Bayesian inference – a way to combine statistical data and semantic analysis meaningfully

Eila Lindfors

This article focuses on presenting the possibilities of Bayesian modelling (Finite Mixture Modelling) in the semantic analysis of statistically modelled data. The probability of a hypothesis in relation to the data available is an important question in inductive reasoning. Bayesian modelling allows the researcher to use many models at a time and provides tools to evaluate the goodness of different models. The researcher should always be aware that there is no such thing as the exact probability of an exact event. This is the reason for using probabilistic models. Each model presents a different perspective on the phenomenon in focus, and the researcher has to choose the most probable model with a view to previous research and the knowledge available.

The idea of Bayesian modelling is illustrated here by presenting two different sets of data, one from craft science research (n=167) and the other (n=63) from educational research (Lindfors, 2007, 2002). The principles of how to build models and how to combine different profiles are described in the light of the research mentioned.

Bayesian modelling is an analysis based on calculating probabilities in relation to a specific set of quantitative data. It is a tool for handling data and interpreting it semantically. The reliability of the analysis arises from an argumentation of which model can be selected from the model space as the basis for an interpretation, and on which arguments.

Keywords: method, sloyd, Bayesian modelling, student teachers,

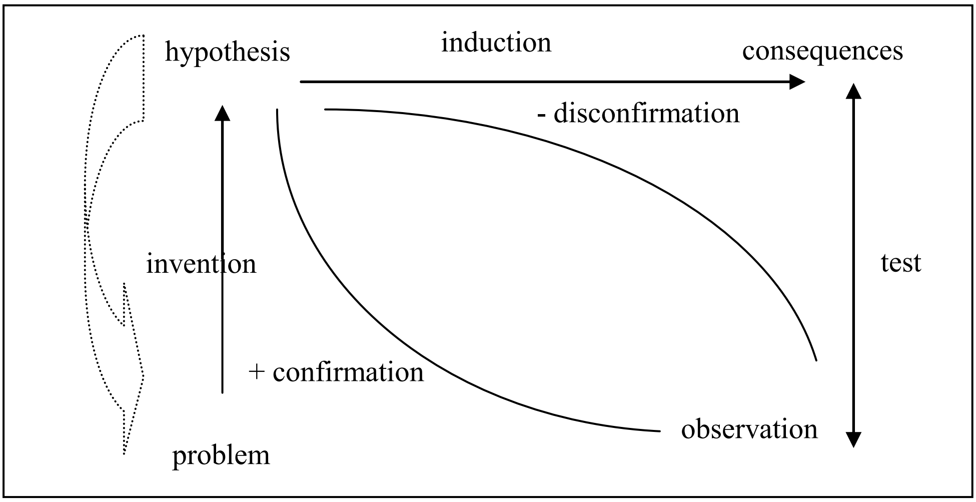
Background and theoretical perspective

If we take a look at craft science research during the last two decades and sloyd education research, we have to admit that much of it is qualitative. This is obvious, for if we are dealing with philosophical questions (e.g., Kojonkoski-Rännäli, 1995;

Karppinen, 2007) or in some cases with time (e.g., Luutonen, 1997; Kaipainen, 2008), or try to analyse interviews (Kokko, 2007), the qualitative method and concept and content analysis are the most suitable ways of gaining a more profound understanding of the problems. If the data is quantitative, collected by questionnaires or by exact measurements, and the volume is large and includes different variables, the researcher must resort to computer power for examining the data and for finding answers to the problems with the help of the data. Factor analysis, regression analysis and variance analysis are probably the most commonly-used linear methods with quantitative data. Critical voices have claimed that young researchers do not know how to use quantitative methods and handle large bodies of data (Törmäkangas, 2004). Due to this, some interesting information will never be discovered in large data spaces.

In social sciences in general, modern educational research is constantly faced with the problem of reasoning based on incomplete and uncertain information (Tirri, 1999). Researchers cannot usually put people in a laboratory to separate the independent and dependent variables and linear and non-linear dependencies. As we are dealing with quantitative data, our main problem is to find linear and non-linear dependencies in the data space to be able to classify the information units in order to make judgements. In craft science and in sloyd education researches large data spaces are created e.g. by using questionnaires or measuring properties of materials, equipments and tools. To find and to understand the relations between variables the researcher has to classify the data. In classification, the units (cases, data vectors, subjects, individuals) must be grouped on a specific basis. The task is to build a model of the problem domain for predicting the group membership (classes) of previously unseen units (cases, data vectors, subjects, individuals), given the descriptions of the units (Silander & Tirri, 1999). The discovery of a previously unknown structure occurs most frequently when there are many relevant variables describing each case (Tirri, Silander & Tirri, 1997). The method for finding classes has traditionally been based on linear models. Non-linear modelling techniques have developed only recently as a result of increased computing power. They are already applied on industrial, economic and biological phenomena, but are still almost unknown and rarely used in the social sciences. From the viewpoint of modelling in social sciences, Bayesian inference is one of the most appealing approaches (Ruohotie & Tirri, 1999; Silander, 2009).

Bayesian inference can be illustrated on the basis of hypothetic-inductive reasoning (figure 1). In inductive reasoning, the probability of a hypothesis in relation to a given set of data is a basic principle. This means that in the data space, there are many possibilities of modelling the data. On the basis of premises, the researcher has to decide how to model the data. However, in most cases it is impossible to model the data as an exact picture of reality, which is why the concept of predictive modelling is important. Consequently, the researcher should make choices and decisions which will lead to the most predictive models. The researcher must consider the phenomenon in focus in relation to earlier research results and the data available, which both include uncertainty on many levels. The modelling and decisions will be made in the framework of uncertain information (Tirri, 1999). The question to be asked is which explanation is the most probable.



Figur 1: The hypothetic-inductive method (Niiniluoto, 1983, s. 130).

In quantitative analysis, uncertainty is typically described with probabilities (Tirri, 1999). The main point of Bayesian inference is its predictive value. The interpretation of Bayesian probability takes into account the uncertainty of information (Berger, 1988; Gelman et. al., 1995). It helps us to make decisions that can affect the phenomenon. The result of Bayesian inductive inference is not the confirmation or rejection of a hypothesis (Niiniluoto, 1983). Rather, it is an evaluation of the probability of the hypothesis on the basis of the model selected. The modelling is successful if the models are capable of predicting the cases that we can observe. In calculating probabilities, the Bayes pattern is applied (see B Course, 2002; ISBA, 2010).

This article focuses on presenting principles of Bayesian inference by two research examples. The first example (Lindfors, 2002), a craft science research (167 respondents, 3 questionnaires with 100 variables in each) is an expert evaluation of textile properties (figure 3). In the second example (Lindfors, 2007), the data concerned student teachers' concepts (63 respondents, two questionnaires with 40 variables in each) on educational sloyd before and after a problem-based user-centred design course (figure 4). In the end of the article the use of Bayesian inference is considered critically and some recommendations are presented.

Data collection and empirical investigation

Research usually starts by defining the problems in a problem space. The second phase is to collect and model data for making predictions and conclusion on what the construction of the data is. Does the data reveal some unknown relations? The final phases consist of an evaluation of the deductions and predictions made on the basis of the statistical analysis, following which the researcher may proceed to ask new questions. Certain factors recommend a Bayesian approach (Nokelainen, 2008; Silander, 2009; Tirri, 1999). It is highly appropriate in situations with small data sets with many variables, data sets that include discrete values and data that involves latent structures. The Bayesian framework offers the researcher:

- time-efficient prediction in model construction
- use of few technical parameters
- good prediction performance even with small data sets
- principled approach to avoid over-fitting the data in the model construction
- possibility of controlling the model construction time

The benefits of using Bayesian inference are:
- it allows a study of the model structure and its parameters
- the generalizability of the models is good
- the data can be studied in many ways with semantic inference
- the amount of data needed can vary
- the models can be found clearly
- it can be used flexibly in many different situations
- it is possible to use and mix discrete and continuous variables
- missing data can be marginalized
- it offers the possibility of comparing different ways of doing things

(Myllymäki & Tirri, 1998; Nokelainen, 2008; Silander, 2009; Tirri, 1997)

The Bayesian approach is flexible and produces clear and direct inferences. It allows the use of all the information available. It enables the combining of statistical data and expert knowledge in a natural way. Without this aspect, the researcher cannot choose a meaningful model from a model family. In the Bayesian inference, to get the best possible decision there is always an attempt to combine the sample information with other relevant aspects of the problem. Due to this, the researcher has to take into account knowledge of the possible consequences of the decisions and prior information which arises from sources other than statistical information (Berger, 1988).

Methods and analysis

Faced with a research problem, we try to model some part of the universe and make decisions based on that model to get answers to the research problems. In Bayesian inference, data is examined with several models (Figure 2). The basic idea is that there are many answers (models) to the research problems. In a typical research situation, the researcher has some background information about the phenomenon, some related information from previous studies, and some gathered data, all of which contain uncertainty on many levels. The researcher must deal with uncertain information. Bayesian modelling attempts to build probabilistic models to keep this subjective interpretation in mind. The researcher has to decide which model gives the best possible answer to the research problems, and give explicit grounds for the decision.

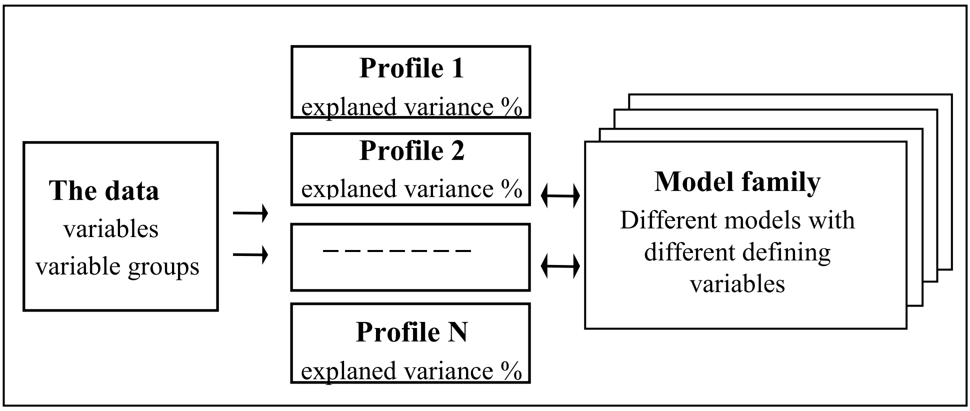


Figure 2. Model construction in Finite Mixture Modelling (Lindfors, 2002, s. 73).

In the construction of models, the aim is to create probabilistic models which truly describe the problem space. Models are usually constructed in the framework of all the data available, instead of including only some defined variables or groups of variables in the analyses. The criteria chosen for the model construction define the features/properties which the researcher is interested in. The criteria are defined either on the basis of counting the example data or on prior knowledge, or both. Models are means for describing the interesting features/properties of the cases in the data. To be able to choose a model as the basis of the analysis, the researcher has to understand the principles of creating models and their tasks. The two main points in Bayesian modelling are: 1) to construct probabilistic models for the cases in the problem space and 2) to apply a subjective view in the semantic interpretation of the models.

In Bayesian modelling, the modelling itself as a starting point offers several models based on available data to each research problem (Nokelainen, 2008; Silander, 2009). The first task in modelling is to build a family of models which broadly describes the whole problem space. The variables which define the different models inside a model family are not the same. The different variables and variable groups create different models, and the number of profiles in each model varies (Figure 2). Every model opens a different perspective on the research object. For every model in the family, there are some variables which define that model to a greater extent than other variables. To find the best model, the researcher has to consider three important questions (Tirri, 1999): 1) What models are possible? 2) How is it possible to compare models? and 3) How is it possible to find good models?

The model is always chosen from a set of models because for every problem there are an infinite number of models (Figure 2). The models and model families define the limits for observing similarities and differences in the data. (Tirri, 1999.) The models are constructed on the basis of example cases. These cases (e.g., persons, data sets, combinations of variables) differ from each other but, at the same time, they act as example cases for different profiles. The profiles are formed by the other cases joining an example case on the basis of being similar to the example case of a given profile and, at the same time, as different as possible from the example cases of the other profiles . In this way, Bayesian modelling constructs profiles which the cases can join with the highest probability. Models with different numbers of profiles explain the data in different ways (Figure 2), because the variables defining the model and the profiles are different. The cases included in a given profile join its example case by means of specific variables or groups of variables. The order of importance of a given variable is different in different profiles.

When modelling, we try to build the best possible model on the basis of prior knowledge and the data available. The researcher has to evaluate previous information in order to decide what models are possible for a specific problem. To compare the possible models, the researcher can use scoring (see Silander, 2009). Usually, simpler models are preferred over more complex ones. (Tirri, 1999;

Nokelainen 2008). Models with a better prediction power are better. (Heckerman, 1996; Silander & Tirri, 1999).

Selecting a model as the basis of interpretation requires that the model is a good one both in terms of calculations and in the light of prior knowledge. The principle of Occam's Razor calls for selecting the simplest one from among equally good models (Silander, 2009). It is a challenge to identify a model which presents the structure of the data without allowing too much explanation power to individual cases in small profiles. The model chosen here was capable of being interpreted in relation to prior knowledge. The background reasoning must also be meaningful, as well as the conclusions made on the basis of modelling (Berger, 1988). The models are used for predicting features/properties of interest and for deciding on actions based on these predictions. (Tirri, 1999). One computer application of Bayesian modelling is called Bayesian Mixture Modeling.

Presentation of results

The following presents two authentic models created by Finite Mixture Modelling. Figure 3 shows a model used in a craft science dissertation to analyse how important experts in the textile branch rated the knowledge of the properties of textile products for consumers as textile users (Lindfors, 2002). Several models were calculated, the number of profiles and their explanation variance assessed, the variables defining the different models evaluated, the order of importance of variables in different profiles estimated and the example cases in each profile considered, the probabilities as basis of inference evaluated and mean values rated. Finally, the five-profile model was chosen as the basis of analysis.

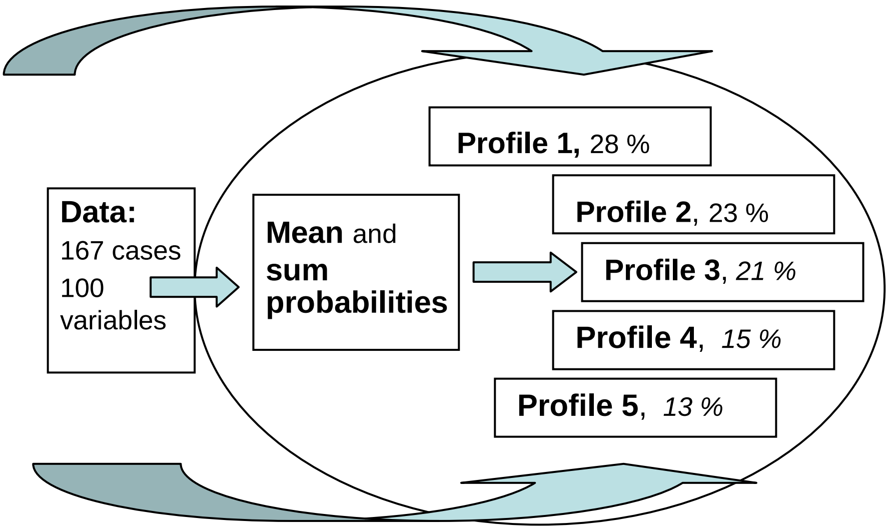


Figure 3. Crafts science research: importance of knowledge of properties of textile products from consumer viewpoint (Lindfors, 2002). Explanation variance of each profile in 5-profile model.

The data itself consisted of many classes of textile technological knowledge in relation to several product groups. Rating statements were used. This meant that there were numerous different perspectives on the phenomenon which had to be combined and separated by finding the variables and groups of variables which defined the cases included in the modelling. The analysis started with finding the mean probabilities of the different variables in the model (Figure 3). This was a way to assess variables and summary variables in relation to all data. The analysis continued by a description and naming of the profiles on the basis of the most important variables in each profile (Table 1). A comparison of the profiles to the mean profile and to each other revealed that the profile with the greatest explanation variance did not resemble the mean profile. In this model, the second smallest profile was closest to the calculated mean value, even though its explanation variance was only 15% of the model. This profile stressed the care properties as the most important for consumers in terms of the use of textiles. This is similar to official regulations concerning care instruction in product labels. This means that all the other four profiles highlighted completely different ideas than the profile closest to the mean (Figure 1).

Table 1. Explanation variance of profiles in 5-profile model (Lindfors, 2002).

Prof.	**Explan. variance**	**Profile name based on defining variables**	*Relation to mean profile*
1	28 %	Knowledge of properties of leisure, casual and outerwear clothing very important for consumers	4
2	23 %	Aesthetic appearance, safety of use and biological resistance of textiles important for consumers	2
3	21 %	Knowledge of care properties of textiles and casual, leisure and outerwear clothing properties important for consumers	5
4	15 %	Knowledge of care properties of textiles very important for consumers	1
5	13 %	Knowledge of properties of outerwear important for consumers	3
Mean	20 %	Mean profile	0

The biggest profile, with 28-% explanation variance, revealed a totally different view, which considered the use of textiles from a product group perspective. This example shows that modelling classifies the data effectively and brings to daylight different perspectives on the research problems by means of profiles. In the second example, the data concerned student teachers' concepts on educational sloyd before and after a problem-based user-centred design course in the didactics of sloyd (Lindfors, 2007, figure 4). The data collected before the studies in sloyd didactics was modelled in a four-model family (see Figure 2), each model having 2-5 profiles. A dichotomous scale was used to force the student teachers to choose their side: yes or no! On assessing the four models, the two-profile model was found to be the most usable. The other models did not reveal other structure in the data than that shown by the two-profile model (Table 2). The profiles were exactly the same size, half of the data. The first profile was mainly defined by positive attitudes on sloyd at school. The other profile consisted of student teachers who had only taken obligatory sloyd studies at school, having their earlier experiences in sloyd from the 7th grade of the comprehensive school. Negative experiences and attitudes against sloyd were also registered in this profile.

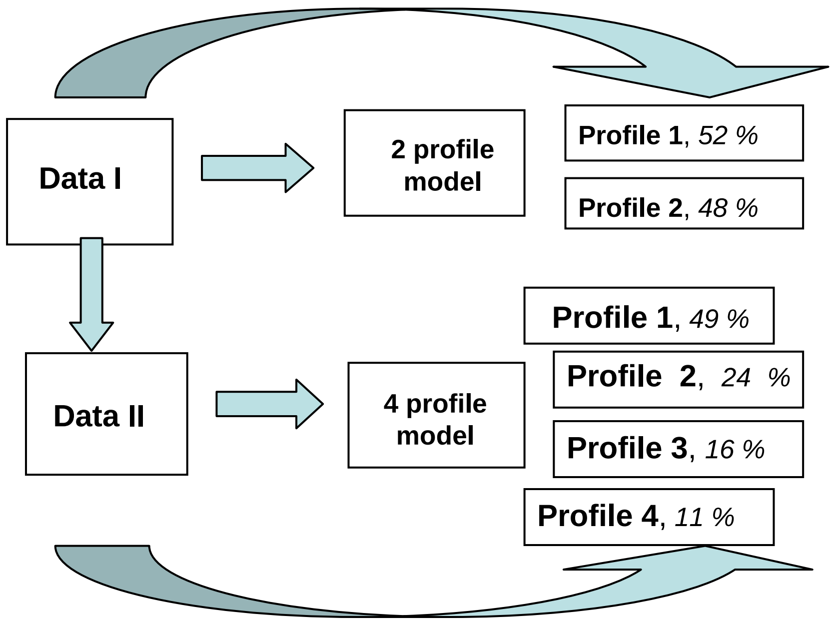


Figure 4. Student teachers' concepts on educational sloyd before and after a problem-based user-centred design course (Lindfors, 2006).

Table 2. Student teachers' concepts on educational sloyd before didactical studies in sloyd (Lindfors, 2006).

Pro file	Explanati-on variance	Student teachers' concepts on educational sloyd before didactical studies in sloyd	Relation to mean profile
1	52%	Positive experiences of school sloyd as a background.	1
2	48%	Obligatory sloyd studies at school as a background.	2

This two-profile model was the only meaningful way to interpret the data. After the didactics course in sloyd, new data was gathered and the new 2–5-profile model family was calculated. In the light of previous knowledge and the structure of the explanation variance, the most usable and the easiest to interpret was the four-profile model (Figure 4, Table 3). The profiles were formed on the basis of the student teachers' earlier school experiences and the didactical ideas of sloyd received in the course of teacher education.

Table 3. Student teachers' conceptions on educational sloyd after didactical studies in sloyd (Lindfors, 2006).

Pro file	**Expla-nation variance**	**Student teachers' conceptions on educational sloyd after didactical studies in sloyd: four-profile model.**	**Relation to mean profile**
1	49%	Sloyd is solving problems by hands.	1
2	24%	Sloyd is creative working by hands.	2
3	16%	Sloyd is useful therapy.	3
4	11%	Sloyd is making things by different techniques.	4

In the largest profile (Table 3), nearly half of the student teachers (cases) stressed the meaning of sloyd in education as a problem-based, holistic design and production process. This profile described half of the data and was closest to the mean profile. Those students, one fourth of the cases, who had had negative school experiences, emphasized the importance of support for the students' creativity and innovativeness in sloyd. The third profile considered sloyd as therapy in the form of making useful things. The fourth profile consisted of students who saw sloyd as the learning of techniques. The two smallest profiles together represented one fourth of the cases.

The four-profile model differentiated the data in such a way that the researcher could make conclusions of the student teachers' conceptions before and after the didactics course in teacher education and could also understand the reasons for the different conceptions. Compared to earlier studies, the new clarifications and background knowledge made it possible to consider the reasons for different types of cases, i.e., the profiles. The analysis also helped to understand the student teachers' thinking in sloyd before and after the didactics course.

Reflection

Bayesian inference is based on the degree of belief in the interpretation of probabilities. Thus, the uncertainty of information in the Bayesian framework is represented as probabilities. Bayesian probability can be defined as a subjective assessment of whether the cases, events or things in question will occur. (Nokelainen 2008; Tirri 1999.) The researcher should always be aware that there is no such thing as the exact probability of an exact event. In traditional linear modelling, probability is based on frequency calculations. In Bayesian modelling, probability always depends on the state of knowledge of the one who believes: what is the structure of a model under which premises. So, depending on previous information, the probabilities can be different. There is always the possibility that our initial information is faulty. In a Bayesian context, subjectivity does not mean arbitrariness. With different information one may get different probabilities, but if the information is the same, the probabilities should also be the same. It can be said that the state of knowledge determines the value of the probability. (Nokelainen, 2008; Tirri, 1999). This means that all decisions made have to be documented very carefully to point out the basic underlying ideas and so enable the possibility of evaluating the decision, both empirically and analytically. One of the benefits of Bayesian modelling is the combination of expert knowledge and statistical learning. (Myllymäki & Tirri, 1998.) Inductivity means rational reasoning in a framework of uncertainty. The researcher has to evaluate the models and their consequences in the light of previous knowledge. Is the new structure of data in a model relevant and does it reveal information which logically helps to understand the phenomenon in a meaningful way? An earlier example (Table 1) revealed that there were different ways of considering the properties of textiles. Different variables defined the example cases in each profile and the biggest profile was not close to the mean. If the findings are not logical when compared with prior knowledge and previous research results, the researcher must discuss the issues critically and understand explicitly what causes the conflict.

To choose the model used for interpretation, the researcher has to consider the whole model family (Figure 2) and to compare the structure of different models (Nokelainen & Tirri, 2002). The advance criteria for a good model are its value in explaining the phenomenon (goodness of fit) and its interpretability. The number of profiles and their explanation variance have to be assessed, the variables defining the different models have to be evaluated, the order of importance of the variables in different profiles estimated and the example cases in each profile considered; moreover, the profiles must be viewed in relation to the mean profile (see Lindfors, 2002). If there are many profiles included in the interpretation, the answer to the research problem will be detailed. However, if the profiles are too small, they will not present a differentiated structure inside the data; instead, they present individual cases which cause problems when evaluating the validity of the research. If the cases in the model happen to be homogeneous, even a two-profile model is possible (Table 2), but if the model evaluation reveals that there are heterogeneous cases, a model with small profiles is needed as a basis of the explanation and semantic analysis of the phenomenon. The mean probabilities (Figure 2) describe the data on model level, and the explanation variance of each profile reveals part of a model. A variable can define a certain profile while being useless in another profile.

Research in craft science and in sloyd education, as well as in other social sciences, is done in uncertain and varying circumstances, seldom in a laboratory where many variables can be controlled by the researcher. In pedagogy, the learning and teaching contexts, the individual learners and groups of learners cannot be placed in a laboratory to conduct the research at hand. Even if this were possible, the results gained in a laboratory would not describe the phenomenon in real circumstances. There is no exact truth. Instead, phenomena differ from each other as different dimensions and variables affect them. The research results must always be considered in a specific context with specific constraints.

Bayesian modelling includes the idea that the analysis is made under uncertain conditions. The conclusions will initially be considered with the help of several models. Finite Mixture Modeling describes the data by means of a family consisting of several models (Figure 2). Some of the models show such results that the researcher cannot infer them logically from previous research and theory. Some of the models reveal interesting views which can bring new structures and classifications to light. The researcher has to consider the different models carefully and choose as the basis of inference a model which reveals interesting features from the viewpoint of the research problems (Figures 3 and 4). Bayesian modelling seems appropriate in situations where the data consists of small samples with many variables with discrete values (Tirri, 1999). Bayesian inference is flexible and thorough, and the requirement of explicit argumentation in the model construction makes the analysis valid.

All the profiles in one model describe the case inside the model differently. The profiles are formed of such variables and such ordering of variables which best describe specific cases in the profile (see Tables 1–3). There is no reason to restrict the number of variables. The non-linear connections between the variables will become visible even if the parameters for the model construction were defined with some specific variables. Thus, even if the models are calculated on the basis of specific variables, they can reveal unknown connections and classes inside the data. This is why the researcher has to consider the models critically: What does each of the different models reveal? Do they reveal something which supports the known facts or something that was not expected? Bayesian modelling allows the use of many models and the means to evaluate the goodness of the models.

The two examples of Finite Mixture Modelling presented (Figures 3 and 4 and Tables 1–3) demonstrate the principles of Bayesian modelling (Figures 1 and 2).

Both examples present different perspectives on examining the data. The researcher has to know that the model families do not offer a technical solution as the basis of analysis. The model families insist that the researcher presents an explicit argumentation for choosing a model as the basis for analysis. The use of previous research results, consideration of the structure of the models and profiles, and consideration of the independent and dependent variables led to the selection of the models in the example data (Lindfors, 2007, 2002) presented here. An understanding of phenomena in education via Bayesian modelling includes the concepts of relativity, diversity and uncertainty in modelling and inference.

In developing research methods one should consider and understand the character, possibilities and opportunities, as well as the limitations of methods. Consequently, using different methods side by side is also justified (see Kaartinen, 2005). A technical analysis without consideration of the phenomena, the environment and the data will lead to mechanical solutions instead of understanding. Bayesian modelling is a type of modelling not often used. To understand the nature of Bayesian modelling, to gain user experiences and to develop methods further, researchers should have the curiosity and creativity to apply and use new methods. User experiences of Bayesian modelling reveal that it provides the researcher with probabilistic methods for understanding the data and the phenomenon being studied (see Nokelainen & Ruohotie, 2002; Ruohotie et al., 2001, 2002; Tirri & al., 1997). It also challenges the researchers to argue explicitly for the reason why the data was modelled in the way chosen.

References

Berger, J. O. (1988). *Statistical Decision Theory and Bayesian Analysis*. Springer Series in Statistics. New York: Springer-Verlag.

B-Course (2002). In http://b-course.cs.helsinki.fi/obc/, read on 15.11.2010.

Cheeseman, P. & Stutz, J. (1996). Bayesian classification AutoClass: Theory and results. In U. M. Fayyad, & G. Piatetsky-Shapiro & P. Smyth & R. Uthurusamy (eds.) *Advances in Knowledge Discovery*. Massachusetts: AAAI Press, 153–180.

Eskola, J. (2004). On valintojen aika. *Kasvatus* 35 (5), 566–568.

Gelman, A. & Carlin, J.B. & Stern, H. S. & Rubin, D. B. (1995). *Bayesian Data Analysis.* London: Chapman & Hall.

Heckerman, D. (1996). Bayesian Networks for Knowledge Discovery. In U. M. Fayyad, & G. Piatetsky-Shapiro & P. Smyth & R. Uthurusamy (eds.) *Advances in Knowledge Discovery*. Massachusetts: AAAI Press, 273–305.

Heikkinen, H.L.T., Huttunen, R., Niglas, K. & Tynjälä, P. (2005). Kartta kasvatustieteen maastosta. *Kasvatus* 36 (5), 340–354.

ISBA (2010) International Society for Bayesian Analysis. In http://www.bayesian.org/bayesexp/bayesexp.html, read on 15.11.2010

Kaartinen, T. (2005). *Itsesäätelyvalmiudet musiikin opiskelussa*. Acta Universitatis Tamperensis 1125. University of Tampere. Tampere: Tampereen yliopistopaino.

Kaipainen, M. (2008). *Ken tilauspukua käyttää, hän herrasmieheltä näyttää. Eteläkarjalainen maalaisvaatturi ja vaatturitoiminta Suomessa 1920–1960-luvuilla.* Joensuun yliopiston kasvatustieteellisiä julkaisuja N:o 125.

Karppinen, S. (2007). *'Mitä taide tekee käsityöstä?' Käsityötaiteen käsitteellinen analyysi.* Helsingin yliopisto. Tutkimuksia 263.

Kojonkoski-Rännäli, S. (1995). *Ajatus käsissämme. Käsityön käsitteen merkityssisällön analyysi.* Turun yliopiston julkaisuja. Sarja C. 109.

Kokko, S. (2007). *Käsityöt tyttöjen kasvatuksessa naisiksi.* Joensuun yliopiston kasvatustieteellisiä julkaisuja n:o 118.

Lindfors, E. (2007). Sloyd in education – Student teacher perspective. In M. Johansson & M. Porko-Hudd (eds.) *Knowledge, Qualities and sloyd. Research in Sloyd Education and Crafts Science.* Techne Series. A:10/2007, 53–73.

Lindfors, E. (2002). *Tekstiilituotteen teknologiset ominaisuudet.* Joensuun yliopisto. Kasvatustieteellisiä julkaisuja n:o 77.

Myllymäki, P & Tirri, H. (1998). *Bayes-verkkojen mahdollisuudet*. Teknologian kehittämiskeskus TEKES. Teknologiakatsaus 59/98.

Niiniluoto, I. (1984). Johdatus tieteenfilosofiaan. Käsitteen- ja teorianmuodostus. Keuruu: Otava.

Niiniluoto, I. (1983). Tieteellinen päättely ja selittäminen. Keuruu: Otava.

Nokelainen, P. (2008) Modeling of Professional growth and Learning. Bayesian

Approach. Acta Electronica Universitatis Tamperensis 728. In http://acta.uta.fi/pdf/978-951-44-7328-9.pdf, red on 15.11.2010

Nokelainen, P. & Ruohotie, P. (2002). Modeling students' motivational profile for learning in vocational higher education. In H. Niemi & P. Ruohotie (eds.) *Theorethical undestandings for learning in the virtual university*. Saarijärvi: Saarijärven offset Oy, 177–205.

Nokelainen, P. & Tirri, H. (2002). Issues in designing an interactive personalized self assessment tool. In H. Niemi & P. Ruohotie (eds.) *Theorethica undestandings for learning in the virtual university.* Saarijärvi: Saarijärven offset Oy, 73–89.

Ruohotie, P. (1999). Professional Growth and Development. In Ruohotie, P. & Tirri, H. & Nokelainen, P. & Silander, T. (eds.) *Modern Modeling of Professional Growth*. Vol. 1. Saarijärvi: Saarijärven Offset Oy.

Ruohotie, P., Nokelainen, P., Tirri, H. & Silander, T. (2001). *Modeling Individual and Organizational Prequisites of Professional Growth*. Saarijärvi: Saarijärven Offset Oy.

Ruohotie, P. & Tirri, H. (1999). Introduction. In Ruohotie, P. & Tirri, H. & Nokelainen, P. & Silander, T. (eds.) *Modern Modeling of Professional Growth*. Vol. 1. Saarijärvi: Saarijärven Offset Oy.

Silander, T. (2009) The Most Probable Bayesian Network and Beyond. University of Helsinki. Department of Computer Science Series of Publications. Report A-2009-2. In https://www.doria.fi/bitstream/handle/10024/44974/themostp.pdf?sequence=1, read on 15.11.2010

Silander, T. & Tirri H. (1999). Bayesian classification. In Ruohotie, P. & Tirri, H. & Nokelainen, P. & Silander, T. (eds.) *Modern Modeling of Professional Growth*. Vol. 1. Saarijärvi: Saarijärven Offset Oy.

Tirri, H. (1997). *Plausible Prediction by Bayesian Inference.* University of Helsinki. Department of Computer Science. Series of Publications A. Report A-1997-1.

Tirri, H. (1999). What the Heritage of Thomas Bayes has to Offer for Modern Educational Research. In P. Ruohotie & H. Tirri, P. Nokelainen & T. Silander (eds.) *Modern Modeling of Professional Growth.* Vol. 1. Saarijärvi: Saarijärven Offset Oy.

Tirri, H., Silander, T., Tirri, K. (1997). *Bayesian Finite Mixtures for Nonlinear Modeling of Educational Data.* A paper presented at the 1997 Annual Meeting of the American Educational Research Assosiation (Chicago, IL, USA, March 1997).

Törmäkangas, K. (2004I). Kvantitatiiviset menetelmät ja ”menetelmät” opinnäytetöissä ja tutkimuksissa. *Kasvatus* 35 (4), 445–447.

Eila Lindfors, PeD, lektor vid pedagogiska enheten vid Tammerfors universitet, Finland. Ett av Lindfors forskningsintressen är användarcentrerad design i pedagogisk kontext. Lindfors har anpassat metoden Bayesian modelling till slöjdvetenskaplig och slöjdpedagogisk forskning.