variable. Explanatory variables are also called predictor variables, independent variables, input variables, regressors, etc. as the terminology is not highly harmonized

(Dodge et al., 2006; Everitt & Skrondal, 2010). In particular, the term "independent variables" when being used interchangeably with "explanatory variables" can be very confusing. In the general sense, explanatory variables may or may not be independent from each other, and may or may not be independently experimentally controlled. In the Cambridge Dictionary of Statistics (Everitt & Skrondal, 2010), it is recommended to abandon the use of this term for explanatory variables. When the overall causal structure among the variables is of interest, it is important to differentiate the variables being independently experimentally controlled (or manipulated) from the variables being explanatory. Therefore, we define here explicitly the exact use of terminology.

In this chapter, the term "primary explanatory variables" refers to explanatory variables that are used to explain the outcome variable in statistical models, and that can be independently manipulated in the experimental setting. The term "secondary explanatory variables" is used for explanatory variables that cannot be manipulated as such but are calculated/derived from primary explanatory variables. In studies having rather similar experimental setups but different goals, often the primary explanatory variables are the same (but not always reported), while the secondary ones might be different. It is thus crucial to differentiate the two types of explanatory variables when combining data of those studies. Special attention should be paid to the experimental designs whose intention is to control or manipulate secondary explanatory variables by the primary ones. Such designs have major influence on the causal structure among all variables observed in the experiments and may render a meaningful analysis on combined data with other experiments impossible.

The current work consisted of two main tasks. The first task was to revisit the experimental designs of the two original studies. The role of each explanatory variable being primary or secondary in the design was clarified. The second task was to consider the possibility to combine data from related studies to build a larger model, as described earlier by Phan et al. (2012), from the available data. The data combination was first analyzed solely using the two current original studies (Bolhuis et al., 2011; Bolhuis et al., submitted for publication), and then analyzed together with an additional related study (Bolhuis et al., 2012). These two tasks were performed using the qualitative aspect of Bayesian network modeling, i.e. the network structure or the causal relationships among the variables.

4.3 Results

4.3.1 Revisiting the experimental designs

Secondary explanatory variables versus primary explanatory variables. Study II used 'Food exposure time' and 'Food bite number' as explanatory variables of 'Intake'. However, looking into the details of the experimental setup, the variables that were actually manipulated to obtain a constant 'Eating rate' and varied 'Food bite number' and 'Food exposure time' were 'Bite interval', 'Bite size', and 'Bite residence time'. The network structure in Figure 4.3 illustrates the dependency among these six variables. In this network, each arrow is assumed to point from cause to effect, therefore, it represents a causal relationship.

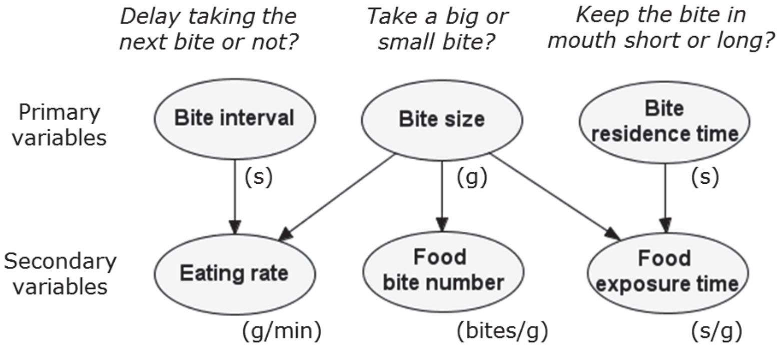


Figure 4.3: Network structure for the design of Study II.

Given the definition of the variables, it is straightforward that 'Eating rate' is equal to **60*Bite size/Bite interval**, 'Food bite number' is equal to **1/Bite size**, and 'Food exposure time' is equal to **Bite residence time/Bite size**. These deterministic relationships allowed us to draw the arrows from 'Bite interval' and 'Bite size' towards 'Eating rate', the arrow 'Bite size' → 'Food bite number', and the arrows from 'Bite size' and 'Bite residence time' towards 'Food exposure time'. 'Bite interval', 'Bite size', and 'Bite residence time' were taken as the *causes* because they are primary explanatory variables in the experimental setting. Also, as it is true in most cases, they are most easily manipulated in more natural settings of eating. That makes them easy to translate to consumer advice or to product design related to eating behavior in real life. Bite interval tells the consumers to delay or speed up in taking the next bite; bite size tells the consumers to take a big or small bite; bite residence time tells the consumers to keep the bite in their mouth for a long or short time. 'Eating rate', 'Food bite number', and 'Food residence time' were drawn as the *effects* because they are calculated from the primary variables, and hence are secondary variables. In that sense, 'Food bite number' and 'Food exposure time' were secondary explanatory variables of 'Intake' in Study II.

Experimental design versus natural setting. Eating rate was fixed at 60 g/min for all experimental conditions in both Study I and Study II. This criterion permitted the researchers to rule out the effect of eating rate on ad libitum intake. With the fixing of the eating rate, the two variables 'Bite interval' and 'Bite size' became dependent on each other. The network in Figure 4.4a illustrates this design, in which 'Bite interval' and 'Bite size' appear to be the effects of 'Eating rate'. In Figure 4.4b, the arrows Bite interval → Eating rate and Bite size → Eating rate could be justified by the deterministic relation: **Eating rate = 60*Bite size/Bite interval**. We use the term 'natural setting' because it reflects the causal structure among the three variables as observed during normal eating situations (Figure 4.4b). It can therefore be said that the experimental design of the original studies reversed the causal relationships between 'Bite interval', 'Bite size', and 'Eating rate' as compared to natural eating occasions.

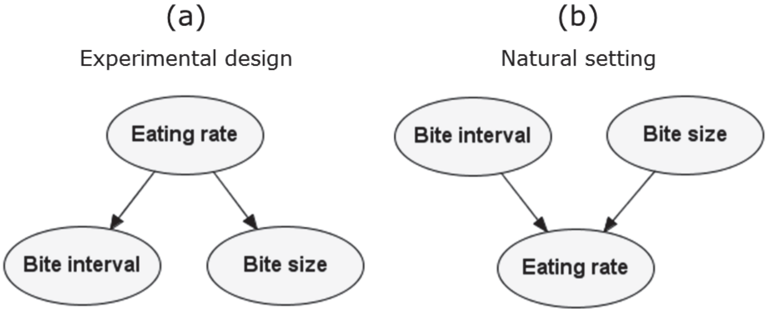


Figure 4.4: Reversal of the causal relationships. A possible causal relationship between 'Bite interval' and 'Bite size' in real life eating occasions was not taken into account.

Similar experimental designs. Study I was designed to investigate the influence of food exposure time and salt intensity on the *ad libitum* intake; Study II was designed to investigate the influence of food exposure time and food bite number on the *ad libitum intake*. Despite these differences, the two experimental designs can be considered as similar because of the two following reasons:

First, despite the absence of ‘Food bite number’ in Study I, the effort made to obtain different values for ‘Food exposure time’ forced the control of the same variables as in Study II. The information on ‘Food bite number’ was thus available in Study I. Second, the absence of ‘Salt intensity’ in Study II can be rectified. This study did not contain the preliminary tasting session related to salt intensity ratings. However, salt concentration for the tested soup was chosen based upon the data of Study I, such that the soup was the most pleasant perceived by most subjects. Hence, we were able to create the variable ‘Salt intensity’ with a single state ‘ideal’ for Study II.

As a result, the data of the two original studies can be summarized in one single combined database after some modifications (Table 4.3). In this database, the state ‘ideal’ of ‘Salt intensity’ was assigned to all observations in Study II. The two states of ‘Food exposure time’ were given the same name (short, long) in both studies. However, they did not represent the same value. Then, these ordinal states were converted to arithmetic values using the equation **Food exposure time = Bite residence time/ Bite size** (s/g). The equation **Food bite number = 1/ Bite size** (bites/g) was used to make the same conversion for ‘Food bite number’ in Study II and to obtain the information on ‘Food bite number’ in Study I.

Table 4.3: Combined database. The information related to ‘Salt intensity’, ‘Food exposure time’, and ‘Food bite number’ was modified such that the data of Study I (represented by the first four lines) and Study II (represented by last four lines) can be combined. The state ‘ideal’ of ‘Salt intensity’ was created for all the observations of Study II. The ordinal states of ‘Food exposure time’ and ‘Food bite number’ were replaced by the arithmetic values, which resulted from the calculation containing ‘Bite size’ and ‘Bite residence time’. The old information is put in parentheses.

Salt intensity		Food exposure time (s/g)		Food bite number (bites/g)		Bite interval (s)	Bite size (g)	Bite residence time (s)	Eating rate (g/min)	Intake (g)
(Low)	Low	(Short)	0.2	(NA)	0.07	15	15	3	60	Avail
(High)	High	(Short)	0.2	(NA)	0.07	15	15	3	60	Avail
(Low)	Low	(Long)	0.4	(NA)	0.2	5	5	2	60	Avail
(High)	High	(Long)	0.2	(NA)	0.2	5	5	2	60	Avail
(NA)	Ideal	(Short)	0.2	(Low)	0.07	15	15	3	60	Avail
(NA)	Ideal	(Long)	0.6	(Low)	0.07	15	15	9	60	Avail
(NA)	Ideal	(Short)	0.2	(High)	0.2	5	5	1	60	Avail
(NA)	Ideal	(Long)	0.6	(High)	0.2	5	5	3	60	Avail

NA: data not being identified; Avail: data being observed in the original studies.

4.3.2 Considering the possibility to combine raw data from related studies

Combining raw data of the two original studies.

Since Study I and II have similar designs, and can be combined into one single database (Table 4.3), it is possible to build one model from this database using Bayesian networks. We call the model OroSensory Exposure Network (SEN). The following describes the process of structure specification for this model network.

The eating rate was kept constant so this variable was removed from the SEN model. However, one should note that SEN was only meant for one specific value of eating rate (in the studies it was set to 60 g/min), not for the whole range of this variable. With the knowledge that bite interval and bite size are dependent, we decided to keep 'Bite size' in SEN instead of 'Bite interval' because 'Bite size' contributes to 'Food exposure time', a variable of interest in the original studies. Next, 'Food bite number' was also redundant and thus excluded as this variable can be calculated directly from 'Bite size'. In the end, the SEN model comprised five variables: 'Salt intensity', 'Bite size', 'Bite residence time', 'Food exposure time' (discrete), and 'Intake' (continuous). 'Bite size', 'Bite residence time', and 'Food exposure time' were treated as discrete variables because only few values of each were assessed in the original studies. Figure 4.5 shows the structure of SEN.

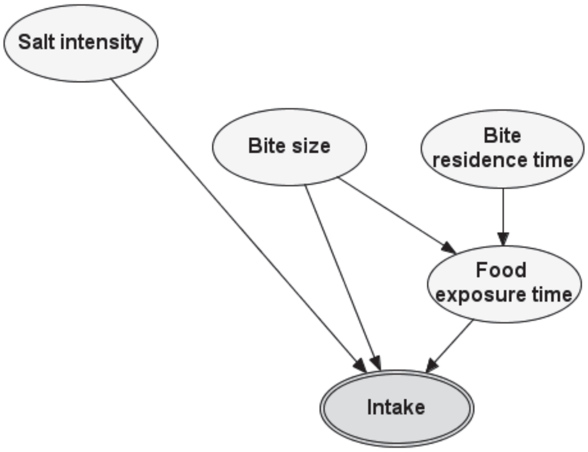


Figure 4.5: Structure of the OroSensory Exposure Network model. The single ovals indicate discrete variables, and the double oval indicates a continuous variable.

The arrows 'Salt intensity' → 'Intake' and 'Food exposure time' → 'Intake' represent the research hypotheses of the original studies. The arrows 'Bite size' → 'Food exposure time' and 'Bite residence time' → 'Food exposure time' could be justified by the deterministic relations: **Food bite number = 1/Bite size** and **Food exposure time = Bite residence time/Bite size**. The direct arrow 'Bite size' → 'Intake' delivers one hypothesis of the current work that 'Food exposure time' does not carry all the information from 'Bite size' towards 'Intake'. The absence of the direct arrow 'Bite residence time' → 'Intake' implies another hypothesis that 'Food exposure time' carries most of the relevant information. There were no arrows connecting 'Salt intensity', 'Bite size' and 'Bite residence time' because these variables were manipulated independently of one another. The current structure of SEN should be adapted if the empirical data do not support the above hypotheses.

Combining raw data from the original studies with those from another related study. Data from another related study (Bolhuis et al., 2012, Study III) was considered for integration with the current combined database (Table 4.3). In this study, 'Salt intensity' (low, ideal, high) was the explanatory variable of 'Intake'. In addition, 'Eating rate', 'Bite size', and 'Bite frequency' could be calculated from observational information. 'Bite frequency' was defined as the mean number of bites per minute and mathematically equivalent to **60/Bite interval** (Phan and Bolhuis et al., submitted for publication). The information on 'Bite residence time' was not observed, and hence 'Food exposure time' was also missing.

Despite having many common variables, data from Study III could not be combined with those from Study I and Study II and analyzed in a meaningful way. The reason was that the experimental design of Study III followed a real-life setting whereas those of Study I and Study II did not. Specifically, the structure of causal relationships among 'Bite interval' (Bite frequency), 'Bite size', and 'Eating rate' in Study III do not match the causal structure of Study I and Study II (Figure 4.4) and hence cannot be combined.

To illustrate that an analysis of the current data combination would be misguided, we describe what would happen if it were to be done. The new database is shown in Appendix 4.A. In constructing the database that is used for any actual analysis, one has to select the variables that are not direct mathematical derivations from each other. That might otherwise cause a redundancy that would lead to numerical and interpretational difficulties during statistical analysis. After the variables have been decided upon, the causal structure is assumed where the selected variables are seen as being primary explanatory variables and independently manipulated. Appendix 4.B shows a network structure suggested for the available variables. In our example, this is the causal structure that is correct for one part of the data (Study III), but not for the other (Study I and Study II). Given the database and network structure, a final network with estimated parameters would predict that a high bite size leads to a lower value of intake than an average bite size does. This outcome would arise because the data from Study I and Study II gives misleading information about the influence of bite size on intake: a rather high bite size (15 g) "causes" an average eating rate (60 g/min) as the bite frequency was artificially kept low. This is combined with the information from the observational data of Study III, which shows a strong positive relationship between eating rate and intake. In combining the three studies, a pattern would emerge from the available data that says: high bite size → average eating rate → average intake. This pattern would be in direct conflict with the empirical data of the original studies (not shown). As such, the model resulting from the combination of the three studies would fail to reflect the true relationships between the variables.

Note that such analysis would be true with any statistical method being applied on this combined dataset (Appendix 4.A). In general, the estimation of parameters in a statistical model relies on the assumption that underlying correlational/causal structure is the same for all data. Bayesian network modeling is just a tool that explicitly reveals causal assumptions with its graphical component.

4.4 Discussion

The two original studies were designed to understand the impact of the food exposure time on the *ad libitum* intake independent from the influence of eating rate. The eating rate, a secondary explanatory variable, was fixed in the experimental designs. We have shown that food exposure time was a secondary variable derived from other primary explanatory variables that were actually manipulated in the experimental setup. It has been also shown that fixing eating rate reversed the causal relationships among bite interval, bite size, and eating rate as compared to more natural settings of eating events. These two design characteristics encountered problems when the results of the individual studies were related with real-life situations (4.4.1). Moreover, this work has provided a better insight into the approach of Global Experimental Design (proposed by Phan et al., 2012), which starts from the real-life situations towards the planning of individual studies (4.4.2).

4.4.1 From individual studies towards real-life situations

The design of the two original studies led to the three following consequences. First, the control (fixing) of a secondary variable makes it impossible to meaningfully combine and analyze the obtained data with those from other related studies having different causal structures. Study I and Study II have a similar design, and their data can be combined to make a larger model network. The model was not larger in term of number of variables, but larger regarding the number of real values or ordinal states of the variables being taken into account. However, making the model larger by adding the data from Study III, a seemingly related study, could not be done as Study III had a different structure. That means data from Study I and Study II cannot be analyzed together with data from any studies where eating rate is not controlled.

Second, strong efforts to generate "explanatory" variables from primary ones often make the obtained results difficult to apply. The causal relationships of food bite number and food exposure time with the *ad libitum* intake were explored in the two original studies. Yet, these relationships may not be straightforwardly intervened in the future. Such studies turn out to be more relevant for basic research, and less so for giving advice to consumers or to the food industry.

Third, derived explanatory variables can be intrinsically dependent. It was the case for food bite number and food exposure time in Study II. Their dependency was induced by the common primary variable bite size. Therefore, it is difficult to interpret their relative importance on the outcome variable *ad libitum* intake.

4.4.2 From real-life situations towards individual studies: Global Experimental Design

We have earlier investigated the use of Bayesian networks to combine raw data from independently performed but related studies (Phan et al., 2012). We have proposed the approach *Global Experimental Design* to avoid the problem of systematic missing data when the related studies were not initially designed for integration.

This approach suggests that the design of related studies should be based on a prior built overall network of the study domain. Gathering more information than needed for the scope of one individual study was the center of that discussion; the design of individual studies was not. The current work has shown clearly that there should be no conflict in causal relationships of the individual designs to enable data combination. This message is inherent within the very first step of building an overall network.

In conclusion, controlling a secondary variable makes it impossible to combine the obtained data with those obtained from related studies following real-life settings. Using a secondary variable as explanatory variable can lead to new mechanistic insights in detailed parts of the study domain. Yet, this approach makes it difficult to apply the findings in real - world problems. Building the overall network within the Global Experimental Design framework is of utmost importance to allow the data integration from related studies. The overall network assists researchers not only to gather the structural linking information (Phan et al., 2012), but also to respect causal relationships when designing individual studies.

4.5 References

Bolhuis, D. P., Lakemond, C. M. M., de Wijk, R. A., Luning, P. A., & de Graaf, C. (2010). Effect of Salt Intensity on Ad Libitum Intake of Tomato Soup Similar in Palatability and on Salt Preference after Consumption. *Chemical Senses, 35*(9), 789-799.

Bolhuis, D. P., Lakemond, C. M. M., de Wijk, R. A., Luning, P. A., & de Graaf, C. (2011). Both Longer Oral Sensory Exposure to and Higher Intensity of Saltiness Decrease Ad Libitum Food Intake in Healthy Normal-Weight Men. *Journal of Nutrition, 141*(12), 2242-2248.

Bolhuis, D. P., Lakemond, C. M. M., de Wijk, R. A., Luning, P. A., & de Graaf, C. (2012). Effect of salt intensity in soup on ad libitum intake and on subsequent food choice. *Appetite, 58*(1), 48-55.

Bolhuis, D. P., Lakemond, C. M. M., de Wijk, R. A., Luning, P. A., & de Graaf, C. Both higher number of bites and longer oral residence duration increase the oro sensory exposure to food and reduce ad libitum food intake. *Submitted for* *publication*.

de Graaf, C. (2012). Texture and satiation: The role of oro-sensory exposure time. *Physiol Behav*, *107*(4):496-501.

Dodge, Y., Cox, D., Commenges, D., Davison, A., Solomon, P. & Wilson, S. (2006). *The Oxford Dictionary of Statistical Terms. Sixth Edition*. New York : Oxford University Press Inc.

Everitt, B.S. & Skrondal, A. (2010). *The Cambridge Dictionary of Statistics. Fourth Edition*. Cambridge: Cambridge University Press.

Heckerman, D. (1995). A tutorial on learning with Bayesian networks. In, *Technical report MSR-TR-95-06*, Microsoft Research.

Kellen, M. (2010). *The effect of eating rate on food consumption*. Nova Southeastern University, Florida.

Phan, V. A., Dekker, M., Garczarek, U., & van Boekel, M. A. J. S. (2010). Bayesian networks for food science: theoretical background and potential applications. In S. R. Jaeger & H. MacFie, *Consumer-driven innovation in food and personal care products*. Cambrige: Woodhead Publishing Limited.

Phan, V. A., Ramaekers, M. G., Bolhuis, D. P., Garczarek, U., van Boekel, M. A. J. S., & Dekker, M. (2012). On the use of Bayesian networks to combine raw data from related studies on sensory satiation. *Food Quality and Preference, 26*(1), 119-127.

Weijzen, P. L. G., Smeets, P. A. M., & de Graaf, C. (2009). Sip size of orangeade: effects on intake and sensory-specific satiation. *British Journal of Nutrition, 102*(7), 1091-1097.

Zijlstra, N., Mars, M., de Wijk, R. A., Westerterp-Plantenga, M. S., & de Graaf, C. (2008). The effect of viscosity on ad libitum food intake. *International Journal of Obesity, 32*(4), 676-683.

Zijlstra, N., Mars, M., Stafleu, A., & de Graaf, C. (2010). The effect of texture differences on satiation in 3 pairs of solid foods. *Appetite, 55(*3), 490-497.

4.6 Appendices

Appendix 4.A

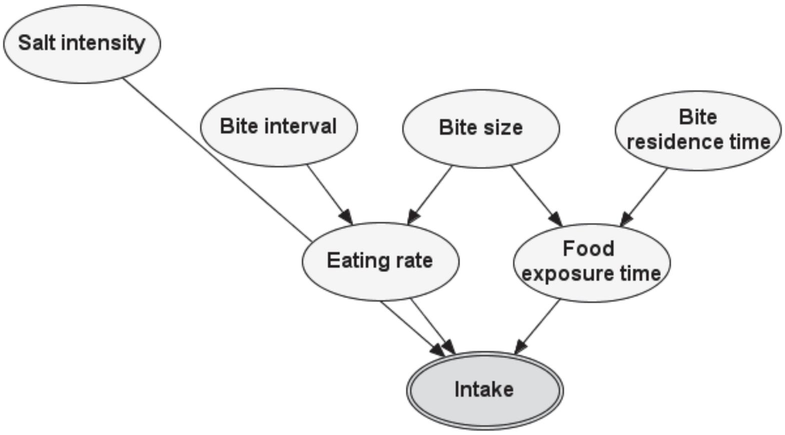
Database resulted from combining raw data from two original studies (Study I and Study II) with those from another related study (Study III). This database does not contain 'Bite number' and 'Bite frequency', for they were represented by 'Bite size' and 'Bite interval', respectively.

Study	Salt intensity	Food exposure time (s/g)	Bite interval (s)	Bite size (g)	Bite residence time (s)	Eating rate (g/ min)	Intake (g)
I	Low	0.2	15	15	3	60	Avail
	High	0.2	15	15	3	60	Avail
	Low	0.4	5	5	2	60	Avail
	High	0.4	5	5	2	60	Avail
II	Ideal	0.2	15	15	3	60	Avail
	Ideal	0.6	15	15	9	60	Avail
	Ideal	0.2	5	5	1	60	Avail
	Ideal	0.6	5	5	3	60	Avail
III	Low	NA	Avail	Avail	NA	Avail	Avail
	Ideal	NA	Avail	Avail	NA	Avail	Avail
	High	NA	Avail	Avail	NA	Avail	Avail

NA: data not available; Avail: data observed (available).

Appendix 4.B

A structure network. This network, which connects all the variable in the database describe in Appendix 4.A, was drawn according to deterministic relations, domain or expert knowledge, and research hypotheses of the original studies.



CHAPTER 5

Generating new insights into the effect of eating behavior and sensory aspects on *ad libitum* intake by Bayesian network modeling

Submitted for publication by:

V.A. Phan, D.P. Bolhuis, U. Garczarek, M.A.J.S. van Boekel, M. Dekker.

Abstract

Bayesian networks were applied to re-model the data from a published study investigating how the saltiness intensity of a soup influences the *ad libitum* intake. In the original study, the two observed variables bite size and bite frequency were not included due to many missing values. However, with Bayesian networks it was straightforward to handle these variables. Domain knowledge (e.g. scientific literature) could be exploited to specify some highly plausible causal relationships. More information (considered as new hypotheses) was extracted from the data in comparison to the results that were published earlier. For example, the *ad libitum* intake increased about 28 g (7% -17%) or a 10% increase of bite size, and about 8 g (3% - 4%) or a 10% increase of bite frequency. Bite size explained partly the influence of salt intensity on the *ad libitum* intake. Eating rate explained partly the influence of bite size on the *ad libitum* intake, and this pathway accounted for about 60 % of the effect size. Bayesian networks enable scientists to generate new insights into a research domain as this technique allows a clear visualization of complex problems and a transparent quantification of information flows in the model.

5.1 Introduction

Meal termination or satiation has been a research topic of interest due to the prevalence of obesity in developed countries (Blundell et al., 1988; Blom et al., 2004; de Graaf et al., 2004; Ruijschop et al., 2009; de Graaf, 2012). Meal termination is usually assessed by measuring the *ad libitum* intake, which is defined as the amount of food eaten voluntarily by a subject until pleasantly satiated. Researchers have investigated the effect of various food characteristics on the *ad libitum* intake, such as viscosity (de Wijk et al., 2008; Zijlstra et al., 2008), energy content (Weijzen et al., 2009), aroma (Ruijschop et al., 2008), and taste (Griffioen-Roose et al., 2009; D.P. Bolhuis et al., 2010; Bolhuis et al., 2012). Bite size and oral processing time have also been subject to investigation (Spiegel et al., 1993; Weijzen et al., 2009; Zijlstra et al., 2009; Bolhuis et al., 2011). Some additional variables are often observed in such studies, such as liking, appetite ratings, and eating rate. Apart from testing the main hypotheses, different statistical procedures are used to process information related to these observed variables. This practice is not the optimal way to extract all the useful information, especially when a large number of variables are present. We were interested in applying a method that gives an overview on the interplay of all the variables: Bayesian networks.

Bayesian networks are graphical probabilistic models being widely applied in various fields, but not yet popular in food science (Pourret et al., 2008; Phan et al., 2010). This modeling technique can handle incomplete datasets (Heckerman, 1995). It can also make use of expert knowledge in determining causal relationships. Furthermore, the graphical and probabilistic natures make Bayesian networks suitable to model complex problems and to communicate the model with model-users via inference procedures (Phan et al., 2010).

The present paper re-analyzes the data of a published study on satiation using Bayesian networks. All variables of interest can be included in one single model network, even ones containing a large number of missing values. This work focuses on quantifying the relationships among the variables. The objective of this chapter is to investigate whether and how applying Bayesian networks can provide extra information and improve the communication of the outcomes. Specifically, the capability to make the most use of available information and the power of the inference procedure of Bayesian networks would lead to a better insight into the research theme.

5.2 Materials and method

5.2.1 Description of the original study

The original study investigated how the *ad libitum* intake ('Intake') of a tomato soup is influenced by two variables: order of the food course ('Food course') and intensity of perceived saltiness ('Salt intensity') (Bolhuis et al., 2012). A full crossover design was used. Each subject received all six combinations of two states of 'Food course' being 'first' and 'second' and three states of 'Salt intensity' being 'low', 'ideal', and 'high'. Data from 43 subjects were taken into account in the statistical analysis.

The study consisted of two separate parts: a preliminary tasting and the actual crossover design on intake. First, the preliminary tasting was performed to select three soups of different salt concentrations, corresponding to the three states of 'Salt intensity'. Subjects tasted five soups varying in salt concentration, and rated pleasantness and relative-to-ideal intensity of saltiness as described in Bolhuis et al. (2010). From these data, a set of three soups was chosen for each individual subject. The soup with ideal salt intensity was the most pleasant soup, whereas the soups with low and high salt intensity were less pleasant than the ideal but similar in pleasantness ratings (Figure 5.1).

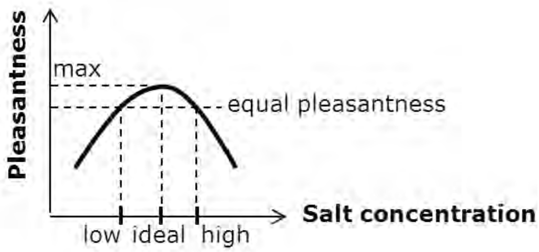


Figure 5.1 : Illustration of the soups with low, ideal, and high salt intensity. The soup with ideal salt intensity is the most pleasant soup. The other two soups are similarly pleasant, though they differ in salt concentration.

In the intake experiment, when the soup was served as the first course, subjects ate the soup *ad libitum* while knowing they could continue voluntarily with buns and different fillings. When the soup was served as the second course, subjects ate a fixed preload of raisin buns before consuming *ad libitum* the soup. Subjects consumed their soup from a self-refilling bowl with a spoon (Wansink et al., 2005; Bolhuis et al., 2010). The total consumption time ('Eating duration') was also recorded along with the *ad libitum* intake. The real-time automatic weighing system allowed the experimenters to read the weight of each bite during the *ad libitum* intake (real-time bite sizes) and total number of bites ('Total bite number'). 'Bite size' was calculated as the mean of the real-time bite sizes for each eating condition. There were 51 random missing data for 'Bite size' and 'Total bite number' due to the instability of the balance that occurred in certain cases. The original paper did not take these two variables into account in the analysis.

5.2.2 Modeling with Bayesian networks

A Bayesian network model consists of two components: structure and parameters. The network structure represents the causal relationships among the variables; the network parameters quantify these relationships through probability expressions (Phan et al., 2010). In this chapter, the model learning was supported with HUGIN Bayesian network software (HUGIN researcher 7.5, http://www.hugin.com). The following presents (1) the formation of the database and (2) the modeling procedure.

Table 5.1 summarizes the database to be used further in building model networks. The variable 'Food course' was concluded not to affect the *ad libitum* intake in the original study. It was therefore not included in order to simplify the model. The variable 'Bite frequency' was introduced as the number of bites eaten per minute (bites/min) and calculated as **Total bite number/Eating duration**.

Table 5.1: Summary of the database. The table represents data obtained from one subject. 'Salt intensity' was controlled with three states: 'low', 'ideal', and 'high'. Data on 'Bite frequency', 'Bite size', 'Eating rate', and 'Intake' were either calculated from the observed information or obtained via direct measurement (represented as 'Avail').

Salt intensity	Bite frequency (bites/min)	Bite size (g)	Eating rate (g/min)	Intake (g)
Low	Avail	Avail	Avail	Avail
Low	-	-	Avail	Avail
Ideal	Avail	Avail	Avail	Avail
Ideal	Avail	Avail	Avail	Avail
High	Avail	Avail	Avail	Avail
High	Avail	Avail	Avail	Avail

The sign, "-", indicates that data is missing. These missing data were found randomly due to technical issues with the balance stability and they accounted for 20% of the observations.

'Eating rate' was introduced as the average amount of soup eaten in one minute (g/min) and calculated as Intake/Eating duration. As 'Bite size' could also be calculated as **Intake/Total bite number**, the following relationship was established:

$$\text{Eating rate} = \text{Intake} * \frac{\text{Bite frequency}}{\text{Bite number}} = \text{Bite size} * \text{Bite frequency} \qquad \textbf{Equation 3.1}$$

Since the information on 'Total bite number' and 'Eating duration' is fully carried by 'Eating rate', 'Bite size', and 'Bite frequency', the former two variables were excluded from the database. Furthermore, one outlier was detected (out of the 6 * Inter Quartile Range) from the records of 'Bite frequency'. This outlier was removed to avoid one data point from having overly large influence on the model and parameter estimates. In summary, the database comprised one discrete variable 'Salt intensity' (controlled) and four continuous variables (observational): 'Bite frequency', 'Bite size', 'Eating rate', and 'Intake'. There were 258 observations in total, including 50 missing values for 'Bite size' and 51 missing values for 'Bite frequency'. The data of the continuous variables needed not being discretized in this chapter because the currently used HUGIN software (HUGIN Researcher 7.5) is capable to deal with continuous data.

The modeling process started with model selection (5.3.1), which made use of both expert knowledge and statistical data. The data-driven judgment of four selected plausible model networks was based on three information criteria: log-likelihood, Akaike information criterion (AIC), and Bayesian information criterion (BIC). In HUGIN software, the model with the highest information scores describes the data best from a statistical point of view. The inferences (5.3.2) were based on the model chosen to be the best. Through this procedure, the influence of salt intensity on bite size and *ad libitum* intake, and the influence of bite size and bite frequency on *ad libitum* intake were examined in detail. Also, the role of eating rate in mediating the effect of bite size on intake was explored by comparing the best model with a model omitting a direct arrow between bite size and intake.

5. 3 Results

5.3.1 Model selection

Four model networks were learned when combining the same set of data (summarized in Table 5.1) with four plausible structures: Net I, Net II, Net III, and Net IV (Figure 5.2). The four network structures differ by their set of arrows implying possible causal relationships. These structures were defined based on deterministic relations, domain or expert knowledge, research hypotheses of the original study, and hypotheses of the current model. The formation of each network is explained as follows.

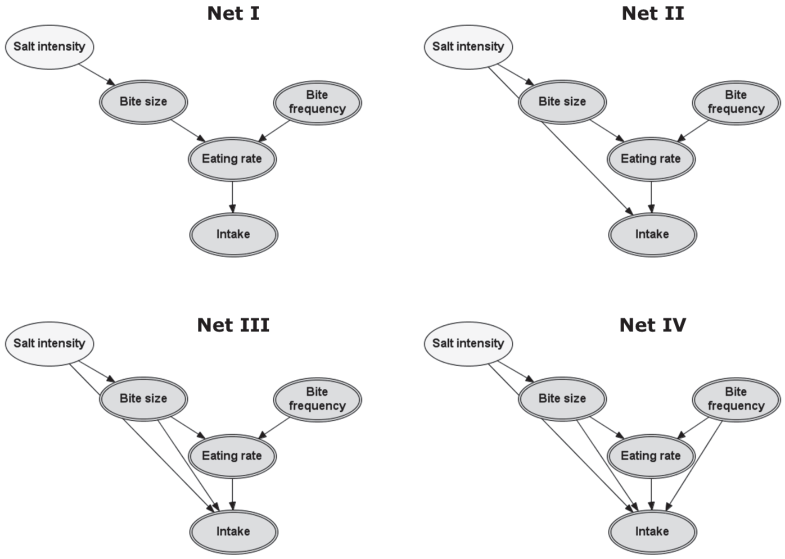


Figure 5.2: Network structure of the four model networks. A single ellipse represents a discrete variable; double ellipses represent continuous variables. Each arrow linking two variables implies a possible causal relationship.

In Net I structure, the arrows from 'Bite size' and 'Bite frequency' towards 'Eating rate' originated from their deterministic relation described in Equation [3.1]. In addition, eating rate is often judged as a determinant of food intake due to their strong correlation reported in the literature (Spiegel et al., 1993; Zijlstra et al., 2010; Viskaal-van Dongen et al., 2011). The causal effect of the eating rate on food intake has been proven also by empirical evidence (Zijlstra et al., 2008; Kellen, 2010). This relationship can be considered as domain or expert knowledge, and it was expressed by the arrow 'Eating rate' → 'Intake'. Furthermore, the perceived salt intensity was hypothesized to affect bite size. We introduced this hypothesis according to the proven effect of aroma intensity on bite size (de Wijk et al., 2009). In this network, 'Bite size' and 'Bite frequency' were assumed to be independent since there was no apparent correlation between these two variables (raw data inspected). Net II carries all the arrows from Net I and contains an additional arrow 'Salt intensity' → 'Intake' that is a direct link of the investigated relationship in the original study. In turn, Net III differed from Net II by the direct pathway from 'Bite size' to 'Intake'. This direct pathway accounts for the extra information from 'Bite size' that 'Eating rate' cannot pass on to 'Intake' via the connection 'Bite size' → 'Eating rate' → 'Intake'. With the same argument, Net IV was added with the direct pathway from 'Bite frequency' to 'Intake' onto the Net III structure. The two arrows from 'Bite size' and 'Bite frequency' towards 'Intake' were seen as the hypotheses of this chapter.

Table 5.2 shows the scores of the information criteria of the four network models. Net IV had the highest log-likelihood score and AIC score; Net I had the highest BIC score. The log-likelihood criterion is always in favor of more complex models, as complexity translates to a higher flexibility in fitting any given dataset. This concept is visualized by the increase in the log-likelihood score when adding one more arrow to the current structure: Net II vs. Net I, Net III vs. Net II, and Net IV vs. Net III. The AIC score takes into account not only the goodness of fit but also the parsimony of the models, i.e. keeping the model simple and avoiding over-fitting of data. This criterion gives a larger penalty for models having more parameters. The BIC score also punishes the model complexity by giving a penalty that is heavier than the AIC.

Table 5.2: Information criteria (IC) of the four model networks. The scores were calculated by HUGIN software. Model with the highest score (formatted bold) represents the best of the statistical data according to the selected criterion.

Network/IC	Log-likelihood	AIC	BIC
Net I	-3974.82	-3991.82	-4022.02
Net II	-3968.88	-3991.88	-4032.74
Net III	-3963.94	-3989.94	-4036.13
Net IV	-3960.47	-3989.47	-4040.99

The BIC and AIC criteria were in favor of different models; BIC tends to prefer Net I while AIC tends to prefer Net IV (the differences between the scores are small). Both AIC and BIC do already penalize model complexity. In the case that after penalizing models of different complexity do have comparable model selection values, the scientist can choose the one that better fits the primary task of the model, e.g. whether that is prediction where one would typically choose the smaller, or gaining new insights where one would typically choose the more complex model. Always choosing the less complex one would mean to double-penalize beyond the scheme of the information criteria, and displays some general mistrust to their capability to fight over-fit. A more complex model would result in more diverse inferences, to examine the relationships of interest. Therefore, Net IV was chosen as the best model and used to perform inferences as shown in the following section. This model network is from now on referred to as Soup Intake Network.

5.3.2 Inferences

In Bayesian networks the step corresponding with hypothesis testing is the structure learning respectively the model selection as described in section 5.3.1. Once that decision is made the model parameters can be learnt and the full model can be used for inference and for quantifying the causal relationships, which corresponds with the interpretation of least square means (also called predictive means) in the more familiar ANOVA setting. Once being fed with the dataset and the pre-defined structure, the HUGIN software automatically calculated the parameters of the Soup Intake Network (information not shown). The initial probability distribution of the network (Figure 5.3) was then calculated based on these parameters (see Phan et al., 2010 for examples). In HUGIN, the continuous variables are presented as normal distributions or a combination of several normal distributions; their probability density function is represented by **sample** mean and **sample** variance. The initial probability distribution represents the model network when no further information on the variables is provided. It acts as the working interface with the model users as well as the base to perform inferences. Performing inference means updating (instantly and automatically) the network probability distribution when certain information on network variable(s) is provided.

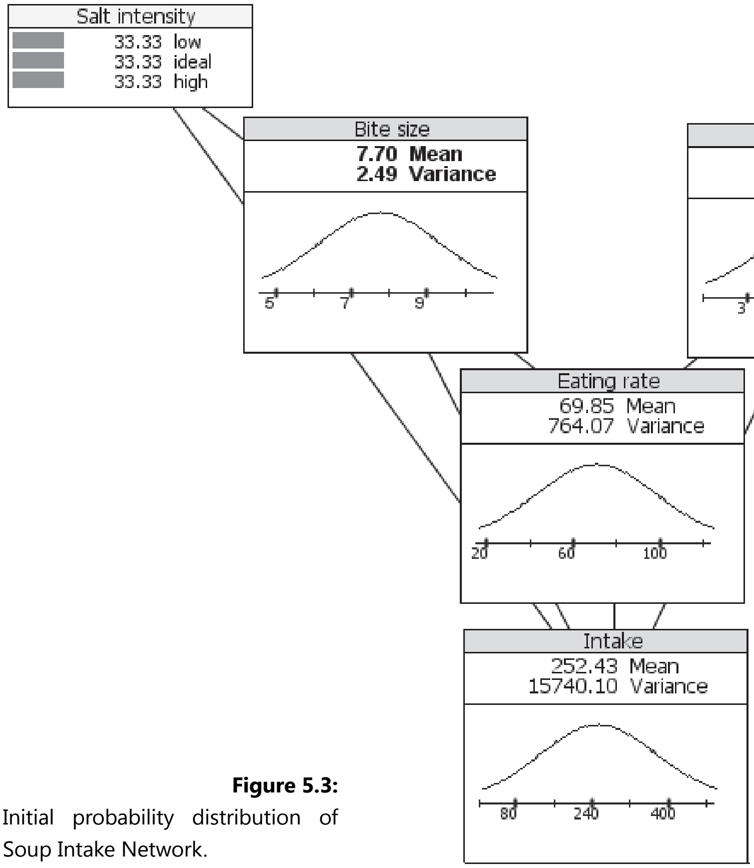


Figure 5.3:
Initial probability distribution of Soup Intake Network.

Influence of salt intensity on bite size and *ad libitum* intake. Figure 5.4 demonstrates the inferences performed when setting evidence on 'Salt intensity' in Soup Intake Network. The soup having ideal salt intensity tended to produce the highest bite size (7.80 g); whereas the soup with high salt intensity led to a smaller bite size than the soup with low salt intensity did (7.51 g vs. 7.78 g). The same pattern was observed for *ad libitum* intake. The highest intake was inferred when the soup had an ideal salt intensity (271 g). A lower *ad libitum* intake was obtained for the soup with high salt intensity compared to the soup with low salt intensity (236 g vs. 250 g), the difference being about 6%.

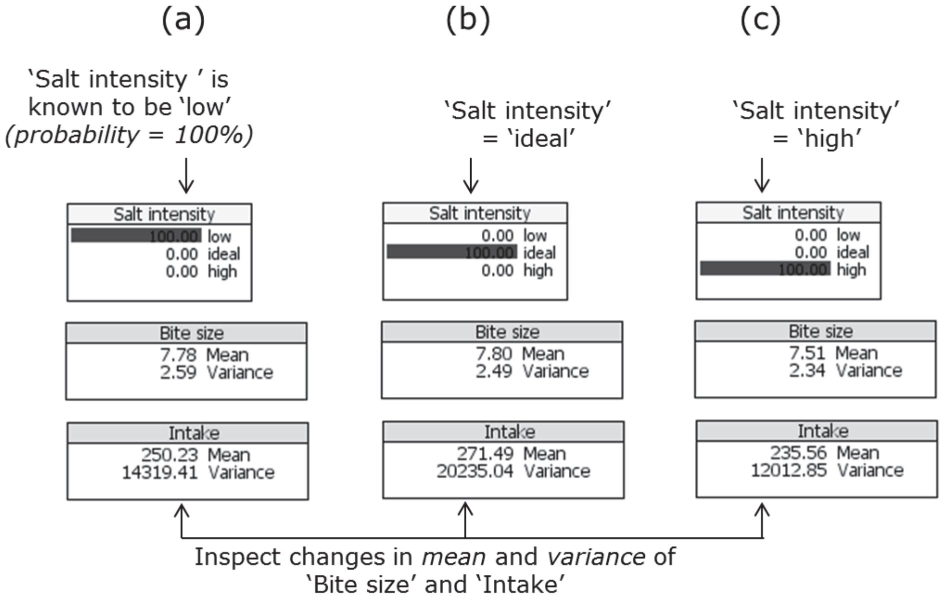


Figure 5.4: Inferences on 'Salt intensity'. Probability distribution of Soup Intake Network (also Net IV in Figure 5.2, only part of the network is shown) was calculated when evidence is provided for 'Salt intensity': (a) 'low', (b) 'ideal', and (c) 'high'. Inspecting changes in mean and variance of 'Bite size' and 'Intake' over the three conditions allows us to examine the influence of 'Salt intensity' on these two variables.

The original paper has reported the same pattern for the influence of salt intensity on intake. The ANOVA analysis showed that the intake of the soup having ideal salt intensity is significantly higher than that of the other two soups. Although being equally pleasant, the soup having high salt intensity yields a significantly lower intake than the soup having low salt intensity. This decrease in intake has been reported to be about 7.5% (235 g vs. 254 g). There was a difference in the extent of decrease in intake found by the original work and that calculated by the current work (7.5% vs. 6%). This difference can be explained by the presence of extra variables (bite size, bite frequency, eating rate) in the Soup Intake Network model compared to the simple model (salt intensity and intake) used in the original study.

Influence of bite size and bite frequency on *ad libitum* intake. A series of inferences were performed to record changes in intake with every step of 10% increase in bite size and in bite frequency over their observed range. The 10%-increase step of each variable was calculated to be one-tenth of the range between the mean +/-

3*Standard Deviation (covering more than 99.7% of the observed variation). This step was equal to 0.93 g for 'Bite size' and 2.25 bites/min for 'Bite frequency'. Figure 5.5 shows the inferred mean and standard deviation of intake depending on bite size (5a) and on bite frequency (5b). The series of input values for bite size and bite frequency started with the minimum among their observed values: 4.64 g and 1 bites/min, respectively. The graphs in Figure 5.5 clearly show that *ad libitum* intake increased with a considerably larger extent when increasing bite size than when increasing bite frequency. The increase of intake was calculated to be 28 g when bite size was increased by 10% (the range of intake increase is from 7% to 17%). The increase of intake was only 8 g when bite frequency was increased by 10% (range from 3% to 4%).

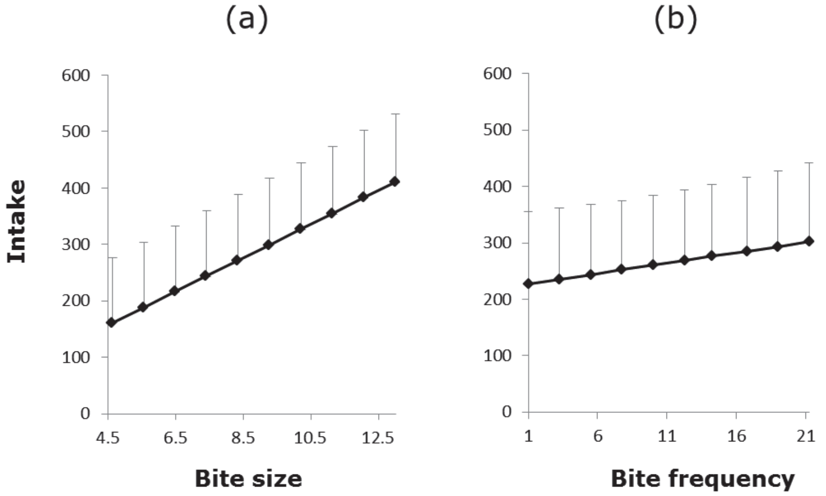


Figure 5.5: Comparing the effect of bite size and bite frequency on intake. The mean and standard deviation of 'Intake' were inferred by Soup Intake Network at every 10% increase of 'Bite size' and 'Bite frequency'.

Role of eating rate in the effect of bite size on intake. We want to quantify the information of bite size carried by eating rate towards intake. Two models Soup Intake Network (also Net IV) and Net II (Figure 5.2) were studied together for this purpose. The direct arrow 'Bite size' → 'Intake' is present in Soup Intake Network but not in Net II. The two networks also differed in the presence of the arrow 'Bite frequency' → 'Intake'. However, this difference did not affect the current inferences because bite size and bite frequency were assumed to be independent.

The relationship between intake and bite size under the condition 'Salt intensity' = 'ideal' inferred by Soup Intake Network was compared with that inferred by Net II (Figure 5.6a). The inferred mean eating rate was also plotted against bite size obtained with both models (Figure 5.6b). In these inferences, the evidence 'Salt intensity' = 'ideal' was set as background to rule out the flow of information from bite size towards intake via salt intensity. It is shown that the changes in eating rate stayed the same with or without the arrow 'Bite size' → 'Intake' (Figure 5.6b). Without that arrow, Net II predicted a higher intake at small bite sizes and lower intake at larger bite sizes than the Soup Intake Network model (Figure 5.6a).

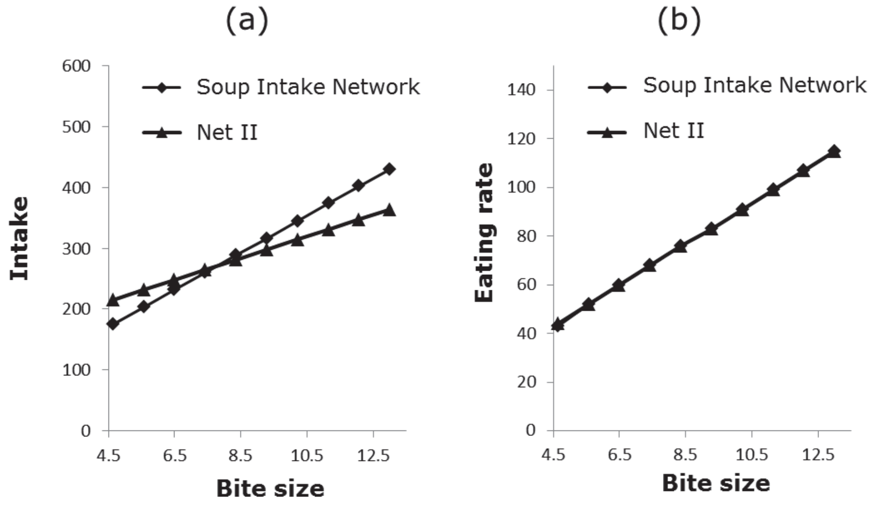


Figure 5.6: Influence of eating rate on the effect size of the relationship between bite size and intake. The increase in mean intake due to increased bite size (at 10% increase) is compared when the arrow 'Bite size' → 'Intake' is present in the model (Soup Intake Network) and when it is not (Net II) (a). The changes in eating rate are also tracked for both cases (b). The variable 'Salt intensity' was fixed at the 'ideal' state for all the inferences above.

Overall, the mean intake increased to a lesser extent when increasing bite size, according to the prediction of Net II. The increase of intake under 10% increase of bite size obtained with Net II was calculated to be about 17 g (increase range from 7% to 16%), while the increase obtained with Soup Intake Network was about 28 g (increase range is from 7% to 17%). Comparing these increases, we can estimate that eating rate carried about 60% (17/28) of the total effect of bite size on intake.

5.4 Discussion

This chapter revisited the dataset of a published study to illustrate the potentials of Bayesian networks in modeling food intake. In this section, we discuss how this modeling technique can make the most use of available information (5.4.1), and how its inference procedure and graphical nature can help to generate new hypotheses (5.4.2) and to visualize and quantify the flow of information (5.4.3).

5.4.1 Making the most use of available information

This chapter has shown that Bayesian networks can use the information on both causal relationships and the statistical data with a large number of missing values. When suggesting a structure for Soup Intake Network, not only the hypothesis being tested ('Salt intensity' → 'Intake') was taken into account, other sources of information supporting causal relationships were also employed. These sources were deterministic relations due to mathematical dependencies ('Bite size' →'Eating rate' and 'Bite frequency' → 'Eating rate'), domain knowledge ('Eating rate' → 'Intake'), and further hypotheses ('Bite size' → Intake' and 'Bite frequency' → 'Intake'). As such, the variables of interest were connected in a systematic and justifiable manner. In addition, the bite size and bite frequency data were included into the model building, even though they contained up to about 20% missing values. This was possible due to the Expectation – Maximization algorithm (Lauritzen, 1995), which is adapted in most Bayesian network software (including HUGIN). The EM algorithm estimates missing values based on the available information; therefore, Bayesian networks can handle missing data directly and competently if the data are missing at random. In contrast, in most classical statistical approaches, one needs either to exclude the observations containing missing values or to do a preceding missing value imputation step. This imputation step is not integrated in the actual analyses, and it is often difficult to decide which is the best or most appropriate procedure among the multitude of available possibilities. This drawback explains why the original work chose not to take into account the data on bite size and bite frequency.

5.4.2 Generating new hypotheses

The inferences in Bayesian networks allow us to predict the outcome variable (or any other variables) given the information on one or more variables of interest. Through this procedure, we found qualitatively the same relationship between salt intensity and *ad libitum* intake as reported in the original paper, but also gained further information. Since the current model has not been validated with new data, the information obtained in addition to the findings published in the original paper is considered as new hypotheses.

First, the new hypotheses concern the effect of salt intensity on bite size. According to the Soup Intake Network model, ideal salt intensity soup leads to people eating with the largest bite sizes. This inferred observation is in line with the finding on smaller bite sizes for less pleasant foods (DiMeglio & Mattes, 2000). It was also shown that higher salt intensity soup results in smaller bite size compared to lower salt intensity soup of similar pleasantness. This result is supported with empirical data obtained by another study using similar test soups (Bolhuis et al., 2011). In this study, bite size has been proven to be significantly smaller for the soup with high salt intensity compared to the soup with low salt intensity during the first half of the soup consumption. Consistently, higher aroma intensities result in smaller bite sizes as de Wijk et al. (2009) concluded for custard desserts, however, one has to note that the pleasantness was not matched for these products. Combining all the above observations, we can illustrate the effect of salt intensity on bite size by Figure 5.7.

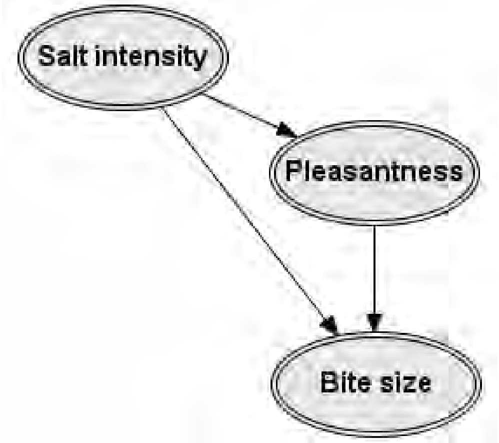


Figure 5.7:
Possible explanation of the effect of salt intensity on bite size. Pleasantness can explain only part of the effect of the salt intensity on bite size.

Salt intensity was generalized to be continuous ratings of perceived saltiness. The effect of saltiness on pleasantness is one pathway that explains the influence of salt intensity on bite size. However, the effect is still observable after ruling out the role of pleasantness. This fact allowed us to draw the direct arrow 'Salt intensity' → 'Bite size'.

Second, the new hypotheses also concern the effect of bite size and bite frequency on *ad libitum* intake. In the original study, the data on bite size and bite frequency were not used, therefore the information related to these variables was absent. The Soup Intake Network model showed that increasing bite size or bite frequency results in higher soup intake. In addition, it was inferred that the effect of bite size on intake is about three times stronger than that of bite frequency. Larger bite size has been reported to increase intake of milk-based chocolate custard (Zijlstra et al., 2009) and that of orangeade (Weijzen et al., 2009). However, to the authors' knowledge, the estimation of the differences in the effect size of bite size and bite frequency on intake has never been done before. This new idea would be interesting to test in future experiments.

5.4.3 Visualizing and quantifying the flow of information

The concept concerning the flow of information can already be perceived via the suggested explanation of the effect of salt intensity on bite size (Figure 5.7). The following elaborates on this concept and its usefulness in extracting and representing information due to the power of Bayesian networks. The discussion is based on the role of eating rate in the effect of bite size on intake.

When setting evidence on bite size in Soup Intake Network (Net IV), the information from bite size towards intake can flow via three pathways: direct, via eating rate, and via salt intensity. When setting evidence on bite size in Net II, the information from bite size towards intake can flow via two pathways: via eating rate, and via salt intensity (Figure 5.2). The background information 'Salt intensity' = 'ideal' allowed us to rule out the via salt intensity pathway. Consequently, the flow of information from bite size towards intake was allowed only via eating rate in Net II, and was possible with both direct and via eating rate pathways in Soup Intake Network.

Without the presence of the arrow 'Bite size' → 'Intake', the effect of bite size on intake was reduced. In other words, the extreme values of predicted intake were pulled closer towards the average, or the prediction lost a certain amount of sharpness. Obviously, eating rate can transfer only part of the information from bite size towards intake as illustrate in Figure 5.8. On the basis of effect size, the transferrable information was about 60%.

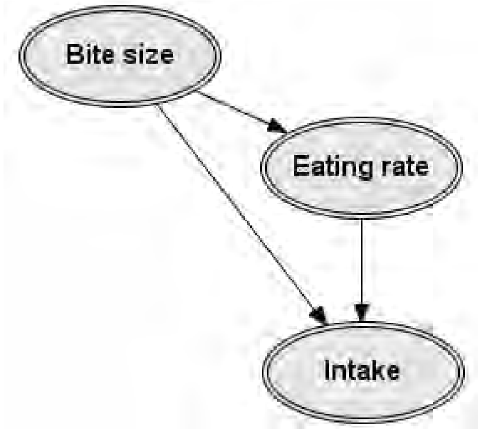


Figure 5.8:
Possible explanation of the effect of bite size on intake. Eating rate can transfer part of the information from bite size towards intake, which accounts for about 60% of the effect size.

In literature, researchers recognize eating rate as a possible explanation for the influence of some food or consumer characteristics on the observed intake (Spiegel et al., 1993; Andrade et al., 2008; Zijlstra et al., 2010; Viskaal-van Dongen et al., 2011). Yet, this explanation has been only communicated at the level of arguments. Bayesian networks, in contrast, enables the explanation to be visualized with the graphical representation of causal relationships and quantified with the inference process. Quantifying the flow of information is beneficial in any situation where a variable is believed to be a mediating factor of a causal relationship.

5.5 Conclusions

As inputs to build a Bayesian network, scientists can make the most use of the available information, from domain knowledge to the current statistical data and even include missing values. As outputs, there are a graph representing dependency among variables of interest and a set of parameters quantifying these dependencies. The graph allows the model-users to quickly grasp the relationships. It is of most value when the model comprises a large number of variables. The parameters support the inference procedure, which allows the model-users to examine the any relationships. With this procedure, the effect sizes can be estimated and compared in a transparent manner. Not to mention, the flow of information from an explanatory variable to an outcome variable can be quantified to estimate the weight carried by an intermediate variable if presents.

By applying Bayesian networks to remodel a previously available dataset (Bolhuis et al., 2012), we could replicate quantitatively the results communicated in the original paper. Furthermore, the current model network also predicted that (1) the *ad libitum* intake increases about 28 g (7% - 11%) at 10% increase of bite size and about 8 g (3% - 4%) at 10% increase of bite frequency, (2) bite size explains partly the influence of salt intensity on the *ad libitum* intake, (3) eating rate explains partly the effect of bite size on the *ad libitum* intake, and this pathway accounts for 60% on the basis of effect size. Since the current model network has not yet been validated with new data, these predictions can be seen as new hypotheses.

To conclude, Bayesian networks enable scientists to generate new insights into a research domain. This technique allows a clear visualization of complex problems and a transparent quantification of information flows in the model.

5.6 Acknowledgements

We would like to thank Dr. Anders L. Madsen, Chief Executive Officer at HUGIN EXPERT A/S, for giving us technical advices.

5.7 References

Andrade, A. M., Greene, G. W., & Melanson, K. J. (2008). Eating slowly led to decreases in energy intake within meals in healthy women. *Journal of the American Dietetic Association, 108*(7), 1186-1191.

Blom, W., de Graaf, C., Smeets, P., Stafleu, A., & Hendriks, H. (2004). Biomarkers of satiation and satiety: A review. *International Journal of Obesity, 28*, S214-S214.

Blundell, J. E., Hill, A. J., & Rogers, P. J. (1988). Hunger and the Satiety Cascade - Their Importance for Food Acceptance in the Late 20th-Century. *Food Acceptability*, 233-250.

Bolhuis, D. P., Lakemond, C. M. M., de Wijk, R. A., Luning, P. A., & de Graaf, C. (2010). Effect of Salt Intensity on *Ad libitum* Intake of Tomato Soup Similar in Palatability and on Salt Preference after Consumption. *Chemical Senses*.

Bolhuis, D. P., Lakemond, C. M. M., de Wijk, R. A., Luning, P. A., & de Graaf, C. (2010). Effect of Salt Intensity on *Ad Libitum* Intake of Tomato Soup Similar in Palatability and on Salt Preference after Consumption. *Chemical Senses*, 35(9), 789-799.

Bolhuis, D. P., Lakemond, C. M. M., de Wijk, R. A., Luning, P. A., & de Graaf, C. (2011). Both Longer Oral Sensory Exposure to and Higher Intensity of Saltiness Decrease *Ad Libitum* Food Intake in Healthy Normal-Weight Men. *Journal of Nutrition, 141*(12), 2242-2248.

Bolhuis, D. P., Lakemond, C. M. M., de Wijk, R. A., Luning, P. A., & de Graaf, C. (2012). Effect of salt intensity in soup on ad libitum intake and on subsequent food choice. *Appetite, 58*(1), 48-55.

Bolhuis, D. P., Lakemond, C. M. M., de Wijk, R. A., Luning, P. A., & de Graaf, C. (2011). Both higher number of bites and longer oral residence duration increase the oro sensory exposure to food and reduce ad libitum food intake. *Submitted for publication*.

de Graaf, C. (2012). Texture and satiation: The role of oro-sensory exposure time. *Physiol Behav*, *107*(4):496-501.

de Graaf, C., Blom, W. A. M., Smeets, P. A. M., Stafleu, A., & Hendriks, H. F. J. (2004). Biomarkers of satiation and satiety. *American Journal of Clinical Nutrition, 79*(6), 946-961.

de Wijk, R. A., Polet, I. A., & Bult, J. H. F. (2009). Bitesize is Affected by Food Aroma Presented at Sub- or Peri Threshold Concentrations. *Chemical Senses, 34*(7), A38-A38.

de Wijk, R. A., Zijlstra, N., Mars, M., de Graaf, C., & Prinz, J. F. (2008). The effects of food viscosity on bite size, bite effort and food intake. *Physiol Behav, 95*(3), 527-532.

DiMeglio, D. P., & Mattes, R. D. (2000). Liquid versus solid carbohydrate: effects on food intake and body weight. *Int J Obes Relat Metab Disord, 24*(6), 794-800.

Griffioen-Roose, S., Mars, M., Finlayson, G., Blundell, J. E., & de Graaf, C. (2009). Satiation Due to Equally Palatable Sweet and Savory Meals Does Not Differ in Normal Weight Young Adults. *Journal of Nutrition, 139*(11), 2093-2098.

Heckerman, D. (1995). A tutorial on learning with Bayesian networks. In, Technical report MSR-TR-95-06, *Microsoft Research*.

Kellen, M. (2010). *The effect of eating rate on food consumption*. Nova Southeastern

University, Florida.

Lauritzen, S. L. (1995). The EM algorithm for graphical association models with missing data. *Computational Statistics & Data Analysis*(19), 191-201.

Phan, V. A., Dekker, M., Garczarek, U., & van Boekel, M. A. J. S. (2010). Bayesian networks for food science: theoretical background and potential applications. In S. R. Jaeger & H. MacFie, *Consumer-driven innovation in food and personal care products*. Cambrige: Woodhead Publishing Limited.

Pourret, O., Naïm, P., & Marcot, B. (2008). *Bayesian Networks: A Practical Guide to Applications*: Wiley.

Ruijschop, R. M. A. J., Boelrijk, A. E. M., de Graaf, C., & Westerterp-Plantenga, M. S. (2009). Retronasal Aroma Release and Satiation: a Review. *Journal of Agricultural and Food Chemistry, 57*(21), 9888-9894.

Ruijschop, R. M. A. J., Boelrijk, A. E. M., de Ru, J. A., de Graaf, C., & Westerterp-Plantenga, M. S. (2008). Effects of retro-nasal aroma release on satiation. *British Journal of Nutrition, 99*(5), 1140-1148.

Spiegel, T. A., Kaplan, J. M., Tomassini, A., & Stellar, E. (1993). Bite Size, Ingestion Rate, and Meal Size in Lean and Obese Women. *Appetite, 21*(2), 131-145.

Viskaal-van Dongen, M., Kok, F. J., & de Graaf, C. (2011). Eating rate of commonly consumed foods promotes food and energy intake. *Appetite, 56*(1), 25-31.

Wansink, B., Painter, J. E., & North, J. (2005). Bottomless bowls: Why visual cues of portion size may influence intake. *Obesity Research, 13*(1), 93-100.

Weijzen, P. L. G., Smeets, P. A. M., & de Graaf, C. (2009). Sip size of orangeade: effects on intake and sensory-specific satiation. *British Journal of Nutrition, 102*(7), 1091-1097.

Zijlstra, N., de Wijk, R. A., Mars, M., Stafleu, A., & de Graaf, C. (2009). Effect of bite size and oral processing time of a semisolid food on satiation. *American Journal of Clinical Nutrition, 90*(2), 269-275.

Zijlstra, N., Mars, M., de Wijk, R. A., Westerterp-Plantenga, M. S., & de Graaf, C. (2008). The effect of viscosity on ad libitum food intake. *International Journal of Obesity, 32*(4), 676-683.

Zijlstra, N., Mars, M., Stafleu, A., & de Graaf, C. (2010). The effect of texture differences on satiation in 3 pairs of solid foods. *Appetite, 55*(3), 490-497.

CHAPTER 6

General discussion

This general discussion consists of two parts. The first part (6.1) synthetizes the outcomes of Chapter 3 and Chapter 4, which leads to the development of Global Experimental Design approach. The second part (6.2) gives a conclusion on the whole of the thesis and discusses the outlook of this thesis. The approach Global Experimental Design is represented separately and thoroughly because it stands out as an important product or message of this thesis.

6.1. Towards Global Experimental Design in food science using Bayesian networks

6.1.1 The need to combine data from related studies

To understand the underlying mechanisms of complex phenomena in real life, scientists of different disciplines approach it from different perspectives. Within each perspective, a reductionist approach is adopted, where the problem is typically broken down into sub-problems that answer specific questions. It is usually tacitly assumed that the reverse process is possible, i.e. integrating the obtained information on the sub-problems to rebuild the complex phenomenon. This process is unfortunately difficult to achieve as independent scientific studies, despite having similar objectives, are in general not designed in a way that allows their data to be combined. The ability to combine raw data from various studies would be a big advantage in answering real-life problems. It is investigated here what would be needed to develop such an approach, proposed as the Global Experimental Design.

Previous work on modeling sensory satiation using Bayesian networks demonstrated a need for Global Experimental Design (Phan et al., 2012; Phan and Garczarek et al., submitted for publication). In this work, data was used from studies that were independently designed and conducted to investigate the impact of different sensory aspects on *ad libitum* food intake (definition of all variables mentioned in this chapter can be found in Appendix 6.A). We faced two hurdles that need to be overcome if the goal to combine data in a meaningful way is to be achieved. Hurdle One is a lack of Structural Linking Information (Phan et al., 2012); Hurdle Two is a conflict in causal relationships underlying the experimental designs (Phan and Garczarek et al., submitted for publication). To avoid these hurdles, specific actions need to be taken from the very beginning when designing the experiments. The objective of the present paper is to describe the Global Experimental Design approach and to demonstrate its importance in the process of integrating data from independent but related studies.

6.1.2 Hurdle One: lack of Structural Linking Information

The first effort in combining data was based on two controlled studies that investigated how different sensory aspects influence *ad libitum* food intake of a tomato soup. First, the Aroma study investigated the effect of aroma concentration and aroma duration (Ramaekers et al., submitted for publication). Second, the Taste study investigated the effect of perceived intensity of saltiness (salt intensity) by manipulating the salt concentration (Bolhuis et al., 2010).

In order to understand the assumed underlying causal relationships in those experiments, Bayesian networks were used (see Chapter 4). In short, Bayesian networks are graphical probabilistic models consisting of two components: structure (graph) and parameters (probabilities) (Heckerman, 1995). The network structure represents the causal relationships among the variables; the network parameters quantify these relationships through probability expressions (Phan et al., 2010). The graphical nature of Bayesian networks makes it easy to communicate and comprehend the overall picture of a research domain, even if a large number of variables are involved.

The graphical representation is shown in Figure 6.1a for the Aroma Study and in Figure 6.1b for the Taste Study. It is rather straightforward to draw the structure (causal dependencies among the variables) of the combined network from the two single networks (Figure 6.1c).

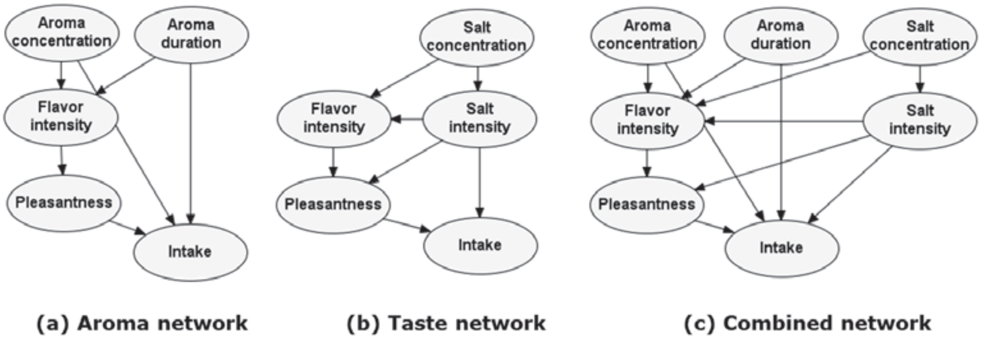


Figure 6.1: A combined network from single networks. Aroma network (a) represents the design of the Aroma study. Taste network (b) represents the design of the Taste study. The combined network (c) was drawn taking into account the specified causations. This figure was adapted from Phan et al. (2012).

The causal relationships (i.e. arrows) were based on the domain knowledge (literature) and the research hypotheses of the original studies. Measurements on variables beyond those that were addressing the main objectives of the studies were also included, namely ratings on the overall flavor intensity and pleasantness. However, the challenge is in the combination of the two independent datasets, to integrate them as one combined database. In this combined database, there was no information on salt aspects in the Aroma study and no information of aroma aspects in the Taste study. This missing information was deemed to be necessary to allow a reliable estimation of model parameters. Such information was therefore called Structural Linking Information.

The problem that Structural Linking Information is lacking occurs when studies are designed to answer their own specific research questions without the intention of combining data with related studies. Some variables might be relevant for one question but not for the other, and hence are typically not measured and/or not documented. To rectify this problem, the missing information needs to be derived by different means - by making use of all available information or performing extra experiments. Some information can directly be derived from the original studies. In the current case study, the value of salt concentration in the Aroma study was calculated from the ingredient information of the soups used. Some information can also be logically deduced through reasoning on the experimental designs. For instance, a state (level) on the aroma concentration and aroma duration could be assigned for the observations of the Taste study. Such decisions can be justified by consensus among the experimenters and modelers. In some cases, performing extra experiments is required. This was needed to obtain the individual ratings of salt perception for the observations of the Aroma study. Given the Structural Linking Information, the parameters (i.e. conditional probabilities) of the combined network can be learned (estimated) from the combined database. This combined network provides a broader view over the problem of interest in terms of dependency among the variables. It also allows the examination of possible combined effects of the variables that have been manipulated in the individual experiments (e.g. aroma duration and salt concentration). Such extra information cannot be obtained if the two sets of data are analyzed separately.

It is, however, not always possible to provide Structural Linking Information, once the individual studies have been completed. The costs of extra experiments can be prohibitive or it may not be possible to attain consensus among experts while creating new states for a variable. It is shown below in section 6.4 that the approach Global Experimental Design enables researchers to circumvent this problem.

6.1.3 Hurdle Two: conflict in underlying causal relationships

The second effort in combining data was based on two other controlled studies investigating the *ad libitum* intake of a tomato soup (Bolhuis et al., 2011; Bolhuis et al., submitted for publication). Both studies used food exposure time and a second, different, variable as explanatory variables of the intake. In order to rule out the influence of eating rate on the *ad libitum* intake, the experimenters kept the eating rate constant in all experimental conditions of these studies. Bite interval, bite size, and bite residence time were manipulated to obtain the desired conditions. Phan and Garczarek et al. (submitted for publication) have analyzed these experimental designs to judge the possibility to combine their data with other related studies. This work used only the graphical component of Bayesian networks, i.e. the representation of causal relationships.

Eating rate is a derived variable of bite interval and bite size because of the deterministic relationship: **Eating rate = Bite size*60/Bite interval** (see Appendix 6.A). Translated into graphs, bite interval and bite size are represented as the causes and eating rate as the effect (Figure 6.2a). The causes are seen as interventional causes and are the ones that can be directly manipulated in the experiments, and hence makes them easy to translate into advice or actions in real life. For example, to reduce the meal intake, it is advisable for a consumer to take a rather small bite (Weijzen et al., 2009; Zijlstra et al., 2009), and food companies may consider the products of small portion sizes or serving sizes (Rolls et al., 2002; Ledikwe et al., 2005; Flood et al., 2006). The causal relationships among bite interval, bite size, and eating rate as shown in Figure 6.2a represent well their relationships in the studies and also reflect real-life situations. Examples of such natural settings can be either those three variables, which are all observational, or eating rate being observed while manipulating bite size and/or bite interval.

However, eating rate was fixed in the two original studies and this fact reversed the natural causal relationships (Figure 6.2b). Bite interval and bite size were manipulated together to keep the eating rate constant. As such, the two primary variables have become dependent on each other in these settings. The eating rate then appears to be the cause, and bite interval and bite size appear to be the effects. Phan and Garczarek et al. (submitted for publication) have shown that it is possible to combine raw data of the two original studies described above, for the structures (causal relationships) underlying their experimental design were similar. However, it is impossible to combine their data with those from related studies that reflect real-life situations, such as in the study by Bolhuis et al. (2012), as depicted in Figure 6.2a. This is a conflict in causal relationships and there is no way to rectify it at this stage. That means these designs are useful only in relation to their own scope but not for the general research topic of understanding food intake. This is the background for Hurdle Two. The Global Experimental Design approach is proposed to avoid facing

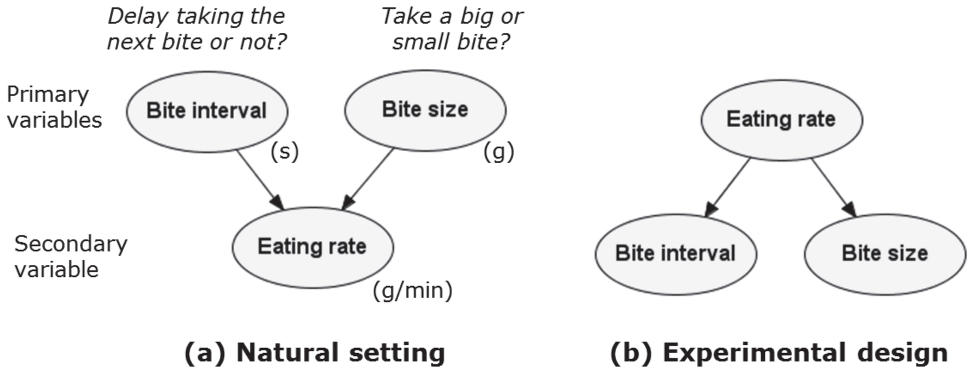


Figure 6.2: Reversal of causal relationships. Figure 2a represents the natural setting in eating. The experimental design (b) that manipulated bite interval and bite size when keeping eating rate constant reverses the causal relationships among these variables found in the natural setting.

6.1.4 Solution: Global Experimental Design

Global Experimental Design is proposed to give guidance on how to design independent but related studies to allow the integration of their raw data at a later stage. This approach involves two main steps. The first step is to build an overall network structure. The second step is to design individual studies, as before, but now derived from the overall network.

Building an overall network structure

The objective is to identify variables of interest for a specific research problem and then to draw possible causal relationships among them. This task is similar to what scientists need to do before designing a controlled study, but on a much larger scale, i.e. concerning many more variables. Building a network structure on a large scale consists of two main challenges. The first is to select certain variables among many for the network. The second is to suggest the causal relationships among the variables. To overcome these two challenges, knowledge and consensus from domain experts are of utmost importance. One should keep in mind, however, that the first overall network is only a formal suggestion with expert support, and is subject to improvement when more data are available.

To illustrate the approach, an overall network for Food Intake (a model predicting *ad libitum* intake) has been generated with a professionally facilitated workshop gathering more than 20 domain experts. The participants were asked to generate variables contributing to *ad libitum* intake in four areas: physical chemical properties of foods, sensory perception, oral processing, and gut feedback.

The workshop consisted of a group session and a plenary session. In the group session, the experts worked in pairs to write down the most relevant variables belonging to the four areas mentioned; the plenary session was led by the facilitator. All these variables were first shared and selected for each area based on the agreement among the participants. Important remarks and unsolved points were written down to ensure the time line and main focus. This was followed by the specification of the cause-effect relationships connecting the selected variables. To simplify the process, the causal relationships toward the outcome variable (Intake) were prioritized over the relationships among the explanatory variables, i.e. possible interactions between texture, taste, aroma variables.

The decision on causal relationships can be based on deterministic relations, causations confirmed in the scientific literature, or common beliefs and hypotheses of the experts. Such a practice has been clearly described in our previous papers (Phan et al., 2012; Phan and Bolhuis et al., submitted for publication; Phan and Garczarek et al., submitted for publication). Figure 6.3 shows the primary outcome of the workshop just for illustration of the size of the obtained network.

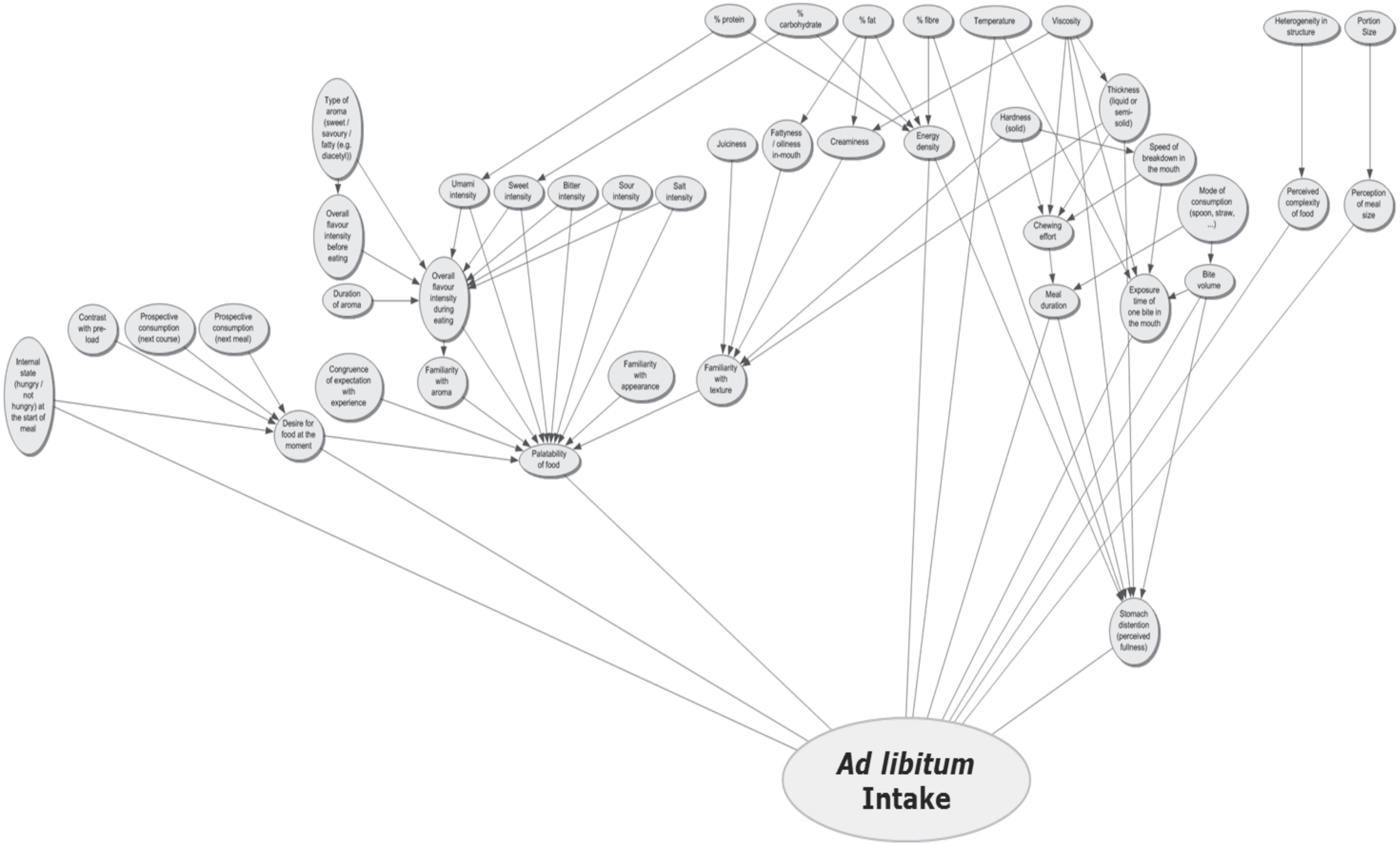

% protein
% carbohydrate
% fat
% fibre
Temperature
Viscosity
Heterogeneity in structure
Portion Size
Type of aroma (sweet / savoury / fatty (e.g. diacetyl))
Thickness (liquid or semi-solid)
Hardness (solid)
Juiciness
Fattyness / oiliness in-mouth
Creaminess
Energy density
Speed of breakdown in the mouth
Umami intensity
Sweet intensity
Bitter intensity
Sour intensity
Salt intensity
Mode of consumption (spoon, straw, ...)
Perceived complexity of food
Perception of meal size
Overall flavour intensity before eating
Chewing effort
Overall flavour intensity during eating
Duration of aroma
Bite volume
Meal duration
Exposure time of one bite in the mouth
Contrast with pre-load
Prospective consumption (next course)
Prospective consumption (next meal)
Internal state (hungry / not hungry) at the start of meal
Congruence of expectation with experience
Familiarity with aroma
Familiarity with appearance
Familiarity with texture
Desire for food at the moment
Palatability of food
Stomach distention (perceived fullness)
Ad libitum Intake

Figure 6.3: Illustration of a Food Intake overall network.

Designing individual studies based on the overall network

The overall network allows the experimenters to see each (future) individual study as a part of a bigger picture. Hence, it can help to avoid not only a conflict in causal relationships (Hurdle Two) but also the problem of missing information (Hurdle One).

First and foremost, the overall network visualizes how the information is assumed to flow among the variables, i.e. from causes to effects, which helps in the formulation of hypotheses. An individual experiment can contribute to the data integration only if the causal relationships underlying its design do not violate those of the overall network. To put it another way, a pair of cause and effect variables found in the overall network must conserve their role in the experimental design. In some cases, this prerequisite is not applicable because a specific design must be followed to answer specific scientific questions. One should be aware that these cases cannot be considered for the data integration. Phan and Garczarek et al. (submitted for publication) have illustrated this situation earlier with the reversal of the causal relationships due to the control of a secondary or derived variable.

Second, the overall network shows explicitly how the variables of primary interest in an individual experiment connect to the rest. The Structural Linking Information can then be recognized and taken into account in the experimental design. That is, each experiment might have to gather more data than necessary for its own scope. For example, subjects participating in the Aroma study could have been asked to rate the perceived salt intensity in the testing sessions (Phan et al., 2012). This information was not of direct value for the hypothesis testing procedure in the Aroma study, but it plays an essential role when combining data from Aroma study with those from Taste study. Moreover, the effort to identify the structural linking information also requires a global consideration on each variable of interest regarding its range of possible values. The variables should be defined and measured in a standardized manner. In this way, the states (levels) of the explanatory variables manipulated in the individual experiments can be handled adequately and judged independently from the experiments. For example, the manner in which Aroma concentration and Aroma duration were defined and manipulated could have been made more explicit in relation with these aspects in the real-life situations (Phan et al., 2012).

6.1.5 Discussion, conclusions, and perspectives

Food problems are complex (van Boekel, 2008). Therefore, integrating data from different studies is highly desired to obtain a more holistic view. The use of Bayesian networks is emerging in food science. For example, food safety has used Bayesian networks to assess microbial risk in along the production chains (Barker et al., 2002; Barker et al., 2005). Another potential application in food safety is foreseen for assessing the performance of different food safety management systems (Sampers et al., 2010; Sampers et al., 2012; Luning et al., 2013). Food product design can also gain benefit from this modeling technique as Corney (2000) has discussed. This is because product design requires various sets of information - physical and chemical properties of foods (instrumental analyses), sensory attributes (sensory panel) and consumer preferences (consumer panel) - which are intricately linked but are difficult to capture in entirety in a single study.

The two hurdles identified in this chapter are not specific to the studies on sensory satiation; they can occur in any research domain when combining raw data from independently performed but related studies. Hurdle One, i.e. missing information, occurs naturally because of the specific or focused needs of individual studies. Hurdle Two is more likely to occur only with sophisticated experimental designs involving the control of secondary variables.

It has been shown that data integration is possible only when special efforts are taken during the early phases of experimental design. The Global Experimental Design approach offers a framework which would guide the designing of individual experiments. Building an overall network is the core of this approach and requires taking the initiative to identify and gather experts. Such an initiative could be feasible within a research group or possibly with larger national research or EU projects where more complex and multidisciplinary themes are involved. The overall network acts as a guideline to avoid conflicts in causal relationships and overlooking of essential linking information.

This is the first time an approach to allow the combination of raw data from related controlled experiments has been communicated. Global Experimental Design could make a new impact on scientific practices in food science and technology as well as in other fields. This approach promotes the sharing or publishing of raw data as well as the standardization in data collection.

6.2 Conclusions and outlook

The objective of this thesis was to explore the use of **Bayesian networks** to **combine raw data** of independently performed but related experiments to **build a quantitative model** of sensory satiation.

The biggest challenge was that no framework has yet been published supporting the data combination from related controlled experiments to build a quantitative model, either using Bayesian networks or any other tools. Meta-analysis is a popular statistical procedure that assists the combination of results from related studies. This procedure has been mainly used in the medical field where the general goal is to get a good estimate of either the effect size of a specific drug or the strength of a risk factor (Sutton & Higgin, 2008). Meta-analysis is typically based on summary characteristics such as effect size, sample size, mean, and variance and thus suffers from the loss of information (among other things). This current thesis aimed at combining raw data to increase the understanding of the combined influences from several factors by providing a quantitative model of sensory satiation. This goes beyond what could be achieved by standard meta-analysis as working with raw data would minimize the loss of information and the bias that might result from working with the summary characteristics.

6.2.1 Reflections on the main outcomes

First, a tutorial on Bayesian networks has been, for the first time, written for the field of food science (Chapter 2). Statistical methods are most often associated with engineering, mathematics, and the medical sciences. As a result, food researchers are forced to use methods that were originally aimed at other disciplines (Pripp, 2013). This could hinder the use of many statistical innovations in the field of food research due to a lack of understanding of these statistical tools. This chapter makes the theoretical background of Bayesian networks accessible to food researchers by gently introducing it through a food example.

Second, two hurdles have been identified in the process of combining data of related studies that are performed independently without the intention of combining their data for a pooled analysis (Chapter 2 and Chapter 3). The first hurdle is a lack of information, which was termed as Structural Linking Information. This hurdle becomes apparent when building the combined data table based on separately obtained datasets. This missing information is typically measured and recorded only in some but not all of the studies as it was only relevant for the specific objective of some studies. Such information appears to be necessary to allow a reliable estimation of parameters of the combined model network. The second hurdle is a possible conflict in causal relationships underlying the experimental design of the individual studies. This hurdle mainly concerns sophisticated designs that involve the control of secondary (or calculated) variables. In such cases, one or more causal relationships found in real-life settings may be reversed. The obtained data are then beneficial only in the framework of those studies, and not for being integrated with the data from other studies having a different causal structure.

Third, the Global Experimental Design approach has been proposed as a potential solution to overcome the identified hurdles (Chapter 6, part I). The core of this approach is to build an overall network structure prior to designing individual related studies. This overall network visualizes all variables of interest for a specific research problem and the causal relationships among them. It provides guidance on which extra information to be gathered and on the causations to be respected when designing individual studies. Moreover, it also assists scientists to design (individual) studies such that the obtained results can be translated directly into actions in real life. In practice, the overall network can be built with the participation of domain experts to maximize the consensus.

Fourth, it has been shown that scientists are able to gain more insights into a research domain when using Bayesian networks (Chapter 5). The graphical component of this modeling technique allows a quick grasp of the flow of information even with a complex problem; its probabilistic component allows a transparent quantification of any information flows of interest through inferences. Such outcomes can be obtained with observational data. These powerful features have been illustrated when comparing the information extracted from the same single dataset by using Bayesian networks versus using classical statistical procedures.

6.2.2 Fulfillment of the objective

Given the limited availability of published applications of Bayesian networks in Food Science, this research has been of an explanatory nature. The use of Bayesian networks to combine raw data of independently performed but related experiments has been explored. A quantitative model of sensory satiation has not yet been attained.

The major outcome of this thesis is the development of the approach of Global Experimental Design using Bayesian networks (Chapter 4). This approach incorporates and facilitates the combination of raw data from related studies. Therefore this approach has the potential to overcome the current limitations in developing an overall integrative model of sensory satiation.

Bayesian networks have been demonstrated throughout the thesis to be a powerful tool in supporting the design of experiments, analyzing data, and communicating the results.

6.2.3 Conclusions

It is possible to combine raw data from related studies for a meaningful analysis if extra effort is made in the phase of experimental design. The approach of Global Experimental Design outlines this phase with the building of an overall network. Using Bayesian networks as an exploratory analysis tool, scientists are able to gain more insights into a research domain.

6.2.4 Outlook

The motivation of this thesis was to combine raw data from the related studies that have already been completed. Yet, its outcomes have drawn our full attention to the phase of experimental design. It was found that in order to combine data from separate studies, it is of critical importance that scientists work together towards the common goal of understanding a particular research theme. In the proposed Global Experimental Design, we have demonstrated that it is possible to form a clear network structure to guide specific actions towards this goal.

Global Experimental Design has been developed from the attempt to use Bayesian networks to model sensory satiation. Yet, the core of this approach – building an overall network – makes use of only the graphical representation of this modeling technique. The graphical representation has indeed been widely used in reasoning and organizing information, e.g. mind mapping (Budd, 2004), or in other modeling techniques, e.g. path modeling (Tenenhaus, 2004) and neural network (Lacy, 1989; Fasel, 2003). Thus, the approach Global Experimental Design can be applied independently from using Bayesian networks to model the obtained combined database.

Scientists have the freedom to use any other statistical procedures to extract the information from the combined database (e.g. obtained with Global Experimental Design). However, this thesis has found much reason to support the use of Bayesian networks in studying complex food problems (Chapter 5). The power of this modeling technique encourages scientists to first fully discover a complex scientific problem with observational studies. Controlled experiments should then be carried out only on relevant relationships to the problem. This practice is indeed in line with the common approach of performing scientific research as illustrated in Figure 6.4.

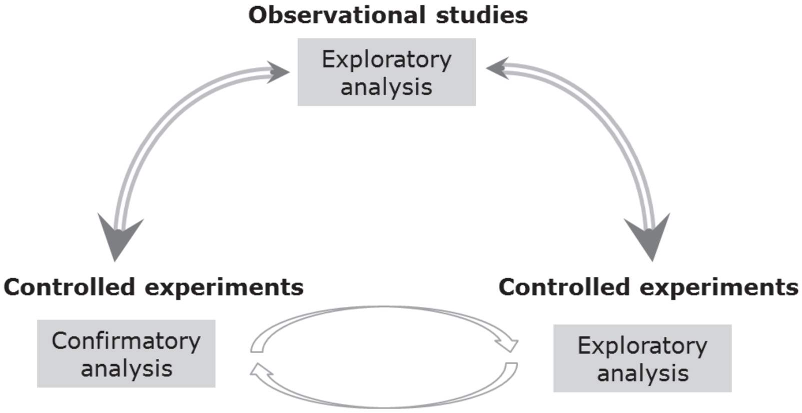


Figure 6.4: Common approach of performing scientific research.

The exploratory analysis on observational studies provides guidance to designing controlled experiments. Data obtained from the controlled experiments are subject to not only confirmatory analysis but also explanatory analysis. The second type of data analysis (explanatory) has received much less attention compared to the first type (confirmatory). The potential use of Bayesian networks emphasizes the possibility to perform exploratory analysis with data from controlled experiments. The information obtained from both types of analysis on data of controlled experiments, in turn, gives feedback to improve the design of future observational studies.

In short, this work has provided new insights and has demonstrated a tool that enable scientists to integrate related information. The approach of Global Experimental Design using Bayesian networks is universal. It can be beneficial not only to satiation studies or other food problems but also to any other fields of research where data integration is of interest.

6.3 References

Barker, G. C., Malakar, P. K., Del Torre, M., Stecchini, M. L., & Peck, M. W. (2005). Probabilistic representation of the exposure of consumers to Clostridium botulinum neurotoxin in a minimally processed potato product. *Int J Food Microbiol, 100*(1-3), 345-357.

Barker, G. C., Talbot, N. L. C., & Peck, M. W. (2002). Risk assessment for Clostridium botulinum: a network approach. *International Biodeterioration & Biodegradation*(50), 167-175.

Bolhuis, D. P., Lakemond, C. M. M., de Wijk, R. A., Luning, P. A., & de Graaf, C. (2010). Effect of Salt Intensity on *Ad Libitum* Intake of Tomato Soup Similar in Palatability and on Salt Preference after Consumption. *Chemical Senses, 35*(9), 789-799.

Bolhuis, D. P., Lakemond, C. M. M., de Wijk, R. A., Luning, P. A., & de Graaf, C. (2011). Both Longer Oral Sensory Exposure to and Higher Intensity of Saltiness Decrease *Ad Libitum* Food Intake in Healthy Normal-Weight Men. *Journal of Nutrition, 141*(12), 2242-2248.

Bolhuis, D. P., Lakemond, C. M. M., de Wijk, R. A., Luning, P. A., & de Graaf, C. (2012). Effect of salt intensity in soup on *ad libitum* intake and on subsequent food choice. *Appetite, 58*(1), 48-55.

Bolhuis, D. P., Lakemond, C. M. M., de Wijk, R. A., Luning, P. A., & de Graaf, C. (2011). Both higher number of bites and longer oral residence duration increase the oro sensory exposure to food and reduce *ad libitum* food intake. *Submitted for publication*.

Budd, J. W. (2004). Mind Maps as classroom exercises. *Journal of Economic Education, 35*(1), 35-46.

Corney, D. P. A. (2000). Designing food with Bayesian Belief Networks. Parmee, I.ed. , *Adaptive computing in design and manufacture*, 83-94.

Fasel, B. (2003). An introduction to bio-inspired artificial neural network architectures. *Acta Neurologica Belgica, 103*(1), 6-12.

Flood, J. E., Roe, L. S., & Rolls, B. J. (2006). The effect of increased beverage portion size on energy intake at a meal. *Journal of the American Dietetic Association, 106*(12), 1984-1990.

Heckerman, D. (1995). A tutorial on learning with Bayesian networks. In, Technical report MSR-TR-95-06, *Microsoft Research*.

Lacy, M. E. (1989). An Introduction to Neural Network Methods and Software. Abstracts of Papers of the American Chemical Society, 197, 26-Comp.

Ledikwe, J. H., Ello-Martin, J. A., & Rolls, B. J. (2005). Portion sizes and the obesity epidemic. *Journal of Nutrition, 135*(4), 905-909.

Luning, P. A., Chinchilla, A. C., Jacxsens, L., Kirezieva, K., & Rovira, J. (2013). Performance of safety management systems in Spanish food service establishments in view of their context characteristics. *Food Control, 30*(1), 331-340.

Phan, V. A., Dekker, M., Garczarek, U., & van Boekel, M. A. J. S. (2010). Bayesian networks for food science: theoretical background and potential applications. In S. R. Jaeger & H. MacFie, *Consumer-driven innovation in food and personal care products. Cambrige*: Woodhead Publishing Limited.

Phan, V. A., Ramaekers, M. G., Bolhuis, D. P., Garczarek, U., van Boekel, M. A. J. S., & Dekker, M. (2012). On the use of Bayesian networks to combine raw data from related studies on sensory satiation. *Food Quality and Preference, 26*(1), 119-127.

Phan, V. A., Bolhuis, D. P., Garczarek, U., van Boekel, M. A. J. S., & Dekker, M. Generating new insights into the effect of eating behavior and sensory aspects on *ad libitum* food intake by Bayesian network modeling. *Submitted for publication.*

Phan, V. A., Garczarek, U., Bolhuis, D. P., van Boekel, M. A. J. S., & Dekker, M. Bayesian networks as a tool to analyze causal relationships in experimental design: a case study on oro-sensory exposure studies. *Submitted for publication*.

Pripp, A. H. (2013). Statistics in Food Science and Nutrition. In: Springer Briefs.

Ramaekers, M. G., Luning, P. A., Ruijschop, R. M. A. J., Lakemond, C. M. M., Bult, J. H. F., Gort, G. Aroma exposure time and aroma concentration in relation to food intake. *Submitted for publication*.

Rolls, B. J., Morris, E. L., & Roe, L. S. (2002). Portion size of food affects energy intake in normal-weight and overweight men and women. *American Journal of Clinical Nutrition, 76*(6), 1207-1213.

Sampers, I., Jacxsens, L., Luning, P. A., Marcelis, W. J., Dumoulin, A., & Uyttendaele,

M. (2010). Performance of Food Safety Management Systems in Poultry Meat Preparation Processing Plants in Relation to Campylobacter spp. Contamination. *Journal of Food Protection, 73*(8), 1447-1457.

Sampers, I., Toyofuku, H., Luning, P. A., Uyttendaele, M., & Jacxsens, L. (2012). Semi-quantitative study to evaluate the performance of a HACCP-based food safety management system in Japanese milk processing plants. *Food Control, 23*(1), 227-233.

Sutton, A. J., & HigginS, J. P. I. (2008). Recent developments in meta-analysis. *Statistics in Medicine, 27*(5), 625-650.

Tenenhaus, M. (2004). PLS regression and PLS path modeling for multiple table analysis. *Compstat 2004: Proceedings in Computational Statistics*, 489-499.

van Boekel, M. A. J. S. (2008). Models and modeling. In, Kinetics modeling of reations in foods.

Weijzen, P. L. G., Smeets, P. A. M., & de Graaf, C. (2009). Sip size of orangeade: effects on intake and sensory-specific satiation. *British Journal of Nutrition, 102*(7), 1091-1097.

Zijlstra, N., de Wijk, R. A., Mars, M., Stafleu, A., & de Graaf, C. (2009). Effect of bite size and oral processing time of a semisolid food on satiation. *American Journal of Clinical Nutrition, 90*(2), 269-275.

6.4 Appendix 6.A

Variable name	Definition	Unit	Relationship
Intake (ad libitum)	Amount of food eaten voluntarily by a subject until pleasantly satiated	g	
Bite size	Weight of each bite	g	
Bite interval	Time interval between the start of two subsequent bites	s	
Bite residence time	Residence time of each bite in the oral cavity	s	
Bite frequency[i]	Number of bites taken per minute	bites/min	= 60 / Bite interval
Eating rate	Average amount of food eaten per minute	g/min	= Bite size * bite interval
Food exposure time[ii]	Average oral residence time calculated for one gram of food.	s/g	= Bite residence time / Bite size

[i] referred to as bite number in the original paper(s); [ii] referred to as oro-exposure time in the original paper(s).

Summary

Background

Food science problems are complex as they involve multiple disciplines and a multitude of factors that are interconnected with one another. Individual studies are not capable of capturing this complexity, but if integrated with related studies, a better representation of the investigated theme can be achieved. Unfortunately, individual studies are usually not designed to allow such integration, and the commonly used statistical methods cannot be used for analyzing integrated data.

The modeling technique of Bayesian networks has gained popularity in many fields of application, but emerged only recently in food science. A Bayesian network model has two components: graphical (structure) and probabilistic (parameters). The network structure represents the causal relationships among the variables of interest; the network parameters quantify these relationships through probability expressions. These features allow scientists working with Bayesian networks to deal with complexity.

Aim

This thesis was part of a larger project where various controlled experiments were independently designed and conducted to understand sensory satiation, i.e. how different sensory aspects influence the amount of food eaten in a meal. The development of satiation during a meal is a highly complex process that involves interaction of the chemical and physical properties of foods, sensory factors, cognitive factors, and also environmental factors. The specific objective of this thesis was to explore the use Bayesian networks to combine raw data of those studies to build a quantitative model of sensory satiation.

Methods

The biggest challenge was that no framework had yet been published on supporting the data combination from related experiments to build a quantitative model, either using Bayesian networks or any other tools.

A tutorial on Bayesian networks was first written to address the field of food science (Chapter 2). This chapter made the theoretical background of this modeling technique accessible to food researchers by gently introducing it through a food example. The available data from various independent studies on sensory satiation were then examined for their potential combination. Finally, the outcomes of Bayesian networks on single data set were compared with those published using classical statistical procedures when analyzing the respective set of data.

Main outcomes

Two hurdles were identified in the process of combining data of related studies that were performed independently without the intention of combining their data. A framework was proposed to avoid these hurdles when designing future related studies to enable the integration of their data. It has also been shown that Bayesian networks allowed scientists to extract more information by generating new hypotheses and to communicate the outcomes in a transparent manner.

The first hurdle was a lack of essential information, which was termed as Structural Linking Information (Chapter 3). This hurdle became apparent when building a combined data table based on separately obtained datasets. The missing information is typically measured and recorded only in some but not all of the studies as it is only relevant for the specific objective of some studies. The Structural Linking Information is necessary for reliable estimations of parameters of the combined model network. It could be obtained by deriving it from existing information or by performing extra experiments; these practices are, however, not always feasible. Given the Structural Linking Information, model parameters of the combined network can be estimated. This allows the examination of possible combined effects of the variables that are independently manipulated in the individual studies. Such effects cannot be studied if the two sets of data are analyzed separately.

The second hurdle was a possible conflict in causal relationships underlying the individual experimental designs (Chapter 4). This hurdle occurred with some experiments that involved the control of secondary explanatory variables. The graphical component of Bayesian networks was used to illustrate that one or more causal relationships found in real-life settings may be reversed in such experiments. The obtained data are thus useful only to the framework of those specific studies, but cannot be integrated with the data from other studies that have different causal structures. The latter effort can cause misleading analyses of the combined dataset.

In addition, the results obtained from such experiments were shown to be difficult to be translated into advice for actions in real-life occasions.

The Global Experimental Designed approach, developed in this thesis, was proposed as a solution to avoid the two hurdles mentioned above (Chapter 6, section 6.1). The core of this approach is to build an overall network structure prior to designing individual related studies. This overall network visualizes all variables of interest for a specific research problem and the causal relationships among them. It provides guidance on which extra information to be gathered and on the causal relationships between variables that is to be respected when designing individual studies. Moreover, it also assists scientists to design (individual) studies such that the obtained results can be applied in real life. In practice, the overall network can be built with the participation of domain experts to reach a consensus.

A justification for the goal to use Bayesian networks as an exploratory analysis tool in food science was presented in Chapter 5. With this technique scientists could make use of the domain knowledge in a transparent manner (e.g. to specify causation) and handle missing data. From the same single dataset, more information (considered as new hypotheses) was extracted with Bayesian networks as compared to classical statistical methods. The graphical component allows a clear visualization of a complex model network, and the probabilistic component allows a clear quantification of information flows in the model. Such a powerful tool could be of great value when working with the combined data from related studies (e.g. supported by Global Experimental Design).

Conclusions

This thesis has yet to attain a quantitative model of sensory satiation. However, it has provided new insights and has demonstrated a tool that would enable scientists to reach that final goal.

It is possible to combine raw data from related studies for a meaningful analysis if effort is made in the phase of experimental design. The approach of Global Experimental Design outlines this phase with the building of an overall network. Using Bayesian networks as a tool for exploratory analysis, scientists are able to gain more insights into a research domain.

Potential applications

The two hurdles identified in this thesis are not specific to the studies on sensory satiation; they can occur in any research domain when combining raw data from independently performed but related studies. As such, the approach of Global Experimental Design is universal and can be beneficial to any research where complex and multidisciplinary themes are involved. This approach promotes the sharing or publishing of raw data, as well as the standardization in data collection.

Samenvatting

Achtergrond

Het wetenschappelijke onderzoek aan levensmiddelen is vaak complex door een veelheid van onderling verweven factoren; om al deze factoren een plaats te kunnen geven vereist de inzet van verscheidene disciplines. Deze complexiteit kan niet goed worden geadresseerd als een bepaald probleem wordt gereduceerd tot één factor. Integratie van meerdere factoren tegelijk zou kunnen helpen om tot een betere benadering van de werkelijkheid te komen. Helaas zijn de meeste één-factor studies niet ontworpen om tot een dergelijke integratie te komen en de gebruikelijke statistische technieken zijn dan ontoereikend.

Modelleren volgens de techniek van Bayesiaanse netwerken wint aan populariteit in verschillende onderzoeksvelden, maar wordt nog nauwelijks toegepast in het levensmiddelenonderzoek. Een Bayesiaans netwerk bestaat uit twee componenten: een grafische deel (dat de structuur weergeeft) en een probabilistisch deel (dat de parameters weergeeft). De netwerk structuur vertegenwoordigt de causale relaties tussen de variabelen die men onderzoekt en de netwerk parameters kwantificeren deze relaties door middel van waarschijnlijkheidsverdelingen. Deze aanpak is in principe geschikt om met complexiteit om te gaan.

Doel van het onderzoek

Het onderzoek beschreven in dit proefschrift maakte deel uit van een groter project over sensorische verzadiging waarin verschillende gecontroleerde experimenten onafhankelijk van elkaar werden opgezet, bijvoorbeeld een onderzoek op het effect van smaak en een ander op et effect van geur. Het doel van het grotere project was om te begrijpen welke sensorische aspecten van invloed zijn op de hoeveelheid gegeten voedsel tijdens een maaltijd. Het optreden van verzadiging is een zeer gecompliceerd proces waarbij chemische en fysische eigenschappen van levensmiddelen interacteren met sensorische, cognitieve en omgevingsfactoren. Het specifieke doel van het hier beschreven onderzoek was om te onderzoeken of het gebruik van Bayesiaanse netwerken het mogelijk maakt om de data uit de verschillende onderzoeken te combineren en aldus een model te bouwen dat sensorische verzadiging kwantitatief beschrijft.

Methoden

De grootste uitdaging was dat er nog geen raamwerk beschreven is in de literatuur om een kwantitatief model te bouwen gebaseerd op individuele maar gerelateerde data, noch met Bayesiaanse netwerken, noch met welke andere techniek dan ook.

Begonnen werd met het beschrijven van een voorbeeld van hoe Bayesiaanse netwerken werken en hoe dit zou kunnen worden toegepast in levensmiddelen onderzoek (Hoofdstuk 2). De theoretische achtergrond werd duidelijk gemaakt in een voorbeeld aangaande een levensmiddel. Vervolgens werden de data die beschikbaar kwamen vanuit verschillende onafhankelijke studies onderzocht op hun mogelijkheid om ze te combineren in één kwantitatief model; daarbij werd de vergelijking gemaakt met resultaten verkregen m.b.v. de klassieke statistiek.

Belangrijkste resultaten

Het combineren van data van gerelateerde studies die onafhankelijk zijn uitgevoerd zonder bedoeling vooraf om ze te combineren leverde twee potentiele hindernissen op, beschreven in Hoofdstuk 3 en 4. De eerste hindernis kwam aan het licht bij het samenstellen van een gecombineerd data bestand gebaseerd op de afzonderlijk verkregen data bestanden. Informatie die verzameld was in de ene studie voor een bepaald doel bleek niet essentieel te zijn geweest voor een andere studie en werd daar dan niet gemeten. Bij het combineren van datasets bleek dat dan achteraf toch essentiële informatie te zijn voor een betrouwbare schatting van parameters van een gecombineerde model netwerk. Deze informatie die nodig is om onafhankelijke studies met elkaar te kunnen verbinden werd getypeerd als Structural Linking Information. Voor een deel kon deze benodigde informatie alsnog boven tafel gehaald worden of te kunnen worden verkregen uit additionele experimenten, maar dat is uiteraard niet altijd mogelijk. Het expliciet maken van Structural Linking Information maakt het wel mogelijk om uit gegevens van afzonderlijke, onafhankelijke studies gecombineerde effecten van variabelen te schatten, iets wat niet mogelijk is uit de afzonderlijke studies. Het werk beschreven in Hoofdstuk 3 was erop gericht een raamwerk te ontwikkelen dat deze eerste hindernis kan vermijden in toekomstige studies. Ook werd aangetoond dat het toepassen van Bayesiaanse netwerken het mogelijk maakt om extra informatie te verkrijgen uit de data, om nieuwe hypotheses te genereren en om de resultaten op een transparante manier te communiceren.

De tweede hindernis is beschreven in Hoofdstuk 4 en bestaat uit een mogelijk conflict in de causale relaties die ten grondslag lagen aan de individuele experimentele ontwerpen. Het ging daarbij om secundaire verklarende variabelen in sommige experimenten. De grafische component van Bayesiaanse netwerken maakte duidelijk dat een of meer causale relaties afgeleid uit de realiteit soms omgedraaid werden in experimenten. Het gevolg daarvan is dat de verkregen data alleen maar gebruikt kunnen worden voor de experimentele setting van dat bepaalde experiment en niet kunnen worden geintegreerd met data van andere studies met andere causale relaties, hetgeen zou resulteren in misleidende informatie als dat toch gebeurt. Ook kan de verkregen informatie niet meer gegeneraliseerd worden naar de realiteit waarvan de experimenten waren afgeleid.

Een motivering voor het gebruik van Bayesiaanse netwerken als een verklarend hulpmiddel voor de analyse van een probleem in het levensmiddelenonderzoek werd gepresenteerd in Hoofdstuk 5. Wetenschappers kunnen gebruik maken van kennis uit een bepaald domein op een transparante manier (bijv. om causale relaties te specificeren) en aldus met missende informatie om gaan. Vergeleken met klassieke statistische methoden bleek er meer informatie uit een dataset verkregen te kunnen worden met Bayesiaanse netwerken, leidend tot nieuwe hypotheses. De grafische component maakt het mogelijk om een complex netwerk overzichtelijk te visualiseren. De probabilistische component maakt het mogelijk om de informatie stroom in een netwerk te kwantificeren. Dit blijkt een krachtig hulpmiddel te zijn om data uit verschillende studies zinvol te kunnen combineren.

In Hoofdstuk 6 wordt een benadering gepresenteerd als Global Experimental Design die onderzoekers in staat stelt om de twee eerder genoemde hindernissen te kunnen vermijden. De kern hiervan is dat een algeheel netwerk structuur wordt voorgesteld voordat individuele studies worden uitgevoerd. Met andere woorden, de individuele studies moeten worden afgeleid uit de algehele netwerk structuur. Dit algehele netwerk visualiseert alle relevante variabelen voor een bepaalde onderzoeksvraag, inclusief de causale relaties tussen de variabelen. Dit geeft richting aan welke informatie echt benodigd is en het laat ook toe om de causale relaties intact te houden voor alle studies. Op die manier kunnen de resultaten uit afzonderlijke studies ook gecombineerd worden om uitspraken te doen over relaties in de realiteit. Een dergelijk algeheel netwerk kan tot stand komen door experts uit het betreffende domein met elkaar tot consensus te laten komen.

Conclusies

Het werk beschreven in dit proefschrift heeft de moeilijkheden in kaart gebracht die ontstaan als geprobeerd wordt om gegevens uit verschillende afzonderlijke studies met elkaar te combineren om tot een algeheel model te komen. Deze moeilijkheden waren van dusdanige aard dat het nog niet mogelijk is om met de bestaande gegevens een kwantitatief model voor sensorische verzadiging te bouwen. Niettemin heeft het proefschrift wel tot inzichten geleid hoe dat in de toekomst bereikt kan worden en er is een protocol/methode Global Experimental Design ontwikkeld dat wetenschappers in staat stelt om dat doel van een algeheel model op basis van een Bayesiaans netwerk te kunnen bereiken.

Mogelijke toepassingen

De twee hindernissen die zijn vastgesteld in dit proefschrift zijn niet typerend voor onderzoek aan sensorische verzadiging alleen. Ze zullen ook voorkomen in andere onderzoeksgebieden waar geprobeerd wordt om data uit verschillende maar samenhangende onderzoeken te combineren. Het concept van Global Experimental Design is universeel en toepasbaar op onderzoeksvragen die een complex terrein bestrijken waar een multidisciplinaire aanpak zinvol is. De voorgestane aanpak is ook relevant voor het delen en publiceren van onbewerkte onderzoek data, en het standaardiseren van het verzamelen van data.

Acknowledgements

Dear supervisors, I have a lifetime to thank you for all.

Tiny, my involvement with Bayesian networks started when you asked "Do you like mathematics?" during my interview for another project. Each time my brain went blank, I was convinced that a person with an appropriate background would do a much better job for the project. In the end, I recognized my contribution: to produce and deliver messages accessible for food scientists (not only for statisticians). My PhD time was, therefore, a journey to find my self-confidence in research, and I succeeded. What I keep in mind is that you were the one who introduced this journey to me.

Matthijs, you stopped me from digging deep into the theory of Bayesian networks and guided me towards their potential in solving food problems. You reminded me constantly of this task. You were the reason why we get close to our targeted readers.

Ursula, I found you in the second year of my PhD after searching at different places for an expert in Bayesian networks. You became the soul of the project since then. More than that, you were also like a friend to me.

Tiny, Matthijs, and Ursula, a thank-you word is not enough for your guidance and support. I am so grateful to have you, a perfect trio-supervisors. I believe that we will have opportunities to work together in future.

Dear members of the Sensory Specific Satiation project, it was not only about the data.

I learned that the project of Bayesian networks had been not approved when standing alone. This fact was also true during the journey of my project.

Maartje, Sanne, Marielle, and **Dieuwerke**, that was a long way but we made it! Maartje, I said to myself when watching you on TV: wow, this lady can make a career as a scientific journalist. Sanne, you were already an independent scientist from the beginning of your PhD. Your confidence, know-how, and skills amazed us all. Marielle, I find it easy to talk to you. You were so sweet to come to tell me that I really made some progress. Dieuwerke, your presence in my WUR experience was so vivid: contribution to 80% of my thesis, office mate, co-organizers of the PhD trip, SSS mother club together with Marielle. Ladies, I would like to thank you for your collaboration and friendship. I treasure the journey that we had together and wish you a lot of success in both personal and professional lives.

Kees, you supported us all the way. I really appreciate that you made a bridge between me and many of your students to access their data. You were always enthousiastic about a sensory satiation model. It has not yet been attained in the end of this project; however, I believe that we are on the right track to get there.

My special thanks to all the participants of the Experts workshop.

One of the achievements in my PhD time was to successfully organize an experts workshop to build a network structure for Food Intake (January 13, 2011 in Wageningen, the Netherlands).

I deeply acknowledge the contribution of all the participants: the professional facilitator Dr. Jan Vaessen (Jan Vaessen facilitator), Emeritus Prof. Jan Kroeze (Utrecht & Wageningen University), Dr. Jeff Brunstrom and Dr. Hal MacFie (University of Bristol), Dr. Ursula Garcrazek, Dr. Liesbeth Zandstra, and Dr. Clair Boucon (Unilever Food and Health Research Institute), Dr. Annette Stafleu (TNO), Dr. Paul Smeets and Maartje Spetter (University Medical Center Utrecht), Prof. Tiny van Boekel, Prof. Kees de Graaf, Dr. Rene de Wijk, Dr. Matthijs Dekker, Dr. Catriona Lakemond, Dieuwerke Bolhuis, Sanne Griffioen-Roose, Pleunie Hogenkamp, Marielle Ramaekers (Wageningen UR).

I would also thank Dr. Jeff Brunstrom and Dr. Hal MacFie for bringing this method to the 9th Pangborn Sensory Science Symposium.

More collaborations

Pascale, my first experiment with Bayesian networks was performed on your data. I have learned a lot from this work. Adrian, your set of data was the most beautiful one I got: four large related experiments were designed with a big picture. Most of the breakthrough ideas for this thesis originated from the analysis of this dataset. Maimunah, I enjoyed analyzing your data and was inspired by the approach that you used for your research. **Mirre, Nicolien, Pleunie**, and **Monica**, you have also shared your data with me, which helped me gain a better picture on the possibilities of Bayesian networks in your field of research.

I am truly thankful to all of you for your kind support and collaboration. I hope to be able to publish some of the works above in future.

More support from modeling people

Anand, you were such an inspiring and positive person. You spent many hours to teach me about Bayesian networks in my first year. **Jimmy**, you did not mind neither to teach me basic knowledge on statistics in my second year. Later, it was so kind of you to send me an email every now and then asking about my work progress. **Anders**, you were always prompt in answering my queries on Bayesian networks with the HUGIN software.

Anand, Jimmy, and Anders, please consider this thesis book as a thank-you present from me!

Wonderful colleagues and working environment

Dear colleagues at Food Quality and Design and Food Physics, I had a great time learning, working, and also enjoying life in the Netherlands with you.

Teresa and **Kristin**, sitting next to each other for a long time, we shared the ups and downs in our PhD lives. I am so happy that you accepted to be the Paranymphs for my public defense ceremony even though you would be very busy with finalizing your thesis at that time.

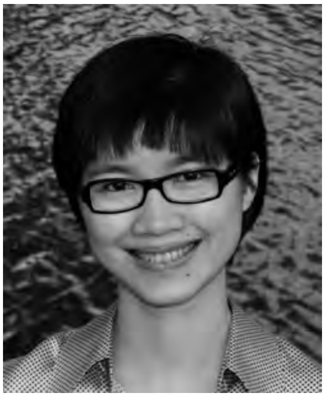
Grace, you influenced me a lot even though we had only few months sharing the office. You went through my whole thesis just within a weekend to help me improve the text. I thank you so much and wish you a lot of success with your PhD project.

Yvonne, Vesna, and **Fre**, I truly appreciated my two-year activities being a chair-woman of the VLAG PhD council with your full support.

Gia đình thân yêu
Bố mẹ và **anh chị** yêu quí, con cảm ơn bố mẹ, em cảm ơn anh chị đã luôn dõi theo con/em trong hành trình trau dồi tri thức và kỹ năng của một người làm khoa học. **Anh Cường**, cảm ơn anh đã luôn đồng hành và hỗ trợ em trong cuộc sống gia đình. Em rất trân trọng tấm lòng của anh. **Thức Đan**, cảm ơn con đã đến bên bố mẹ để cuộc sống của bố mẹ trở nên trọn vẹn.

Wageningen 2013, Phan Vân Anh

About the author



Phan Vân Anh was born on March 1st, 1980 in Thai Nguyen, Vietnam. She graduated as an Engineer in Food Technology at Hanoi University of Science and Technology (Hanoi, Vietnam) in 2003. Subsequently, she pursued a Master's degree in Food Science at ENSBANA, Bourgogne University (Dijon, France). Investigating how the structure of a cheese imitation influences salt release and perception while chewing was the subject of her master internship at INRA Dijon. From 2005 to 2008, Vân Anh worked at Nestle Research Center (Lausanne, Switzerland) in a large project entitled "Multi-modal mechanisms of fat perception". Afterwards, she began her PhD journey at Wageningen University (Wageningen, the Netherlands), within the group of Food Quality and Design. She explored during this journey how Bayesian networks, a machine learning technique, can be used to connect data from related studies. Most data for her work were produced by her PhD fellows in the same project entitled "Sensory specific satiation: linking product properties to obesity prevention". Currently, Vân Anh works at Solanic/AVEBE (Veendam, the Netherlands) as a Protein Application Developer.

Patent

Phan, Van Anh; Godinot, Nicolas; Sagalowicz, Laurent; Leser, Martin; Robert, Fabien. Oil-in-water emulsion and its use for the delayed release of active elements. WO2008145183. 2008-12-04

Publications

Phan V.A., Garczarek U., van Boekel M.A.J.S., Dekker M. Towards Global Experimental Design in food science using Bayesian networks. Submitted for publication.

Phan V.A. and Garczarek U., Bolhuis, D.P., van Boekel M.A.J.S., Dekker M. Bayesian networks as a tool to analyze causal relationships in experimental designs: a case study on oro-sensory exposure studies. Submitted for publication.

Phan V.A. and Bolhuis, D.P., Garczarek U., van Boekel M.A.J.S., Dekker M. Generating new insights in the effect of eating behavior and sensory aspects on food intake by Bayesian network modeling. Submitted for publication.

Phan V.A., Ramaekers, M.G., Bolhuis, D.P., Garczarek U., van Boekel M.A.J.S., Dekker M. 2012. On the use of Bayesian networks to combine data from related studies on sensory satiation. Food Quality and Preference, 26 (1): 119-127

Phan V.A., Kole A.P.W., Garczarek U., Dekker M., van Boekel, M.A.J.S. 2012. Bayesian networks to explain the effect of label information on product perception. Procedia Food Science 1: 1084 – 1090

Phan V.A., Garczarek U., Dekker M., van Boekel M.A.J.S. 2010. Bayesian networks for food science: theoretical background and potential applications. In S. R. Jaeger & H. Mac Fie, Consumer-driven innovation in food and personal products . Cambrige: Woodhead Publishing Limited, pages 487-513

Phan V.A., Liao Y., Antille N., Sagalowicz L., Robert F., Godinot N. 2008. Delayed volatile compound release properties of self-assembly structures in emulsions. Journal of Agriculture and Food Chemistry, 56(3): 1072–1077.

Phan V.A.,, Yven C., Lawrence G., Chabanet C. , Reparet J.M., Salles C. 2008. In-vivo sodium release related to salty perception during eating model cheeses of different textures. International Dairy Journal, 18: 956-963.

Salles C., Phan V.A., Yven C., Chabanet C., Reparet J.M., Le Quere J.L., Lubbers S.

2006. Decourcelle N., Guichard E. In-vivo flavour release from dairy products: relationships between aroma and taste release, temporal perception, oral and matrix parameters. Developments in Food Science , 43: 385-390

Pham Thu Thuy, Nguyen Thuy Huong, Phan Van Anh. 2006. Synthesis of enzyme invertase from Saccharomyces cerevisiae in industrial molasses of Hoa Binh. Vietnamese Journal of Science and Technology, 44: 85-89

Overview of completed training activities

General courses	Credit	Year
Philosophy and Ethics of Food Science and Technology, VLAG	1.5	2008
VLAG PhD week	1.5	2008
Information literacy and EndNote	0.6	2008
PhD competence assessment, WGS	0.3	2008
Project and Time Management, WGS	1.5	2008
Techniques for Writing and Presenting Scientific Papers, WGS	1.2	2009
Effective behavior in your scientific surroundings, WGS	0.7	2010
Academic writing II, Language Services WU	4.0	2011
Scientific writing, Language Services WU	1.8	2011
Career perspectives, WGS	1.6	2012
Interdisciplinary Research: Crucial knowledge and skills, WGS	1.1	2012
Discipline courses and workshops		
Regulation of Food Intake and its Implications for Nutrition & Obesity, VLAG	0.9	2008
Bayesian Machine Learning workshop	0.3	2008
HUGIN training course	0.9	2008
Introduction to R for statistical analysis	0.6	2008
Bayesian statistics	0.6	2009
Basic statistics	1.5	2009
Exposure Assessment in Nutrition Research, VLAG	1.5	2010
Sensometrics workshop	0.6	2010
Orientation on Mathematical Modeling in Biology	1.5	2011
Linear models	0.9	2011
Mixed linear models	0.6	2011
Generalized Linear models	0.6	2011
WinBUGS workshop: Bayesian Modeling for Cognitive Science	1.5	2011
Expert workshop for Food Intake Bayesian network	0.3	2011
International conferences and presentations		
8th Pangborn Sensory and Consumer Research Symposium	1.4	2009
Oral presentation at the 8th Pangborn	1.0	2009
10th Sensometrics	1.1	2010
Oral presentation at the 10th Sensormetrics	1.0	2010
11th Internaltional Congress on Engineering and Food (iCEF)	1.4	2011
Oral presentation at the 11th iCEF	1.0	2011
Other activities		
Four-month internship at Unilever Vlaardingen	6.0	2009
Organizing a PhD trip for PDQ group	1.0	2009
PhD trip	2.8	2010

Colophon

The research described in this thesis was financially supported by the Dutch Technology Foundation STW (grant 07438).

Financial support from Wageningen University and the Dutch Technology Foundation STW for printing this thesis is gratefully acknowledged.

Cover design and layout: agilecolor design studio/atelier (www.agilecolor.com)
Printed by: Gildeprint Drukkerijen, Enschede (NL) (http://www.gildeprint.nl)

FSC
www.fsc.org
MIX
Paper from
responsible sources
FSC® C012854