PETRI NOKELAINEN

Modeling of Professional Growth and Learning

Bayesian approach

ACADEMIC DISSERTATION
To be presented, with the permission of the Faculty of Education of the University of Tampere, for public discussion in the Auditorium of Research Centre for Vocational Education, Korkeakoulunkatu 6, Hämeenlinna, on June 17th, 2008, at 12 o’clock.
For in much wisdom is much grief: and he that increaseth knowledge increaseth sorrow.

(Ecclesiastes 1:18)
PREFACE

First, I would like to acknowledge the examiners of this article dissertation, Professor Erno Lehtinen (University of Turku) and Professor Paul Ilsley (Northern Illinois University). Their comments helped to improve and clarify this work.

Next, I wish to give credits to three Professors who have been mentors, co-authors and colleagues for my whole academic career, that is, well over ten years: Pekka Ruohotie (University of Tampere), Henry Tirri (Nokia Research Center) and Kirsi Tirri (University of Helsinki).

Pekka has provided me with challenging working opportunities in numerous vocational education research projects since 1995 in my first academic home, the Research Centre for Vocational Education (RCVE, University of Tampere). He has always given me true collegial support and a great amount of academic freedom by allowing me to do distance work. I thank him for being the co-author in the first, third and fourth original research publication.

I have had a great pleasure to work with Henry in the field of applied statistics and educational technology applications. I wish to thank him for being a co-author in the fourth original publication and a supportive colleague in my second academic home, the Complex Systems Computation Group (CoSCo, University of Helsinki), while I was working there as a visiting researcher in 1997 – 2005.

I wish to express my gratitude to Kirsi as a co-author of the second original publication, mentor and colleague with whom I have been successfully working in the research field of gifted and religious education since 1996. She has always been a supportive scholar and given me many opportunities to reach for the higher research standards.

My special thanks goes to a computer scientist and statistician M.Sc. Tomi Silander (CoSCo) not only for being my co-author in the fourth original publication, but also for his valuable feedback on too-many-to-mention-here statistical issues (especially in section 4.4 in this dissertation) and his incredible programming skills.
I also wish to thank my other colleagues during the CoSCo years: Ph.D. Jaakko Kurhila, M.Sc. Jussi Lahtinen and M.Sc. Miikka Miettinen. Miikka and Jaakko have joined their efforts with me to design an online questionnaire that I used in data collection for many years. Jussi has given me valuable comments on data visualization techniques and helped me with the data analysis.

I wish to thank Ph.D. Hanna-Leena Merenti-Välimäki (EVTEK polytechnic institution of higher education) as a co-author in the second original publication. Ph.D. Seppo Kolehmainen (HAMK polytechnic institution of higher education) has given valuable comments regarding the Growth-oriented Atmosphere Questionnaire. Working with Professor Erkki Komulainen (University of Helsinki) improved the quality of the third original publication considerably. I am also grateful to Ph.D. Jorma Saarinen (HAMK polytechnic institution of higher education) and Professor Hannele Niemi (University of Helsinki) for our past projects that helped me to test my models with empirical samples. Ph.D. Bruce Beairsto has been proof-reading all of the original publications and he has given some valuable comments.

I want to thank all the researchers and staff at the RCVE. My special thanks goes to Airi, Hilkka, Jaana, Kaja, Lea, Mika and Tarja for tolerating my flexible timing, unpredictable administrative standards and helping me with many practical matters for all these years.

My research work with all these aforementioned scholars has been funded by aforementioned universities together with numerous funding partners (e.g., European Union, Finnish Academy, Finnish Work Environment Fund, Ministry of Education, Tekes). I would also like to thank Ella and Georg Ehrnrooth foundation for a grant that helped me to complete this work.

Finally, thank You, God Almighty, for giving me Raila and two lovely children, Seela and Ruut.

Tuulos, April 2008

Petri Nokelainen
The major goal of the study was to contribute both to the basic research on professional growth and learning, and to the development and use of quantitative research methodology in these research areas. The research goal was addressed with four empirical studies conducted between 2002-2007.

The research questions were as follows: 1) What is the optimal number of dimensions and items in the Growth-oriented Atmosphere Questionnaire (GOAQ) to describe the theoretical model of growth-oriented atmosphere? (Study I); 2) Are the theoretical dimensions of the Self-confidence attitude attribute Scales (SaaS) questionnaire identified in the domain of three groups of mathematically gifted participants: Academic mathematics Olympians, polytechnic institute of higher education students, and elementary school students who have participated in mathematical competitions? (Study II); 3) What is the optimal number of dimensions and items in the Abilities for Professional Learning Questionnaire (APLQ) to describe the theoretical model of learning experiences and motivation? (Study III); 4) Is there a difference between substantive interpretations of the results of Bayesian dependency modeling and linear bivariate correlational analysis with professional growth data that has both linear and non-linear dependencies? (Study IV); 5) Is there a difference between substantive interpretations of the results of Bayesian dependency modeling and linear confirmatory factor analysis with professional growth data that has both linear and non-linear dependencies? (Study IV).

Results of the first study showed that the theoretical four group classification of the growth-oriented atmosphere factors was supported by the empirical evidence: 1) Support and rewards from the management; 2) Incentive value of the job; 3) Operational capacity of the team; 4) Work related stress. Further, the results of categorical factor analysis showed that the 67-item and thirteen-factor solution was the most interpretable in terms of correspondence to the theoretical GOA model.

Results of the second study showed that the theoretical four group classification of the self-confidence attitude attribute factors was supported by
the empirical evidence: 1) Success due to ability; 2) Failure due to a lack of ability; 3) Success due to effort; 4) Failure due to a lack of effort. Further, the results of exploratory factor analysis showed that the eight item and four factor solution was the most interpretable in terms of the attribution theory.

Results of the third study showed that the theoretical six group classification of the motivational factors was supported by the empirical evidence: 1) Intrinsic goal orientation; 2) Extrinsic goal orientation; 3) Meaningfulness of study; 4) Control beliefs; 5) Self-efficacy; 6) Test anxiety. Further, the results of confirmatory factor analysis showed that the 21-item solution was the most interpretable in terms of correspondence to the baseline model.

Results of the fourth study showed that in general Bayesian network models were congruent with the correlation matrixes as both techniques found the same variables independent of all the other variables. However, non-linear modeling found with both linear and non-linear samples a greater number of strong dependencies between the GOA factors. Results further showed that by using linear methods with non-linear data may lead to different substantive interpretations when using non-linear methods with the same data.

**Keywords:** Professional growth and learning, organizational atmosphere, learning motivation, self-attributions, quantitative research methods, Bayesian modeling
TIIVISTELMÄ


Tutkimuskysymykset olivat seuraavat: 1) Mikä on kasvuorientaatiomittarin (Growth-oriented Atmosphere Questionnaire, GOAQ) kasvuorientoituneen ilmapiirin teoreettisen mallin kuvauksen kannalta optimaalinen ulottuvuuksien ja väittämien lukumäärä? (Tutkimus I); 2) Ovatko itseluottamusta ja attribuutioita mitaavaan Self-confidence attitude attribute Scales (SaaS) –kyselyn teoreettiset ulottuvudet tunnistettavissa kolmessa matemattisesti lahjakkaista ihmisisä (matematiikan olympiisit, teknillisen ammattikorkeakoulun matematiikkalinjalaiset, peruskoulu ja lukion matematiikkakilpailun osallistujat) koostuvassa ryhmässä ryhmässä? (Tutkimus II); 3) Mikä on ammatillisen oppimisen valmiuksien keskittyvän mittarin (Abilities for Professional Learning Questionnaire, APLQ) oppimiskokemuksiin ja motivaatioon liittyvän teoreettisen mallin kuvauksen kannalta optimaalinen ulottuvuuksien ja väittämien lukumäärä? (Tutkimus III); 4) Onko bayesilaisen riippuvuussuhdemallin ja lineaarisen korrelaatioanalyysin tulosten tulkinnan väliä eroa, jos ammatillista kasvua kuvaavassa aineistossa on sekä lineaarisia että epälineaarisia vaikutustyahteita? (Tutkimus IV); 5) Onko bayesilaisen riippuvuussuhdemallin ja lineaarisen konfirmatorisen faktorianaalyysin tulosten tulkinnassa eroja, jos ammatillista kasvua kuvaavassa aineistossa on sekä lineaarisia että epälineaarisia vaikutustyahteita? (Tutkimus IV).

Ensimmäisen tutkimuksen tulokset osoitti että teoreettinen neljän pääulottuvuuden kasvuorientaatiomalli löytyi myös aineistosta: 1) Johdon tuki ja kannustus; 2) Työn kannustearvo; 3) Ryhmän toimintakyky; 4) Työhön liittyvää stressi. Tulokset osoittivat, että 67 väittämän ja 13 aliulottuvuuden faktoriratkaisu oli parhaiten tulkittavissa teoreettisen mallin valossa.

Toisen tutkimuksen tulokset osoittivat, että teoreettinen neljän itseluottamusta ja attribuutioita kuvaavan pääulottuvuuden malli sai tukea
empirisestä aineistosta: 1) Menestyminen kykyjen vuoksi; 2) Epäonnistumisen kykyjen puutteen vuoksi; 3) Menestyminen ponnistelujen puutteen vuoksi; 4) Epäonnistumisen ponnistelujen puutteen vuoksi. Tulokset osoittivat, että neljän faktorin ja kahdeksan väittämän ratkaisu oli tulkinnallisin attribuutioteorian valossa.

Kolmannen tutkimuksen tulokset osoittivat, että teoreettinen kuuden motivationaalisen pääulottuvuuden ratkaisu oli löydettävissä myös empirisen aineiston tarkastelun perusteella: 1) Sisäinen tavoiteorientaatio; 2) Ulkoinen tavoiteorientaatio; 3) Opintojen mielekkyyys; 4) Kontrolliuskomukset; 5) Tehokkuususkomukset; 6) Kohermostuneisuus. Tulokset osoittivat, että 21 väittämän ratkaisu oli riittävä mallin tulkinnan kannalta.

Neljännennen tutkimuksen tulokset osoittivat että bayesilaisten verkkomallien tuottamat tulokset olivat verrannollisia lineaarisen korrelaatioanalyysin tuloksiin. Bayes-mallinnus osoitti epälineaarisenä menetelmänä lisäksi lineaarisia menetelmiä suuremman määrän kasvuorientaatiofaktorien välisiä riippuvuuksia. Tulokset osoittivat, että jos epälineaarista aineistoa analysoidaan lineaarisilla analyysimenetelmissä, voidaan päättyä erilaisiin loppupäätelmiin kuin jos samaa aineistoa analysoitaan menetelmällä joka osaa tulkita sekä lineaariset että epälineaariset vaikutussuhteet.

Avainsanat: Ammatillinen kasvu ja oppiminen, organisaation ilmapiiri, opiskelumotivaatio, attribuutiot, kvantitatiiviset tutkimusmenetelmät, Bayesilainen mallinnus
CONTENT

PREFACE
ABSTRACT
TIEVISTELMÄ

1 INTRODUCTION 15
  1.1 Research Questions .............................................................................................. 19
      1.1.1 Subject-related Research Questions .............................................................. 19
      1.1.2 Methodological Research Questions ............................................................. 20
  1.2 Outline .................................................................................................................. 20

2 PROFESSIONAL GROWTH 22
  2.1 Professional Development .................................................................................... 25
      2.1.1 Organizational Triggers................................................................................. 28
      2.1.2 Work Role Triggers ....................................................................................... 30
      2.1.3 Personal Triggers ........................................................................................... 31
  2.2 Organizational Factors Supporting Professional Development ........................... 32
      2.2.1 Learning Organization ................................................................................... 34
      2.2.2 Empowerment ............................................................................................... 39
  2.3 Modeling of Organizational Atmosphere ............................................................. 42
      2.3.1 The Initial Model of the Growth-oriented Atmosphere ................................ 43
  2.4 Growth-oriented Atmosphere Questionnaire (GOAQ) ........................................ 45
      2.4.1 Relation between the Dimensions of Empowerment and Growth-oriented Atmosphere .............................................................................................. 49
      2.4.2 Operationalization of the First Research Question........................................ 51

3 PROFESSIONAL LEARNING 53
  3.1 Learning as a Process ........................................................................................... 53
  3.2 Self-regulatory Processes ..................................................................................... 56
      3.2.1 Models of Self-Regulation ............................................................................ 57
      3.2.2 Concept of Self-Regulation ........................................................................... 59
      3.2.3 Self-Regulation of Learning ........................................................................ 61
      3.2.4 Empirical Support for Self-Regulation .......................................................... 65
      3.2.5 Self-Regulatory Abilities as Metacompetencies .......................................... 67
      3.2.6 The Development of Self-Regulation ............................................................ 70
  3.3 Attribution Theory ................................................................................................ 72
  3.4 Self-confidence Attitude Attribute Scales (SaaS) ................................................ 74
      3.4.1 Operationalization of the Second Research Question .................................... 74
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.5 Learning Strategies</td>
<td>76</td>
</tr>
<tr>
<td>3.5.1 Metacognitive and Cognitive Strategies</td>
<td>77</td>
</tr>
<tr>
<td>3.5.2 Resource Management Strategies</td>
<td>79</td>
</tr>
<tr>
<td>3.6 Motivation and Volition in Learning</td>
<td>80</td>
</tr>
<tr>
<td>3.6.1 Control of Motivation and Volition</td>
<td>82</td>
</tr>
<tr>
<td>3.7 Integrative Model of Motivation</td>
<td>84</td>
</tr>
<tr>
<td>3.7.1 Value Components</td>
<td>85</td>
</tr>
<tr>
<td>3.7.2 Expectancy Components</td>
<td>85</td>
</tr>
<tr>
<td>3.7.3 Affective Components</td>
<td>86</td>
</tr>
<tr>
<td>3.8 Metacognition in Motivation and Learning</td>
<td>86</td>
</tr>
<tr>
<td>3.8.1 Metacognitive Abilities</td>
<td>88</td>
</tr>
<tr>
<td>3.9 Abilities for Professional Learning Questionnaire (APLQ)</td>
<td>90</td>
</tr>
<tr>
<td>3.9.1 Operationalization of the Third Research Question</td>
<td>93</td>
</tr>
<tr>
<td>4 BAYESIAN MODELING</td>
<td>95</td>
</tr>
<tr>
<td>4.1 Typical Problems of Using Parametric Frequentistic Techniques</td>
<td>96</td>
</tr>
<tr>
<td>4.2 Typical Solutions to Problems of Using Parametric Frequentistic</td>
<td>98</td>
</tr>
<tr>
<td>Techniques</td>
<td>99</td>
</tr>
<tr>
<td>4.3 Introduction to Bayesian Modeling</td>
<td>103</td>
</tr>
<tr>
<td>4.4 The Null Hypothesis Significance Testing Procedure</td>
<td>103</td>
</tr>
<tr>
<td>4.5 Essential Benefits of Bayesian Modeling</td>
<td>111</td>
</tr>
<tr>
<td>4.6 Relation of Bayesian Modeling to the Traditional Parametric and</td>
<td>113</td>
</tr>
<tr>
<td>Non-parametric Frequentistic and Non-frequentistic Techniques</td>
<td>113</td>
</tr>
<tr>
<td>4.7 Bayesian Dependency Modeling</td>
<td>120</td>
</tr>
<tr>
<td>4.8 Bayesian Classification Modeling</td>
<td>124</td>
</tr>
<tr>
<td>4.9 Bayesian Unsupervised Model-based Visualization</td>
<td>130</td>
</tr>
<tr>
<td>4.10 Bayesian Modeling in Practice</td>
<td>135</td>
</tr>
<tr>
<td>4.10.1 B-Course</td>
<td>135</td>
</tr>
<tr>
<td>4.11 Operationalization of the Fourth and Fifth Research Questions</td>
<td>137</td>
</tr>
<tr>
<td>5 OVERVIEW OF THE ORIGINAL STUDIES</td>
<td>139</td>
</tr>
<tr>
<td>5.1 Study I</td>
<td>139</td>
</tr>
<tr>
<td>5.1.1 Goals</td>
<td>139</td>
</tr>
<tr>
<td>5.1.2 Participants and Procedure</td>
<td>139</td>
</tr>
<tr>
<td>5.1.3 Measures</td>
<td>140</td>
</tr>
<tr>
<td>5.1.4 Data analysis</td>
<td>140</td>
</tr>
<tr>
<td>5.1.5 Results</td>
<td>140</td>
</tr>
</tbody>
</table>
LIST OF ORIGINAL PUBLICATIONS


---

1 Reprints were made with the kind permission of the publishers.

2 First author was the principal designer, analyst and writer of the article. The second author contributed to the writing of the theoretical background.

3 Authors designed the study together. First author was the principal analyst and writer of the article. Other authors collected the data and commented the manuscript.

4 First author was the principal designer, analyst and writer of the article. The second author participated in writing the sections 2.2 and 3.2.


1 INTRODUCTION

An individual who always tries to represent his certainty by a subjective probability distribution is referred to as a “Bayesian”.

(Leonard & Hsu, 1999, p. 5)

This article dissertation presents a Bayesian approach to the modeling of organizational level differences in professional growth and individual level differences in professional learning with empirical samples. Rationale for the study is based on my three major research areas during the last ten years while working at the Research Centre for Vocational Education (RCVE, University of Tampere, Finland): Firstly, research with self-report instruments on work-related atmosphere\(^1\) in various business and public section organizations (chapter 2); Secondly, research with self-report instruments on student’s self-regulated learning in various universities and polytechnic institutions of higher education\(^2\) (chapter 3); Thirdly, methodological research questions mainly involving application and comparison of linear and non-linear statistical techniques (chapter 4).

The major goal of this research is two-fold: Firstly, to contribute to the basic research on professional growth and learning, and secondly, to contribute to the development and use of quantitative research techniques in aforementioned areas.

In the very core of this work are the four empirical studies conducted between 2002 and 2007. The first and last of the four studies model professional growth, and the second and third studies model professional learning. Next, I will shortly introduce the key concepts and outline of this study.

Beairsto (1996, p. 94) suggests that “‘professional growth’ and ‘professional development’ should be seen as complementary processes rather than synonyms.” According to him, professional growth relates to broadening of expertise and professional development describe the process of moving into new areas of knowledge or ability.

\(^{1}\) Term ‘atmosphere’ is a synonym to term ‘climate’, former is used in this study.

\(^{2}\) ‘polytechnic institution of higher education’ a.k.a. ‘polytechnic’, ‘vocational high school’ or ‘university of applied sciences’ (fi: ‘ammattikorkeakoulu’: se: ‘yrkeshögskola’)
‘Professional growth’ is an important concept in this study as it mirrors how employees feel their knowledge and skills are valued and to what extent they are committed to work and organization. Theoretical model of growth-oriented atmosphere (Ruohotie, 1996b) and its operationalization, the Growth-oriented Atmosphere Questionnaire (GOAQ, Nokelainen & Ruohotie, 2008) provide one way to investigate the current state of professional growth in an organization. The first and last studies presented in this dissertation investigate the dimensions of growth-oriented atmosphere with empirical samples collected from employees of Finnish companies and polytechnic institutions of higher education.

‘Professional learning’ is at its best a continuous self-regulatory process. Self-regulation means that an individual has a control over his/her learning, that is, possibility to regulate his/her motive, method and time consumption in learning process (Zimmerman, 1998). In this dissertation, such self-regulatory abilities enhancing professional learning are seen as conative constructs including motivational (e.g., achievement and career orientations) and volitional (e.g., action controls and personal styles) issues (Ruohotie, 2000a). The theoretical concept of professional learning is operationalized with the Abilities for Professional Learning Questionnaire (APLQ, Ruohotie, 2000c). The second and third of the studies presented in this dissertation investigate professional learning with empirical samples collected from employees of Finnish companies, polytechnic institutions of higher education, universities and elementary schools.

‘Bayesian modeling’ (a.k.a. ‘subjective probability’, ‘non-linear modeling’, e.g., Gill, 2002) is defined as a viable option to both ‘frequentistic’ (a.k.a. ‘quantitative’ or ‘Gaussian’) and ‘linear’ approaches. In all four studies, the empirical data is collected with self-rated five-point Likert-scale questionnaires, and analyzed both with linear frequentistic and Bayesian techniques. The last empirical study in this dissertation compares the Bayesian

---

3 I prefer using the term ‘Likert scale’ (DeVellis, 2003, p. 78-80), named after Rensis Likert (Wainer et al., 2000, p. 243), instead of ‘summated rating scale’ (SRS). According to Johnson and Christensen (2004, p. 132), “another name for a summated rating scale is a Likert Scale because the summated rating scale was pretty much invented by the famous social psychologist named Rensis Likert.”
and frequentistic approach in a way they set assumptions for the data and how much the subjective interpretations drawn from the results vary.

Viability of the Bayesian approach when compared to the frequentistic techniques is the most obvious in situations where it is not possible to collect vast amounts (i.e., sample sizes of 100, 200, 500,…) of data. Viability of the Bayesian approach when compared to the linear techniques is easily addressed by stating that it is capable of understanding non-linear dependencies between observed variables (Nokelainen & Tirri, 2004).

For example, if we have a group of 23 polytechnic institution of higher education headmasters, and we are interested in their metacognitive abilities (i.e., knowledge and regulation of thinking and learning), it sounds more relevant to collect the evidence, analyze it and report the results than to abandon the research problem by saying that ‘we do not have enough data to feed our frequentistic statistical algorithms’. It should be up to the researcher, and not to the statistical technique or research technique, to decide what kind of research question is worth investigation. It should also be up to a research question of scientific interest set by the researcher, and not to the limitations of certain statistical technique or research technique we are familiar with, to dictate what kind of evidence is worth further investigation.

How, then, should we collect the data from the headmasters? Both target (polytechnic institution of higher education headmasters) and sample populations (those headmasters that are willing to participate in the study) are evitable too small for any convincing (parametric or non-parametric) frequentistic approach. At this point a frequentistically oriented researcher starts to look for a larger sample: An international sample of headmasters would surely be larger than a Finnish sample alone? Or, what if we ask headmaster’s subordinates’ views on their superior’s metacognitive abilities? Those thoughts - and a long list of others not worth mentioning here – have crossed my mind dozen times during my ‘frequentistic years’ as an applied statistician, about ten years ago. When I was introduced to Bayesian techniques I learned that they allow, for example, small sample size (theoretical minimum is zero) and even qualitative indicators (i.e., an interview or narrative story coded into classes, Likert–scale questionnaire). With
this new door opened, the decision making process to answer the question becomes more straightforward.

First, I would consider if the phenomenon under investigation (headmasters metacognitive abilities) requires a special data collection approach. A questionnaire is a relevant data collection tool in this case only if we a) believe that metacognition is operationalized successfully in an existing survey, b) are ready to accept self-rated evaluations from our participating headmasters and c) are able to use a non-frequentistic technique. Other possible data collection methods are, for example, an interview, biographical research and observation. Sometimes we even have resources for multimethod approach (triangulation) where all or some of the abovementioned methods are applied together.

Second, I would consider if the phenomenon under investigation is stable or prone to change by nature. If it is considered to be a stable one, I would collect the data from just one group of headmasters at one time point (correlational study). Otherwise, I would like to study the effect of intervention (e.g., consultation) with a control group (quasi-controlled experiment).

Third, I would think who are the most valid informants to provide the evidence needed to answer the research question. If only headmasters will do, the options are limited to ‘qualitative’ (e.g., interview, autobiography) and ‘non-frequentistic’ (e.g., neural networks, Bayesian statistics or fuzzy logic) approaches. If headmasters’ subordinates are valid informants, then also ‘quantitative’ Bayesian and ‘qualitative’ techniques share a lot of common ground as researchers and informants’ subjective experiences and other relevant a priori evidence play a central role in both techniques when interpreting the results. This point is further elaborated in chapter four.

Each of the four original studies have their own specific research questions and tasks which are discussed in detail in chapter five. Next, however, I will introduce the five ‘metalevel’ research questions that connect the four studies together.
1.1 Research Questions

First three research questions tackle the problems of professional growth and learning. The first one is related to the dimensionality of the concept of growth-oriented atmosphere (professional growth). The second research question is about the differences in self-regulative sub processes (professional learning) of three groups of mathematically gifted students. The third research question studies the dimensionality of learning motivation (professional learning).

Two latter research questions form what I call ‘the Bayesian core’ of this dissertation. Generally speaking, they ask if Bayesian modeling is a viable option for traditional linear analysis techniques. First, Bayesian dependency modeling is compared to linear correlations. Second, Bayesian classification modeling is compared to linear discriminant analysis.

After the research questions, I will present a table showing how all the theoretical and methodological aspects discussed above are covered in the four studies.

1.1.1 Subject-related Research Questions

(RQ 1) What is the optimal number of dimensions and items in the Growth-oriented Atmosphere Questionnaire (GOAQ) to describe the theoretical model of growth-oriented atmosphere? (Study I)

(RQ 2) Are the theoretical dimensions of the Self-confidence attitude attribute Scales (SaaS) questionnaire identified in the domain of three groups of mathematically gifted participants: Academic mathematics Olympians, polytechnic institute of higher education students, and elementary school students who have participated in mathematical competitions? (Study II)

(RQ 3) What is the optimal number of dimensions and items in the Abilities for Professional Learning Questionnaire (APLQ) to describe the theoretical model of learning experiences and motivation? (Study III)
1.1.2 Methodological Research Questions

(RQ 4) Is there a difference between substantive interpretations of the results of Bayesian dependency modeling and linear bivariate correlational analysis with professional growth data that has both linear and non-linear dependencies? (Study IV)

(RQ 5) Is there a difference between substantive interpretations of the results of Bayesian dependency modeling and linear confirmatory factor analysis with professional growth data that has both linear and non-linear dependencies? (Study IV)

Table 1 shows how all the theoretical and methodological aspects are considered in the four studies. The row sum of total responses equals 4777, but after omitting the duplicate entries (Häme polytechnic institute of higher education personnel, \( n = 447 \), participated in both the first and third study), the total number of respondents in this dissertation becomes 4330.

1.2 Outline

This dissertation is organized as follows. In the second chapter, I will describe the prerequisites for professional growth in both individual and organizational level. The third chapter is about the processes of professional learning involving, for example, self-regulation. The fourth chapter is an introduction to the Bayesian modeling. I will give a short overview to the Bayesian modeling approach, present its main features and compare them to the features of frequentistic (Gaussian) modeling. Supervised and unsupervised Bayesian modeling approaches are discussed. The fifth chapter presents an overview of the four original studies that were chosen to represent Bayesian modeling of professional growth and learning in this study. In the sixth chapter I will discuss the general findings and validity of this dissertation.
Table 1. Summary of the Major Research Goals and the Empirical Data of the Four Original Publications

<table>
<thead>
<tr>
<th>Theoretical aspect</th>
<th>Study I</th>
<th>Study II</th>
<th>Study III</th>
<th>Study IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>PG</td>
<td>RQ 1</td>
<td>—</td>
<td>—</td>
<td>X</td>
</tr>
<tr>
<td>PL</td>
<td>—</td>
<td>RQ 2</td>
<td>RQ 3</td>
<td>—</td>
</tr>
<tr>
<td>Methodological aspect</td>
<td>BCM</td>
<td>—</td>
<td>X</td>
<td>—</td>
</tr>
<tr>
<td>BDM</td>
<td>—</td>
<td>—</td>
<td>X</td>
<td>RQ 4 &amp; 5</td>
</tr>
<tr>
<td>BUMV</td>
<td>X</td>
<td>—</td>
<td>X</td>
<td>—</td>
</tr>
<tr>
<td>Empirical aspect</td>
<td>COM (N)</td>
<td>—</td>
<td>—</td>
<td>2430</td>
</tr>
<tr>
<td>POL (N)</td>
<td>447</td>
<td>74</td>
<td>971</td>
<td>726</td>
</tr>
<tr>
<td>UNI (N)</td>
<td>—</td>
<td>77</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>SCH (N)</td>
<td>—</td>
<td>52</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Total (N)</td>
<td>447</td>
<td>203</td>
<td>971</td>
<td>3156</td>
</tr>
</tbody>
</table>

Note. Study I = Nokelainen & Raudsepp, 2008. Study II = Nokelainen, Tähti & Merenti-Vähämaa, 2007. Study III = Nokelainen & Raudsepp, 2002. Study IV = Nokelainen, Silander, Raudsepp & Tähti, 2007. RQ 1 = What is the optimal number of dimensions and items in the Growth-oriented Atmosphere Questionnaire (GQAQ) to describe the theoretical model of growth-oriented atmosphere? RQ 2 = Are the theoretical dimensions of the Self-confidence attitude attribute scales (SaaS) questionnaire identified in the domain of three groups of mathematically gifted participants: Academic mathematics Olympians, polytechnic institute of higher education students, and elementary school students who have participated in mathematical competitions? RQ 3 = What is the optimal number of dimensions and items in the Abilities for Professional Learning Questionnaire (APLQ) to describe the theoretical model of learning experiences and motivation? RQ 4 = Is there a difference between substantive interpretations of the results of Bayesian dependency modeling and linear hierarchical correlational analysis with professional growth data that has both linear and non-linear dependencies? RQ 5 = Is there a difference between substantive interpretations of the results of Bayesian dependency modeling and linear confirmatory factor analysis with professional growth data that has both linear and non-linear dependencies?

PG = Professional growth, PL = Professional learning, BCM = Bayesian classification modeling, BDM = Bayesian dependency modeling, BUMV = Bayesian unsupervised model-based visualization, COM = Finnish company employees, POL = Finnish polytechnic institution of higher education staff or students, UNI = Finnish university students, SCH = Finnish elementary school students.
Professional growth is a continuous learning process that enables individuals to acquire the knowledge, skills and abilities needed to cope with changing demands for vocational proficiency throughout their career (London & Mone, 1999). It is, thus, viable to speak about ‘professional career growth’ to distinguish it from the concept of ‘professional development’, which is a collection of concrete developmental strategies and functions that aim to support professional growth. However, these two concepts appear in the research literature intertwined in the form of ‘professional growth and development’. This is natural, as the professional development is *de rigueur* but not *de facto* for professional growth.

London and Mone’s term ‘continuous’ describes strong and durable need or will to learn and also valuation of learning. ‘Learning’ refers in this context individuals will to develop one’s skills via practice and training in order to meet changing challenges of the work. Naturally in most cases such development is possible only if employer/organization shares the same goals. Continuous learning is characteristic of multifaceted career that is in fact defined as growth of know-how (Ruohotie, 1999). People may have at the same time several career paths or consecutive work periods in different companies or even in different professions.

Beairsto and Ruohotie (2003, p. 115) define ‘lifelong learning’ as a “continuous process for professionals who wish to remain abreast of developments in their field and thus retain, or expand, their competence and/or qualifications over the course of a long career”. They state that lifelong learning

---

1 ‘Growth’ is a valid concept for internal evaluation, unlike the concept of ‘development’ that should be evaluated externally (King, 2002; Korpelainen, 2005). Professional growth is a personal process, and thus, comparable to the concept of ‘personal growth.’

2 Ellström (1994) defines *competence* as the potential capacity of an individual or a collective to successfully deal with certain situations or complete a certain task or job according to certain formal or informal criteria set by someone else. He (2001) defines *qualification* as a competence.
is not an option but a necessity in professional fields such as education, health, engineering and law. The pervasive and rapid change that is characteristic of our era has made lifelong learning a necessity in business and other walks of life as well. Thus, the question of how to inculcate the skills and dispositions of lifelong learning during pre-service training, and how to sustain it throughout a career, is vitally important in all areas of vocational education.

Beairsto and Ruohotie (ibid.) further state that one of the most critical periods of lifelong learning comes during professional training. Vocational education institutions need to design their curriculum not only to teach traditional understandings but also to prepare their students for the challenge of maintaining and enhancing competence after graduation. There is an enormous amount of work to be done to ensure that the foundations of lifelong learning are laid in the initial stages of professional education.

In order to be successful, educational organizations must provide effective professional development programs for employees over the entire course of their careers (Lawler, 1994). This notion makes studies of professional updating, and especially those concerning the problems and prerequisites of continual growth in various work communities, most important. These include factors within the individual, the job, the work place and society (e.g., Kauto-Koivula, 1993; Lahti-Kotilainen, 1992; Luoma, 2001; Mäki, 2000; Rantanen, 1997; Ruohotie, 1992, 1996a, 1996b).

Continuous development and updating of skills is important, otherwise they may become useless (or at least obsolete) for the work life. Kaufman (1974, p. 23) has defined the professional obsolescence as “the degree to which professionals lack the up-to-date knowledge and skills necessary to maintain effective performance in either their current or future work role.” According to Pazy (2004), professional updating is a learning response to imminent obsolescence. Like other forms of adult learning beyond formal education, updating is characterized by a problem focus (Knowles, 1990), and it is typically a self-initiated, self-structured, and self-defined activity.

that is actually required by a work task and/or is implicitly or explicitly determined by individual qualities.
Dubin (1990) has studied factors that accelerate this progress and found several indicators that make continuous professional skill development and renewal of professional knowledge essential.

*Rapid creation of knowledge:* Knowledge is being created at an accelerating rate and this new knowledge is displacing the old. Scientific and technical knowledge doubles about every seven years.

*Complexity of knowledge:* The acquisition of new knowledge is complicated by the fact that traditional boundaries are blurring. Development or perceptions in one area may be beneficial in a number of ways in another as disciplines, such as physics and chemistry or psychology and data processing, begin to overlap significantly.

*Technological innovation:* Technological innovation is clearly the process that is primarily responsible for the aging of knowledge. Estimates suggest that in the next 10 to 20 years more technological change may occur than throughout all of previous history.

*Global competition:* At the same time that nations are investing in research and development on an increasing scale, they are looking for management techniques, which will encourage innovation and, through training and other development functions, ensure the productive use of human resources. Restructuring, automation, new methods of management, increased allocation of funds to product development, and radical reductions in the time required for the development of new products are the result.

According to Ruohotie (1996a), these trends, and others, which could be included, affect the economic, social and political environment in fundamental ways. In these situations standing pat means falling back. Both organizations and individuals must respond quickly to changing circumstances in order to retain their competence.

Fossum and Arvey (1990) have developed a hypothetical model which classifies the various factors that can cause an individual’s competence to become obsolete, including motivational factors, individual and organizational factors, and factors outside the work environment. According to this model, new technologies, the setting of new goals, the application of new procedures, and the
changing of the structure of an organization result in changes in tasks, duties and obligations.

Change itself, of course, does not necessarily create obsolescence and change does not always require a response. In fact, automatic adoption of every new development and immediate response to every new idea is no healthier than intransigent self-satisfaction. A problem only exists when changes in tasks and duties result in an employee no longer being able to cope with his/her existing level of professional responsibility.

2.1 Professional Development

Professional development includes all developmental functions that are directed at the maintenance and enhancement of professional competence, and, thus, professional growth. Updating is ideally a continual, lifelong process that addresses such goals as the acquisition of new and up-to-date information, the development of skills and techniques and the elevation of one’s personal esteem (Ruohotie, 1994a; Niemi & Kohonen, 1995).

This is in parallel with Day’s teachers professional development (1999, p. 4) that “… consists of all natural learning experiences and those conscious and planned activities which are intended to be of direct or indirect benefit to the individual, group or school, which contribute, through these, to the quality of education in the classroom. It is the process by which, alone and with others, teachers review, renew and extend their commitment as change agents to the moral purposes of teaching; and by which they acquire and develop critically the knowledge, skills and emotional intelligence essential to good professional thinking, planning and practice with children, young people and colleagues throughout each phase of their teaching lives.”

The maintenance and enhancement of competence is subject to the combined effect of many factors, ranging from personal traits to salient features of the work environment (Fishbein & Stasson, 1990; Kozlowski & Farr, 1988). In Dubin’s (ibid.) analysis, the most important single variable is found to be personal motivation. According to Dubin’s model, participation motivation for a development activity is high if the individual believes that 1) Participation will truly increase his/her ability to cope better in present and future tasks than in
previous ones; 2) Professional competence level affects the amount of rewards and products attainable; 3) Behavioral change leads to the desired results; for example, raise in pay, promotion, approval, increase in prestige, recognition by employer, decrease in routine tasks, increase in free-time.

Farr and Middlebrooks (1990) have examined motivation using the concepts of expectancy theory. In this theory, ‘expectancy beliefs’ refers to the belief that participation in professional development will result in becoming or remaining professionally current; ‘instrumentality beliefs’ refers to the belief that remaining current increases professional status and rewards; and ‘outcome valences’ refers to the strength of the professional’s attachment to the consequences of participation in professional development activities (e.g., promotion, task assignment, praise). Expectancy beliefs, instrumentality beliefs and outcome valences combine in a multiplicative fashion to affect motivation for professional updating.

Expectancy beliefs, instrumentality beliefs and outcome valences are dependent upon, among other things, personalities, motive structures, present circumstances and previous experience. Any one of these may have a significant effect on motivation to learn and/or to participate in professional development programs (Ruohotie, 1993). That is why people working in the same job and in the same circumstances can have different opinions of the utility, necessity and meaning of updating.

Maurer and Tarulli (1994) have identified the following factors affecting the voluntary involvement of workers in development activities.

*Perceptions related to the working environment:* Important factors here include a general organizational orientation towards personnel development and the support of various individuals in the work place, particularly supervisors.

*Perceptions and beliefs regarding the benefits of development:* These perceptions affect participatory motivation, the degree of participation and variations in selected development activities. Influences that affect participation can be divided into internal and external. Examples of internal influences are interesting and challenging work and the tasks themselves. External influences include increased remuneration, promotion and material benefits.
Values and judgments: Perceptions relating to the environment and the benefits of development may have different significance for different people. For example, senior management support may be critical to certain people’s participation and unimportant to others’. Generally speaking it can be said that the more important a given factor is to the individual, the more forcefully it directs his/her actions.

Personality factors: Many personality factors influence learning and development activity. They include: Identification with work, which shows the extent to which an individual assimilates to work/job and how important work is in the individual’s life; personal concept of career, or how the individual has assessed his/her strengths and weaknesses in relation to career; the need for self-development, or how necessary development in skills needed for the job is seen to be; and self-efficacy, or how confident the individual is of his/her ability to learn new skills.

Ideally an individual would know that he or she has the opportunity to grow continuously and to develop personally in his/her work. Tasks and positions would be seen as a continual course that he or she feels is advancing or at least broadening. In practice, however, growth often stops or plateaus at some point, with the result that work motivation decreases, work loses its meaningfulness and the handling of tasks becomes routine. Once a person has settled comfortably into his/her professional role and career routine, one must ask what factors will trigger a change in this routine and create a desire to learn something new.

Hall (1986) has created a model of mid-career sub-identity development which outlines factors that influence professional development (growth triggering factors) and the process through which the professional exploration cycle progresses. It shows that professional growth is dependent on the social and institutional context as well personal attributes and circumstances. Several factors are presented which can trigger in career routine and lead to the acquisition and development of new knowledge and skills.
2.1.1 Organizational Triggers

Changes in organizational structure, areas of responsibility and tasks often require the development of new skills. Individuals respond to such changes both effectively and behaviorally according to their perception of their circumstances, interpreting environmental events or situational change on the basis of personal values and perspectives.

Environmental attributes may be the same for two individuals, and their descriptive perception of them identical, while the valuations associated with them are significantly different. One person may consider a task to be ‘non-challenging’ while the other perceives it as ‘challenging’. Such individual valuations have an effect on affective responses since, together with psychological arousal; they give rise to subjective emotions. Psychological responses are relevant cognitions in emotional terms because they influence the individual’s perception of the extent to which the working environment is conducive or detrimental to personal welfare.

Research conducted as part of the Growth Needs Project shows that the following factors are among the keys to the creation and maintenance of growth and high innovative capacity in an organization (Ruohotie, 1994a).

Creation of a supportive culture: In a supportive environment innovation becomes a natural part of everyday work. Tasks may be intentionally defined in broad terms, encouraging change and emphasizing the possibility of choice. Cooperation is an explicit objective since active interpersonal communication tends to give rise to novel or unanticipated situations, promote elaboration of ideas and support complexity. Because the constant need to seek authorization from others smothers innovation, hierarchical structures are replaced by decentralized networks emphasizing equality and democracy.

Reward of development: In innovative organizations learning, initiative and experiment are prized as inherently valuable. A new task, a new challenge or a new opportunity is seen as desirable in itself, independently of any utilitarian reward that it might offer.

Supportive and participative management: In innovative organizations it is seen as the duty of management to create a workplace in which each individual
can reach his/her full potential. Activation of opportunity, inspiration of a collective vision, enabling of activity and the provision of role-models are characteristic of a supportive and participative management style in which orders, decisions and supervision may be replaced by coaching, training, activation and enablement.

**Intensive communication:** The more intensive the communication, the more effectively new ideas and alternative points of view can be shared and developed. Mechanisms for the rapid dissemination of ideas and for immediate feedback are essential to intensive communication.

**Security:** In an era of intensifying competition, the organizations that will survive and succeed are those in which there is a secure and confident atmosphere for employees. The fear of failure, of blame or of criticism is an effective damper to creative innovation. Familiar procedures and established standards must be open to question if creative capacity is to be unleashed.

Generating continuous enlivening innovation requires at least two things from an organization: First, it must learn to fully develop and utilize the capacity of its personnel, and second, it must show imagination at all times, suspending judgment temporarily when necessary in order to promote the development of new ideas.

These requirements have considerable implications for management style, which must create the conditions in which innovative behavior is possible and then nurture it when it occurs (e.g., Beairsto, 1997). Professional development is stifled by excessive pressure for production and by minimization of risk-taking. Such attitudes encourage people to take refuge in routines that have proven successful in the past, a barrier to innovation.

The overall organizational atmosphere can encourage growth and specific management practices can provide additional impetus, for example, by rewarding managers for supporting the development of their subordinates, rotating personnel into new situations where they have an opportunity to learn new skills, and so on (Hall, 1986). If the organizational culture emphasizes learning, and management practices support the systematic development of individuals, any latent potential for professional growth and development is more liable to be realized.
2.1.2 Work Role Triggers

The nature of work itself can provide opportunities for intellectual development and the way in which work is structured may increase these opportunities. For example, depending upon the individual’s personal characteristics and needs, he or she may be given more room for independence through a loosening of restrictions and requirements (a change in the character of the work) or, conversely, the individual’s uncertainty as to his/her task may be reduced by greater precision in job description and responsibility (Ruohotie, 1993). In an educational context, the way in which teachers and administrators are encouraged to view curriculum and their own role may lead them to perceive instruction as either repetitive or boring delivery of information, which is discouraging, or as variable and creative construction of understanding, which is motivating.

People respond positively to growth opportunities that correspond to their personal growth needs (Loher, Noe, Moeller & Fitzgerald, 1985; Spector, 1985; Graen, G., Scandura & Graen, M., 1986; Champoux, 1991). If they have strong growth needs they will react positively to growth opportunities, whereas if they have weak needs they will probably react neutrally or even negatively to an open-ended challenge. Thus, growth opportunities must be provided in light of individual characteristics, which may vary over time and according to circumstances.

On-the-job learning is most likely to occur when a worker is encountered with a challenge that includes “improving components” (McCauley, Ruderman, Ohlott & Morrow, 1994). Such challenges are beneficial as they provide learning possibilities as well as motivate to learn. Problem solving connected to work is the most productive learning experience for many professionals. Learning takes place when the learner is faced with demands he finds new or with situations that are conflicting. Learner’s motivation may be due to a will to improve his/her working skills or a desire to achieve production based bonuses. It may also come from a necessity to avoid negative results or unpleasant situations.

McCauley and his colleagues (1994) classify challenges at work into three main categories: 1) Job transitions; 2) Task-related characteristics; 3) Obstacles. Job transitions mean changes in working role such as work content or
responsibility, professional status or changes in the workplace. *Task-related characteristics* refer to the various problems and difficulties that the individual runs into at work. They include the demand to create change (developing new directions, solving inherited problems, reduction decisions and problems with subordinates), a high level of responsibility (meeting high demands, managing diversity, overload and meeting outside pressure), and handling individuals over whom the person has no authority. *Obstacles* may be set by a superior, lack of management support or adverse business conditions.

### 2.1.3 Personal Triggers

Events or stages connected to everything from personal factors to life changes can cause an individual to reconsider his/her career priorities and goals. In addition, certain personal characteristics predispose an individual to make changes in order to avoid the negative consequences of work pressure or deal with personal frustration at the status quo (e.g., basic personality disposition, motivation for advancement, initiative, stress on performance, hardiness, flexibility, tolerance of ambiguity, independence) (Hall, 1986; Kautto-Koivula, 1993; Ruohotie, 1993, 1996a).

An important prerequisite for development is the ability to cope with challenging goals. According to Hall (1971), achieving challenging goals increases an individual’s motivation and desire to strive for success in the future - a psychological cycle of success is established. Coping leads to an intrinsic experience of success and progress that in turn enhances a person’s self-esteem and competence. Hall has shown (1986, 1990a; Hall & Goodale, 1986) that intrinsic rewards strengthen one’s commitment and direct later goal choices.

Learning to learn involves the control and application of various learning strategies, skill in practical thought (including creative-critical-analytical thought and logical decision-making), the ability to control resources (such as use of time, learning conditions and one’s own strengths), the ability to apply knowledge in new circumstances, and skills in various problem-solving methods. Innovative organizations consciously seek to develop the learning skills of their key personnel and thus to increase their inclination to growth and ability to take advantage of the existing growth opportunities (Ruohotie, 1994b, 1996c).
Growth motivation, learner’s desire to develop and experience new things, is another key factor in professional development (Ruohotie, 1985, 1994a). London and Mone (1987) have studied growth motivation with specific reference to a career and explained the concept of career motivation. Application of their analysis to the Growth Needs Project shows that personal factors such as self-reliance, need to perform, willingness to take risks, setting of career goals and self-awareness as to weaknesses and strengths, together with commitment to work, correlate positively with active participation in development of professional skills, promotion and professional identity (Ruohotie, 1993, 1994a).

2.2 Organizational Factors Supporting Professional Development

Many different kinds of learning are involved in professional growth and development. Hall (1986, 1990a) classifies these different types according to whether task learning or personal learning is in question and whether the goals are short-term or long-term (Table 2).

| Table 2. Types of Learning Involved in Professional Growth and Development (Hall, 1986) |
|-----------------------------------------------|---------------------------------------------------------------|
| Task learning | Personal learning |
| Short term  | Improving performance-related knowledge, skills and abilities | Resolving issues regarding attitudes toward career and personal life |
| Long term  | Improving adaptability | Developing and extending identity |

Hall’s research (1986, 1990b) shows that most development activities are directed towards short-term rather than long-term learning goals, and they are directed at task learning rather than personal learning. According to Ruohotie (1996b), professional growth and development should be conceptualized in terms of all four quadrants in Hall’s schema and, in general, greater attention should be paid to long-term goals and personal learning dimensions within professional programs.
In addition to program characteristics, individual attributes and personal factors external to work, factors within the work environment are important determinants in motivating and supporting professional development (Baldwin & Ford, 1988; Tannenbaum & Yukl, 1992). Factors that support, or discourage, participation in professional development activity and the more general quest for continuous learning may be direct or indirect, obvious or hidden, intentional or unintentional.

Tracey, Tannenbaum and Kavanagh (1995) describe a culture that supports continuous learning in terms of five characteristics: 1) Each individual is personally accountable for renewal of knowledge and skills required at work; 2) The working community (i.e., colleagues, supervisors, teams, etc.) supports development of skills; 3) The working community has at its disposal operating systems which reward on the basis of performance and encourage or enable professional growth and development; 4) The community supports innovation and competition; 5) The members of the organization view learning as an important part of daily work life.

Further, according to Ruohotie (1994a), challenging, varied and independent work motivates professional growth. The challenge of work is related to matters such as required level of technical skills and knowledge, uncertainty about how to reach goals and the level of self-determination in one’s work; for example, the ability to determine one’s work methods. Making work challenging may not only encourage participation in discrete professional development programs but, if combined with opportunity and support for reflection, may provide opportunities for learning and growth right within the task at hand.

Individual growth and development can also be supported through an individual reward system in some settings. In order for the system to be effective members of the organization must understand performance goals, have the ability to personally influence achievement of the goals, and have realistic expectations for rewards based on that achievement.

In Table 3, Ruohotie (1996b), on the basis of Dubin’s (1990) analysis of technical experts’ personal views, outlines the characteristics of a work community that supports growth. Factors have been categorized in terms of work community.
or work organization, supervisor and worker relations, organizational atmosphere, workplace peer-to-peer relationships and management policies and practices. While not all of these factors would apply equally in an educational setting and not all are equally controllable, they do lead us towards professional growth and development. (Ruohotie, 1996b.)

One central purpose of the Growth Needs Project conducted in Research Centre for Vocational Education (University of Tampere, Finland) since 1990 has been to determine what skills are required by supervisors to encourage development of professional competencies amongst their staff and create a growth-oriented working environment. On the basis of research results so far the following generalizations can be made (Ruohotie, 1994a). A supervisor’s view of people affects his/her management style. Development-minded and encouraging supervisors believe that people have an innate desire and ability to develop. Further, a supervisor’s view of his/her career motivation affects his/her personal achievements and ability to create conditions for personal growth and development by others in the working group. Creation of a growth-oriented atmosphere requires of the management solid managerial skills, self-confidence, high growth motivation, innovativeness and an interest in problem solving.

2.2.1 Learning Organization
Professional growth and development encompasses all four quadrants in Hall’s analysis (see Table 3). It is not intended simply to ensure that an organization survives through what Senge (1990) terms ‘adaptive learning’, but to actively expand its capacity and define its future through ‘generative learning’. Senge uses the term ‘learning organization’ to describe the ideal situation which can be created through the integration of five ‘component technologies’.

**Personal Mastery:** This refers not to the honing of professional skills but to the personal habit of continually clarifying individual vision and aspirations so as to take charge of and responsibility for individual actions.

**Mental Models:** Becoming conscious of their personal assumptions and inner ‘mental models’ allows individuals in an organization to examine those assumptions critically and to expose their thinking to others so that they can influence and/or learn from it.
Building Shared Vision: Translating personal visions into genuinely shared visions generates individual commitment and creates a coherence that increases organizational effectiveness.

Team Learning: Through dialogue individuals can create a synergy that allows groups to discover understandings and accomplish goals that the individuals could not attain on their own.

Systems Thinking: When an organization is understood as a system, the interdependent relationships between the parts become as important as the parts themselves and one begins to see that changing any one thing eventually changes everything. (ibid.)

A learning organization supports both individuals and teams in continuous learning and improvement of their performance (Ruohotie, 1996a). This activity is energized by strong values and clear vision. A learning organization functions close to its consumers, responds rapidly to change, learns from others, continuously questions its methods of operation, accepts error and learns from it. A learning organization, in contrast to a survival strategy organization, is able to maintain process that ensures the creation of new know-how.

The learning organization is a community that is willing and able to question its habits and its outcomes, thus providing creative conflict. Questioning is the positive catalyst for change. This requires a culture within the organization that sees questioners as a prerequisite for and not an obstacle to development.

Brown, Hitchcock and Willard (1994, p. 217-227) present various strategies that can assist in the promotion of a learning organization’s development. They divide the strategies into three groups: 1) Organizational strategies; 2) Management strategies; 3) Team strategies.

The development of organizational strategies fosters learning, including 1) Creation of a learning infrastructure; 2) Promotion and support of experimentation; 3) Empowerment of employees.
<table>
<thead>
<tr>
<th>Job Challenge and Work Assignments</th>
<th>Supervisory Behaviors</th>
<th>Organizational Atmosphere</th>
<th>Peer Interactions</th>
<th>Management Policy and Practices</th>
</tr>
</thead>
<tbody>
<tr>
<td>Job assignments are challenging</td>
<td>Supervisor recognizes good work</td>
<td>Organization fosters use of cutting edge knowledge</td>
<td>Peers are up to date in their knowledge</td>
<td>Financial gain is tied to personal competence</td>
</tr>
<tr>
<td>Work assignments include the latest knowledge in the field</td>
<td>Supervisor recommends promotions based on competence</td>
<td>Innovations are encouraged</td>
<td>Peers often share new useful articles from professional journals</td>
<td>Management supports in-house staff development</td>
</tr>
<tr>
<td>Job rotation encouraged</td>
<td>Individual professional development needs are recognized</td>
<td>Organization rewards professionals for their competence</td>
<td>Peers are willing and able to suggest new ideas to each other</td>
<td>Management provides professionals with up to date equipment</td>
</tr>
<tr>
<td>The professional participates in relevant discussions</td>
<td>Independent thinking is encouraged</td>
<td>Organization holds high professional expectations</td>
<td>Peers are willing to act as a sounding board for others</td>
<td>Management has a systematic job rotation strategy</td>
</tr>
<tr>
<td>The professional is allowed to see a project through</td>
<td>Supervisor seeks solutions from staff</td>
<td>Organization is a leader in professional development</td>
<td></td>
<td>Management maintains a current professional library</td>
</tr>
<tr>
<td>The professional's personal interest is considered when assignments are made</td>
<td>Supervisor provides opportunities for continuing education</td>
<td>Organization has a progressive atmosphere</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job allows free time to explore advanced ideas</td>
<td>Performance reviews point out strengths and weaknesses and offer suggestions for improvement</td>
<td>Organization strives to surpass its competition</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Creation of a learning infrastructure: Learning at work and in the workplace has been found to be more effective in terms of productivity than training separate from the working situation. On the job learning requires a functioning support system, the learning infrastructure, which provides the conditions and opportunities for developing interplay, to allow the exchange of information and the provision of feedback.

Promotion and support of experimentation: An organization that values experimentation and encourages risk-taking tells its employees that new ideas are welcome even if they cannot be seen to immediately enhance productivity.

Empowerment of employees: In traditional organizations, knowledge is equated with power and its sharing reduces this power. The learning organization requires the free sharing of information with all so that it can be used by teams and their members for the benefit of the organization.

Management strategies can support the organizational and team strategies that build a learning organization by 1) Developing a shared vision; 2) Controlling confusion; 3) Being a model learner.

Developing a shared vision: A clear, shared vision should inspire all to commit themselves and work energetically. It is based on collective aspirations, not the charisma of the leader, and helps individuals to accomplish something of significance to them.

Controlling confusion: Confusion usually arises when individuals or a group are confronted with a situation which their knowledge and experience are insufficient to handle. This situation can initiate the learning process and create opportunities for learning and innovative solutions if leaders prevent the confusion from becoming so great that it causes frustration and burnout. Too low a level of uncertainty may also be unproductive, because there is then no perceived need to learn or change.

Being a model learner: A supervisor who works openly on his/her own development is a powerful example. In a learning organization managers may spend as much as 25 per cent of their time on learning activities. Good learners acknowledge their mistakes, realize that any situation can be a learning situation, and know that at some time anyone may be their teacher.
Team strategies can enhance individual and collective learning by 1) Practicing the art of dialogue; 2) Developing reflective skills and habits; 3) Controlling change.

Practicing the art of dialogue: Dialogue as a learning strategy enables people to find new insights that they would not have achieved on their own, thus creating richer understanding of the matters or problems at hand. As Senge (1990) points out, dialogue is not to be confused with discussion. Discussion, which shares its roots with ‘percussion’ and ‘concussion’, is basically a competitive exchange. Dialogue, on the other hand, involves the free flow of ideas based on collaborative, critical thinking.

Developing reflective skills and habits: Argyris (1992) suggests that individuals do not really learn until they are able to reflect critically on their own actions and understand their own cognitive rules and reasoning, which he terms ‘double-loop’ learning. Thus, reflective skills lay the foundation for self-directed growth.

Controlling change: Individuals are increasingly employed by organizations more on the basis of what they can learn than of what they know. Thus, individuals need to develop the skills for self-directed growth that will allow them to shape rather than simply react to change.

Mulford and Silins (2005) investigated in an empirical study involving Australian secondary school students (n = 3500) and their teachers (n = 2500) what leadership practices promote organizational learning. They found that director/head teacher who is transformational, focuses on 1) Providing individual support (e.g., moral support, appreciation for the individual’s work); 2) Culture (e.g., promoting an organizational atmosphere of caring and trust); 3) Structure (e.g., establishing organization structure that promotes participative decision making); 4) Vision and goals (e.g., giving sense of overall purpose to employees); 5) Performance expectation (e.g., having high expectations for employees and expecting them to be effective and innovative); 6) Intellectual stimulation (e.g., encouraging employees to reflect on what they are trying to achieve and how they are doing it). Thus the Leadership for Organizational Learning and Student Outcomes (LOLSO) project’s leadership model is both position based (director/head teacher) and distributive (administrative team and
teacher) stressing on support, care, trust, participation, facilitation and whole staff consensus. (id.)

A learning organization takes advantage of many different development procedures also in problem solving. For example, in decision making adherent to problem solving it is important to rely on knowledge and facts - not presumptions. Literature dealing with leadership presents numerous problem solving techniques or methods and development philosophies such as Total Quality Management (TQM, see Ishikawa, 1985). Scientific methods of problem solving also include problem diagnosis, simple statistical means to organize information and making of decisions (Lindley, 1971).

Transferring knowledge to other parts of organization involves ability to communicate new knowledge quickly and efficiently throughout the organization. Ideas carry maximum impact when they are shared broadly. Written, oral, and visual reports, site visits and tours, personnel rotation and education programs and standardization programs are among the mechanisms available to promote enterprise wide learning.

According to Ashkenes, Ulrich, Jick and Kerr (1995, p. 181-182), organizational learning style can be analyzed with a four component learning model involving learning through: 1) Continuous improvement; 2) Competence acquisition by acquiring new talent or ideas from either inside or outside the company (rotating people into new divisions or buying companies or individuals with certain skills); 3) Experimentation (trying out new ideas or manufacturing tactics right away); 4) Boundary spanning (e.g., sending employees to other countries to learn from their manufacturing technologies).

2.2.2 Empowerment
A wide range of definitions have been used for the concept of empowerment. According to Herrenkohl, Judson and Heffner (1999), components of empowerment include: redistributing authority and control; employees and managers sharing equal responsibility for results; maximizing employees’ contribution to an organization’s success; full participation of workers and leaders in decision making; pursuit of a shared vision and purpose through team effort; self-motivation, which develops through a full understanding of
responsibility and authority commensurate with those responsibilities; the capability to make a difference in the attainment of goals; and a synergistic interaction among individuals, which emphasizes cooperation and leads to expansion of power for the group.

In general, empowerment can be understood as a relational concept that refers to individuals’ personal sense of power and control in relation to others. On the other hand, it can mean sharing power with or redistributing power and control to someone else in a subordinate position. The relational approach to empowerment is based on the idea that creating interdependent and meaningful relations to others improves individual motivation, gives a sense of meaningfulness, and advances self-understanding. (Walsh, Bartunek & Lacey, 1998.) The power sharing approach might be understood as delegation, but according to Mills and Friesen (1995) delegation and empowerment is not the same thing. Delegation refers to a superior passing on some of his/her own tasks for others to undertake. Empowerment means the removal of constraints that prevent individuals from working to their optimal potential (Ruohotie, 2002b).

Empowerment can also be approached as a motivational construct that emphasizes individual cognition, perception and subsequent emotions. Viewed in this way, it is related, among other things, to self-efficacy beliefs and outcome expectations. Perceived empowerment is a process that expands individual power in comparison to status quo or some solid end result. It will occur to varying degrees within any organization, and individuals will experience variable feelings of empowerment at different times (Koberg, Boss, Senjem & Goodman, 1999). Thomas and Velthouse (1990) define psychological empowerment as intrinsic motivation characterized by cognitive structures like relevance, competence, self-direction and influence.

According to Spreizer, DeJanasz and Quinn (1999), there is a connection between empowerment and management of change. Managers should be able to evolve innovative ideas, gain support from his/her superiors, and finally, encourage members of the work community to strive for a common goal.

Niemi (2002, p. 11) has defined empowerment in higher education learning, on the basis of work by Israel, Checkoway, Schulz and Zimmerman (1994), as “... mediative and reciprocal relationships between learners and
between learners and technology-based environments that is intended to increase
the learners’ ability to gain understanding and control over personal, social and
learning environmental force, in order to take action to improve their individual
and collaborative learning opportunities.”

There have been few attempts to operationalize the concept of empowerment. The IQ FORM research group has initially operationalized the concept with four on-line self-rated tests (Niemi, 2002). The first test measures students’ strengths and weakness using a questionnaire (Tirri, K., Komulainen, Nokelainen & Tirri, H., 2002) based on Howard Gardner’s (1993) differentiation of human intelligences. The second test concerns efficacy of learning measured on four dimensions: 1) Expectations of success; 2) Performance anxiety; 3) Inner reward of one’s own studies and concept of usefulness of studies; 4) Self-efficacy and self-confidence. (Pintrich & Garcia, 1991; Pintrich, Smith, Garcia & McKeachie, 1991; Pintrich & McKeachie, 2000; Ruohotie, 2000c.) The third test is for time management, self-management in learning, persistent in learning tasks and help seeking strategies (Zimmerman, 2000; Ruohotie, 2000c). Finally, the fourth test measures learning skills in higher education based on the following learning operations: rehearsal, critical thinking, finding essential points, connecting newer and older knowledge, using keywords and advance organizers, applications of theories and self-assessment of learning skills (Ruohotie, 2000c).

Herrenkohl, Judson and Heffner (1999) developed a measurement instrument to test different dimensions of empowerment as they strove to discriminate empowered and non-empowered employee groups. They found several characteristics that typify empowered groups: solving of relevant problems, search for new ideas and independent functioning. They see empowerment as a relationship between an employee’s actions and an organization’s support for those actions. Based on that assumption, they identify four dimensions of empowerment: 1) Shared vision; 2) Institutional structure and administration that supports responsibility; 3) Employees’ responsibility for knowledge and learning; 4) Institutional recognition. (Table 4.)
Table 4. Dimensions of Empowerment (Herrenkohl, Judson & Heffner, 1999)

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Shared vision</td>
<td>Workers are aware of expectations for performance of duties and institutional goals. They have a sense of responsibility and commitment.</td>
</tr>
<tr>
<td>2. Institutional structure and administration that supports responsibility</td>
<td>Decision-making processes and the chain of command are explicit. Teams are effective and take responsibility for their own actions. Workers are eager to develop their performance and have confidence in superior(s) decisions</td>
</tr>
<tr>
<td>3. Employees’ responsibility for knowledge and learning</td>
<td>The institution encourages workers to seek new, more productive ways of doing their job, to seek new information and to develop their know-how. Workers trust their superiors and believe that failure will not automatically result in punishment.</td>
</tr>
<tr>
<td>4. Institutional recognition</td>
<td>The institution recognizes employee’s achievements. Employees feel that institution cares about them. The rules for rewarding are explicit.</td>
</tr>
</tbody>
</table>

2.3 Modeling of Organizational Atmosphere

Important factors in the development of growth orientation are support and rewards from the management, the incentive value of the job itself, the operational capacity of the team and work related stress (Ruohotie, 1996b, 2000b). Each of these can be further divided into smaller individual factors.

*Management and leaders* face such challenges as how to develop and reward learning, how to empower people, how to support development of professional identity, create careers based on interaction, set goals for learning and how to plan development, evaluate learning and its development and how to create commitment to the job and the organization.

*The incentive value of the job* depends on the opportunities it offers for learning, that is, the developing nature of the job. Therefore, essential factors for professional growth are the developmental challenges, the employees’ chances to influence, opportunities to collaborative learning and valuation of the job.
The operational capacity of a team or a group can be defined by its members’ capability to operate and learn together, by the work group cooperation and by the reputation for effectiveness.

Work related stress might become an obstacle for professional growth. Ambiguity, vagueness and conflicts of role, a too heavy mental load and demand for continuous alterations may stress people and damage the organizational atmosphere. Negative stress quickly suppresses growth and development.

2.3.1 The Initial Model of the Growth-oriented Atmosphere
In 1998, Ruohotie and Nokelainen (2000a) examined theoretical dimensions of growth-oriented atmosphere in a Finnish polytechnic institution of higher education. The sample consisted of 318 employees, 65 per cent out of total population of 492 employees. Both male ($n = 145$) and female ($n = 147$) respondents group sizes were almost identical (46%) with 8 per cent ($n = 27$) missing data. Respondents age was classified into four groups; 20 to 29 year (5%, $n = 17$), 30 to 39 year (25%, $n = 78$), 40 to 49 year (37%, $n = 120$), and over 50 year (24%, $n = 75$) with 9 per cent ($n = 29$) missing data. Job profile contained three groups (13% of missing data): Managers (8%, $n = 25$), teachers (44%, $n = 139$) and other personnel, that is, cleaner, caretaker, librarian, (41%, $n = 131$). The organization consisted of ten geographically separate units. The instrument utilized in the study contained 80 statements in a 5-point Likert-scale from 1 (strongly disagree) to 5 (strongly agree).

Ruohotie and Nokelainen (id.) constructed an initial version of the theoretical dimensions of growth-oriented atmosphere (Table 5). Respondents indicated only moderate differences in preferences for various dimensions as mean values ranged between 3.2 and 3.8. Alpha values ranged from satisfactory (.77) to excellent (.93).
<table>
<thead>
<tr>
<th>Dimension</th>
<th>Definition</th>
<th>$M$ ($SD$)</th>
<th>Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth-oriented atmosphere</td>
<td>Employees attain adequate level of organizational information.</td>
<td>3.16 (.71)</td>
<td>.93</td>
</tr>
<tr>
<td>Job premises and tools</td>
<td>Teaching premises are up-to-date.</td>
<td>3.41 (.80)</td>
<td>.79</td>
</tr>
<tr>
<td>Participative leadership</td>
<td>It is easy to contact and collaborate with the leader of training program.</td>
<td>3.49 (.86)</td>
<td>.85</td>
</tr>
<tr>
<td>Elaborative leadership</td>
<td>Responsibility is distributed into different levels of this educational institute.</td>
<td>3.18 (.82)</td>
<td>.85</td>
</tr>
<tr>
<td>Encouraging leadership</td>
<td>Superior gives support to his/her subjects.</td>
<td>3.28 (.97)</td>
<td>.90</td>
</tr>
<tr>
<td>Work group cooperation</td>
<td>Active discussion about developing work and working environment.</td>
<td>3.43 (.68)</td>
<td>.83</td>
</tr>
<tr>
<td>Student’s attitudes towards teachers</td>
<td>Atmosphere is pleasant and unreserved in the class.</td>
<td>3.49 (.62)</td>
<td>.84</td>
</tr>
<tr>
<td>Developing components of work</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Community spirit</td>
<td>Organization has common goals. Work gives intrinsic fulfillment.</td>
<td>3.28 (.74)</td>
<td>.85</td>
</tr>
<tr>
<td>- Incentive value of the job</td>
<td>Possibility to affect on work and working environment.</td>
<td>3.83 (.78)</td>
<td>.90</td>
</tr>
<tr>
<td>- Potential to influence</td>
<td></td>
<td>3.22 (.87)</td>
<td>.84</td>
</tr>
<tr>
<td>Dignity of the job</td>
<td>Work contribution is respected.</td>
<td>3.63 (.85)</td>
<td>.87</td>
</tr>
<tr>
<td>Growth motivation</td>
<td>To trust ones abilities in difficult situations.</td>
<td>3.65 (.61)</td>
<td>.82</td>
</tr>
<tr>
<td>Commitment to work and organization</td>
<td>To be truly excited about ones work.</td>
<td>3.69 (.77)</td>
<td>.77</td>
</tr>
<tr>
<td>Achievement motivation</td>
<td>Desire to prove ones capabilities.</td>
<td>3.45 (.70)</td>
<td>.77</td>
</tr>
</tbody>
</table>
Ruohotie and Nokelainen (ibid.) found that growth-oriented atmosphere generates togetherness and reflects on developing leadership. Multidimensional scaling and Bayesian unsupervised model-based visualization both provided evidence to conclude that factors representing encouraging leadership and commitment to work and organization are closely situated in, but in different dimensions. Following conclusions were made: 1) Teacher’s professional growth-motivation reflects directly with task value on teacher-pupil relationships and on achievement motivation; 2) Task value has also an effect on growth-oriented atmosphere; 3) Growth-oriented atmosphere is highest in work assignments that offer challenging professional tasks (manager, teacher) and lowest among other workers.

Figure 1 represents initialized organizational growth-oriented atmosphere model as visualized by the job title and work unit (id., p. 164). Likert-scale from 1 to 5 is discretized and placed in the upper part of the figure. Thin bars in the figure describe initial values. The figure illustrates clearly that ill atmosphere is more common within the group of other employees than with managers or teachers. It may also be stated that units A, D and F embody places with low ratings of growth-oriented atmosphere while units E, H and J are out performers considering their initial values. (Figure 1.)

2.4 Growth-oriented Atmosphere Questionnaire (GOAQ)

In this dissertation, I present an updated model of the dimensions of growth-oriented atmosphere (Table 6). These thirteen dimensions are operationalized in the Growth-oriented Atmosphere Questionnaire (Nokelainen & Ruohotie, 2008). Four main categories, namely 1) Supportive and rewarding management; 2) Supportive value of the job; 3) Operational capacity of the team; 4) Personal attitude toward the work, have remained unchanged when compared to the previously presented model, but the dimensions within them have changed to some extent. Next I will briefly describe the main differences between the two models.
Figure 1. Organizational Growth-oriented Atmosphere Visualized by Employees’ Job Title and Work Unit (Ruohotie & Nokelainen, 2000a, p.164)

Note. Job Title = Management, Other (clerk, cleaner, etc.), Teacher. Work Unit = A, … , J.
First dimension of the previous model (see Table 5), “Growth-oriented atmosphere”, overlapped heavily with all the other dimensions (indicated by the inter-item correlations) of the model and was thus removed.

Also the second (“Job premises and tools”) and sixth (“Student’s attitudes towards teachers”) dimensions were omitted as my aim was to produce a more generic growth-orientation model. The second dimension was omitted as I have created a more work premise-oriented instrument (modeled after “Job Satisfaction Barometer” by Finnish Public Administration) to cover those research areas. The sixth removed dimension is problematic to analyze as the number of teachers in various organizations may vary heavily, usually only between 0 to 30 per cent.

Two leadership dimensions (“Participative leadership” and “Elaborative leadership”) were transformed into three dimensions in the new model: “Strategic leadership”, Rewarding of know-how” and “Developing of know-how”.

Previous model’s “Work group cooperation” dimension was further elaborated into “Team spirit” dimension in the new model. “Potential to influence” dimension was joined to the “Incentive value of the job” dimension. “Dignity of the job” was renamed to “Valuation of the job”.

“Achievement motivation” dimension was removed from the new model as it is operationalized via the items in our other measurement instrument, Abilities for Professional Learning Questionnaire (APLQ, see research article III and sections 3 and 5.3 for details).
### Table 6. Dimensions of Growth-oriented Atmosphere

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Supportive and rewarding management</strong></td>
<td>Management of the organization expresses and consolidates values that direct activities, monitors the development processes of units and defines the direction and focus of operations.</td>
</tr>
<tr>
<td>2. Strategic leadership (STR)</td>
<td>Manager supports and motivates personnel to develop know-how, work methods and work community.</td>
</tr>
<tr>
<td>3. Rewarding of know-how (REW)</td>
<td>Organization rewards its employees' professional knowledge and skills. Members of work community gain more responsibility as their know-how increases.</td>
</tr>
<tr>
<td>4. Developing of know-how (DEV)</td>
<td>Organization takes an active interest in its employee's professional growth. Members of work community are interested in self-developing.</td>
</tr>
<tr>
<td><strong>Supportive value of the job</strong></td>
<td>Work gives intrinsic fulfilment by being versatile, autonomous and challenging.</td>
</tr>
<tr>
<td>5. Incentive value of the job (INV)(^a)</td>
<td>Work gives intrinsic fulfillment by being versatile, autonomous and challenging.</td>
</tr>
<tr>
<td>6. Clarity of the job (CLA)</td>
<td>Personnel has a clear picture of goals and responsibilities. They are aware of decision-making processes and personal expectations.</td>
</tr>
<tr>
<td>7. Valuation of the job (VAL)(^a)</td>
<td>Work contribution is respected by the worker itself, colleagues and management.</td>
</tr>
</tbody>
</table>

Note. \(^a\) Common dimension as in the previous study in the same organization (Ruohotie & Nokelainen, 2000a) with an 80-item version of the questionnaire.
Table 6. Dimensions of Growth-oriented Atmosphere (Continued)

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Operational capacity of the team</strong></td>
<td></td>
</tr>
<tr>
<td>8. Community spirit (COS)(^a)</td>
<td>Members of work community learn from each other, for example, via dialogue, by analyzing mistakes, participating in collaborative planning and quality development.</td>
</tr>
<tr>
<td>9. Team spirit (TES)</td>
<td>Good team spirit promotes helping each other and taking responsibility over common goals. Work group members discuss about developing work and working environment.</td>
</tr>
<tr>
<td><strong>Personal attitude toward the work</strong></td>
<td></td>
</tr>
<tr>
<td>10. Psychic stress of the job (PSY)</td>
<td>To what extent work and changes relating to it induce psychic strain like fatiguette and flightiness.</td>
</tr>
<tr>
<td>11. Build-up of work requirements (BUI)</td>
<td>Members of work community are cope with changes in the personal workload.</td>
</tr>
<tr>
<td>12. Commitment to work and organization (COM)(^a)</td>
<td>Members of work community are truly excited about ones work. How important it is to stay in current job.</td>
</tr>
<tr>
<td>13. Growth motivation (GRM)(^a)</td>
<td>Members of work community are trust ones abilities in difficult situations, take new challenges and develop ones know-how.</td>
</tr>
</tbody>
</table>

Note. \(^a\) Common dimension as in the previous study in the same organization (Ruohotie & Nokelainen, 2000a) with an 80-item version of the questionnaire.

2.4.1 Relation between the Dimensions of Empowerment and Growth-oriented Atmosphere

Nokelainen and Ruohotie (2003a) have developed a mapping between the dimensions of empowerment (as defined by Herrenkohl et al., 1999) and growth-oriented atmosphere (as defined by Ruohotie & Nokelainen, 2000a). (Table 7.) Purpose of their work was firstly to study empirically Herrenkohl and colleague’s model, and secondly, to establish an informative ‘empowerment’ sub scale into the Growth-oriented Atmosphere Questionnaire.
Table 7. A Mapping Between the Dimensions of Empowerment (Herrenkohl et al., 1999) and Growth-oriented Atmosphere (Ruohotie, 2000c; Nokelainen & Ruohotie, 2003, p. 156)

<table>
<thead>
<tr>
<th>Empowerment</th>
<th>Growth-oriented atmosphere</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Shared vision</td>
<td>Commitment to Work and Organization (COM)</td>
</tr>
<tr>
<td></td>
<td>Strategic Leadership (STR)</td>
</tr>
<tr>
<td>2. Responsibility supporting structure and administration</td>
<td>Team Spirit (TES)</td>
</tr>
<tr>
<td></td>
<td>Clarity of the Job (CLA)</td>
</tr>
<tr>
<td></td>
<td>Community Spirit (COS)</td>
</tr>
<tr>
<td>3. Knowledge and learning</td>
<td>Developing of Know-How (DEV)</td>
</tr>
<tr>
<td></td>
<td>Incentive Value of the Job (INV)</td>
</tr>
<tr>
<td></td>
<td>Growth Motivation (GRM)</td>
</tr>
<tr>
<td></td>
<td>Rewarding of Know-How (REW)</td>
</tr>
<tr>
<td>4. Institutional recognition</td>
<td>Encouraging Leadership (ENC)</td>
</tr>
<tr>
<td></td>
<td>Valuation of the Job (VAL)</td>
</tr>
</tbody>
</table>

The Bayesian network model supported the conclusion that the first empowerment dimension, shared vision, is not a one-dimensional construct. Results indicated that respondent’s commitment to work and organization depends over three times more on encouraging than strategic leadership. The second dimension, institutional responsibility supporting structure and administration, was separated into two clusters. Encouraging leadership was the common factor between the two, indicating that when management defines direction and focus of operations, employees have a clearer picture of goals and responsibilities. In addition, they are encouraged to participate in collaborative planning and quality development. The third dimension, knowledge and learning, was found to be one-dimensional, indicating that work gives intrinsic fulfillment to those members of the work community who are interested in self-developing. Results indicated that self-development had a more important role in the model than organizational rewards. A weak relationship between growth motivation and know-how development was reported. The fourth empowerment
dimension, institutional recognition, was one-dimensional, indicating that when management is encouraging by defining direction and focus of operations, employees feel that colleagues and management respect their work contribution. (Nokelainen & Ruohotie, 2003a.)

2.4.2 Operationalization of the First Research Question
The first research question of this dissertation is formulated as follows: What is the optimal number of dimensions and items in the Growth-oriented Atmosphere Questionnaire (GOAQ) to describe the theoretical model of growth-oriented atmosphere? Dimensionality of the GOA construct was examined in the first original publication with quantitative analysis based on empirical samples. Figure 2 illustrates the final GOA model and shows it’s operationalization with the GOAQ. Details of the study design and the questionnaire are described in section 5.1 and in the original publication (Study I).
Figure 2. Operationalization of the Growth-orientation Model Relating to the First Research Question

Note. *Study I* = Nokelainen & Ruohotie (2008). Investigating Growth Prerequisites in Organizations: Case Study in a Finnish Polytechnic Institution of Higher Education. *RQ 1* = What is the optimal number of dimensions and items in the Growth-oriented Atmosphere Questionnaire (GOAQ) to describe the theoretical model of growth-oriented atmosphere?
3 PROFESSIONAL LEARNING

A major goal of any modern education, whether of children or adults, should be promotion of the development of self-regulatory skills and thus the creation of opportunities for life-long learning.

(Puustinen & Pulkkinen, 2001, p. 283)

Self-regulation is essential for transforming the inner animal nature into a civilized human being.

(Vohs & Baumaister, 2004, p. 1)

3.1 Learning as a Process

For the purposes of this study, term ‘learning’ is defined as process of construction, reorganization or transformation of knowledge and its use. According to Bailey, Hughes and Moore (2004, p. 30-31), concept of ‘knowledge’ includes: 1) ‘Facts’ (“Thomas Bayes was a reverend.”); 2) ‘Theories’ (“Encouraging leadership is a prerequisite for good work spirit.”); 3) ‘Procedural skills’ (“After job shadowing I am now able to change forklifts tire and change the hydraulic fluids.”); 4) ‘Social skills’ (“I learned not to do more work than the other members of my team.”); 5) ‘Strategies’ (“I will use the Delphi method and let the experts do most of my dissertation.”); 6) ‘Higher-order thinking’ (“I would be most happy to use your inventory, but show me first the scale reliabilities.”); 7) ‘Worldviews’ and ‘values’ (“Qualitative research method is a profound approach, as opposite to quantitative research method that is a superficial approach.”).

In the field of professional learning, the most significant state of knowledge is its use in practice, its contribution to the way people make sense and participate in activities. Thus, we can see learning as the process of a person changing his/her use of knowledge. Individual and groups of learners interact with and through shared forms of knowledge in constructing shared activities (id., p. 32). Learning is not a matter of bringing out the knowledge already in the learner, nor is it a matter of the individual being shaped by the consequences of behavior. Learner takes part in the activities of a community of practice (Wenger, 1998), and in the process engages situated knowledge (Brown, Collins & Duguid, 1989; Lave & Wenger, 1990).
According to constructivist learning theory (e.g., Vygotsky, 1930; Dewey, 1938; Piaget, 1969; Bruner, 1966), learners build meanings actively and they work intentionally. Concepts related to intentionality are cognitive goals, conscious control and purposeful utilization of information. Regarding intentional learning, the central concepts are goal intention and implementation of intention. Pintrich and Sinatra characterize intentional conceptual learning and describe the qualities of intentional cognition as follows: Intentional conceptual change is 1) Goal-oriented action with emphasis on change in conceptual understanding; 2) Related to metacognitive and metaconceptual understanding; 3) Related to internal activity, volitional control and self-regulation (2003; Sinatra & Pintrich, 2003). Intentional learners are, therefore, not only cognitively committed to the learning process; they are also monitoring, and modifying their learning in a metacognitive manner, and their actions are driven by motives, goals, beliefs and emotions.

Researchers agree that a prerequisite for intentional conceptual change is learner’s goal or goal-orientation towards the change (Ferrari & Elik, 2003; Thagard & Zhu, 2003; Linnenbrink & Pintrich, 2003a, 2003b). In addition, motivation and several epistemological beliefs are supposed to have a central role in intentional action (e.g., Andre & Windschitl, 2003; di Sessa, Elby & Hammer, 2003; Limón Luque, 2003; Mason, 2003; Pintrich & Schunk, 2002).

Metacognitive skills or abilities do not guarantee that learners are able to utilize their ability to think; motivational factors are decisive in the actual application and development of metacognition. Application of metacognition involves the understanding of one’s own thinking, active monitoring and controlling of one’s cognitive processes and application of heuristics in problem solving.

Various facilities and skills—such as the application of various learning strategies, utilization of heuristics and control of actions in complex assignments—can help in the execution of a learning assignment, but they do not guarantee that the learner will really consider his/her own thinking or analyze the rational arguments that support the factuality of information.

Metacognitive skills play an important role in the successful execution of an assignment by affecting the utilization of suitable strategies. However, they
do not ensure conscious reflection on which potential strategies, or which one of several competing strategies, should be utilized at each time, or conscious reflection on why some strategy could be useful. According to Hennessey (2003), the ability to reflect the value of specific strategies or heuristics is an intentional level process, whereas the automatic execution of a set of strategies or heuristics is an algorithmic level process.

Chi (1992) makes a distinction between the processes and outcomes of conceptual change. Change processes are those mechanisms that change learners’ earlier knowledge and outcomes are the changes in knowledge. Conceptual outcomes can include enrichment, reassignment of concepts, and change in the framework theory or the radical restructuring of earlier knowledge.

Limón Luque (2003) lists three preconditions that have to be met to make intentional conceptual change possible: 1) Learners have to be aware of their need to change and they have to know what to change (metacognitive precondition); 2) Learners have to desire the change; they have to think of the change as a personal goal (volitional precondition); 3) Learners have to be able to adjust-plan, monitor and evaluate their own change process (self-regulation precondition).

The metacognitive component of self-regulation includes awareness of one’s own knowledge structures, processes and cognitive and affective states, as well as the ability to consciously and deliberately monitor and regulate one’s own knowledge and processes, and cognitive and affective states. (Hacker, 1998, p. 11.)

Schraw (1998) differentiates between two components in metacognition: the knowledge of cognition and regulation of cognition. Knowledge of cognition refers to learners’ own knowledge of their cognition. It can be divided into three kinds of metacognitive awareness: declarative, procedural and conditional knowledge. The components of cognition include motivation and emotion. Limón Luque (2003) uses the terms meta-motivation and meta-emotion to refer to knowledge and regulation of one’s motivation and emotions. Regulation of cognition refers to those activities that learners resort to when they control their own learning. Those activities include, for example, the regulation of their
attention resources and learning strategies or attending to comprehension breakdowns.

Intentional conceptual change is possible if the learner is able to plan, monitor and evaluate his/her own change process. Learners regulating their own change processes are aware of their own knowledge and beliefs, and are willing to change in a way that helps in attaining their goals (ibid.).

The regulation of cognition, motivation and emotion does not by itself guarantee conceptual change. Learners can be skilful in regulating their own activities, but still unable to apply their skills, unless they have the required domain-specific knowledge and skills.

Skilful self-regulators have the ability to create internal prerequisites for the change process. They can regulate their alertness and handle stress, emotions, restlessness and boredom. These skills also promote the will for change.

Self-regulation is a necessary precondition for the intentional conceptual change to take place. Will to change and will to see the change as a personal goal are also central preconditions that are strongly related to domain-specific knowledge and skills. Learners can feel a need for change and they may have a wish to learn new things, but inadequate domain-specific knowledge may impede them from taking advantage of their self-regulation abilities.

3.2 Self-regulatory Processes

Self-regulation involves complex processes such as setting goals for learning, using effective strategies to organize, code, and rehearse information to be remembered, and using resources effectively. It is also affected by performance monitoring, expectancy beliefs, effective time management and beliefs about the value of learning. In this study, self-regulation is examined in relation to the acquisition, use, and control of learning strategies for use in planned learning activities. Learning strategies include any thoughts, behaviors, beliefs, or emotions that facilitate the acquisition, understanding, or later transfer of new knowledge and skills (Weinstein, Husman & Dierking, 2000, p. 727).

Self-regulation refers to an individual’s active participation in his/her own learning process. The key factors in this process that support learning are the
metacognitive, motivational and behavioral processes, such as planning of the learning process, setting personal learning goals, outlining the material to be learned, self-monitoring, efficacy beliefs, expectations of outcomes, the reconstruction of belief systems, constant practicing and the refining of personal skills and work patterns (see Ruohotie, 1994a, 1996b, 2000a, 2000c, 2002a; Ruohotie & Koiranen, 2000). Both individuals’ skills in self-reflection and their knowledge of themselves as learners are important.

From the perspective of phenomenology, the best way to increase one’s ability for self-regulation is to understand the central role of one’s own self-perceptions in self-regulation. As a method, phenomenology helps us understand the importance of our perceptions and thinking about ourselves and about the external world, so that we can direct and regulate our own behavior. It also helps us understand how people may be helped to respect themselves, develop their competence and ability to make decisions and choices, control themselves and to carry the responsibility for continuous learning.

3.2.1 Models of Self-Regulation

According to Zeidner, Boekaerts and Pintrich (2000), self-regulation is an overarching construct covering self-regulated learning (SRL), that is, regulation of one’s health and stress management, which in turn cover, for example, strategy use, self-observation and automaticity. Puustinen and Pulkkinen (2001) have examined and compared five seminal SRL models that have been considerably developed during the past decades and that have empirical evidence to support their theoretical assumptions: 1) Monique Boekaerts’ Model of Adaptable Learning (Boekaerts & Niemivirta, 2000); 2) John Borkowski’s Process-oriented Model of Metacognition (1996); 3) Paul Pintrich’s General Framework for SRL (Pintrich, 2000a; 200b); 4) Philip Winne’s Four-stage Model of Self-regulated Learning (2001); 5) Barry Zimmerman’s Social Cognitive Model of Self-regulation (1994, 1998, 2000, 2001, 2002).

Puustinen and Pulkkinen (id.) compared the five models on four criteria: 1) The background theories of the authors; 2) The definitions of SRL; 3) The components included in the models; 4) The empirical research conducted by the authors.
Examination of the *first criteria* revealed that 1) Borkovski’s model is the purest representative of information processing perspective and the metacognitive research tradition; 2) Boekaert’s model is mostly influenced by action control theory and transactional stress theory; 3) Pintrich’s and Zimmerman’s models reflect social cognitive theory (Bandura, 1991); 4) Winne’s model is the most heterogeneous reflecting most of the aforementioned theories.

When the authors investigated the *second criteria*, they found two kinds of SRL definitions, firstly, a goal oriented definition, and secondly, a metacognitively weighted definition. Boekaerts, Pintrich and Zimmerman define SRL as a goal-oriented process where “monitoring, regulating and controlling one’s own learning includes cognitive, but also motivational, emotional and social factors” (Puustinen & Pulkkinen, 2001, p. 280). Borkowski and Winne include the importance of goal orientations in their models, too, but define SRL more metacognitively governed process oriented way where the main goal is to adapt the use of cognitive tactics and strategies to tasks.

Investigation of the *third criteria* showed that all five models share the same theoretical components, but the relative weight given to each component varies from one model to another. All but one models assume SRL to proceed from some kind of a preliminary phase through to the actual performance phase, and finally, to an appraisal or adaptation phase. For example, Pintrich’s model has four phases (forethought, monitoring, control, reflection) that are in parallel with Zimmerman’s three phases (forethought, performance, self-reflection). Winne, instead, argues that SRL is recursive, and that metacognitive monitoring can produce internal feedback during any phase of the SRL process (ibid.).

Analysis of the *fourth criteria*, empirical research, revealed two major orientations, firstly a motivational orientation, and secondly, a strategy orientation. Motivation oriented researchers (Boekaerts and Pintrich) have both studied the relationships between motivational factors and academic achievement and they have developed questionnaires to assess motivational and cognitive elements influencing students’ learning. Strategy oriented researchers have applied instructional (Borkowski) and trace (Winne) methods. Zimmerman’s research has been both motivation and strategy oriented. (ibid.)
3.2.2 Concept of Self-Regulation

In this study, I apply Zimmerman’s model of self-regulation where the term ‘self-regulation’ refers to the process through which self-generated thoughts, feelings, and actions are planned and systematically adapted as necessary to affect one’s learning and motivation (Schunk & Ertmer, 2000; Zimmerman, 2000).

The presence and quality of observable actions and hidden processes are determined by one’s underlying beliefs and motives. This definition focuses on the process and differs from definitions emphasizing a particular trait, ability, or stage of competence as an explanation for self-regulation. The critical influence of beliefs and motives explains why a person may complete one task but not another. The process definition also differs from metacognitive self-regulation, which emphasizes knowledge aspects and deductive reasoning such as selecting cognitive strategies. Although metacognition plays an important part in learning, self-regulation also depends on self-beliefs and affective reactions, such as doubts (lack of confidence) and fears, related to specific performance contexts. (Zimmerman, 2000.) For example, perceived efficacy-beliefs about one’s ability to organize and implement actions necessary to complete a specific task explains personal motivation to self-regulate one’s performance (Bandura, 1997).

According to social-cognitive theory, self-regulation is dependent on the situation. Therefore, self-regulation is not some general characteristic/facility or a specific development level but is contextually dependent. Students can neither regulate their actions similarly in all branches of study nor while studying different topics. (Schunk, 2001.)

Although some processes of self-regulation (e.g., goal setting) can be adapted to many different situations, the learner has to understand how these processes can be used efficiently in different branches and contents of study.

Zimmerman (1998) discerns six areas in which learners can regulate their behavior: Motives, methods, time consumption, outcomes, physical environment and social environment. Self-regulation is possible to the extent that the learner has the opportunity and ability to affect these six areas. If all six of the above-mentioned aspects are determined by someone other than the learner him/herself, the source of control is external (e.g., supervisors, teachers, parents, and
computer). Naturally, learning can still take place but self-regulation cannot occur.

Feedback loops and other self-monitoring processes are at the center of self-regulation research due to the seminal work of Carver and Scheier (1981) where they emphasized the applications of cybernetic theory to how people monitor their states in relation to goals and other standards. Also Zimmerman (2000) describes self-regulation as cyclical process where the feedback from prior performance is used to make adjustments during current efforts. Because personal, behavioral and environmental factors are constantly changing, an individual has to monitor these changes continuously in order to know whether any adjustments are required. The three feedback loops involved in monitoring one’s internal state, one’s behaviors and one’s environment constitute what Zimmerman (id.) has described as the triadic forms of self-regulation (see Figure 3). Regulation of personal factors, which is referred to as covert self-regulation, involves monitoring and adjusting cognitive and affective states, such as the use of imagery for retrieving information or relaxing. Behavioral self-regulation comprises self-observation and applicable performance processes, such as learning methods. Environmental self-regulation refers to the observation and adjustment of environmental conditions and outcomes.

![Figure 3. Triadic Forms of Self-regulation (Zimmerman, 2000, p. 15)](image-url)
There are a number of factors that affect the effectiveness of self-observation. Delayed feedback prevents a person from taking corrective action in time: one monitors his/her performance after the completion of the task when feedback no longer affects the performance. The information level of performance feedback also influences the quality of performance. Feedback received in a standardized or structured setting increases the information level of the results. A third qualitative feature is the accuracy of self-observation. Misperceptions of one’s performance also detract from the self-regulation of one’s actions.

The continuous and careful self-monitoring of various sources of self-control facilitates the learner’s strategic adjustments and development of self-beliefs. The feedback loops are open: Unlike closed systems that limit self-regulation (for example, decrease performance discrepancies reactively against a constant standard), open systems proactively increase performance discrepancies by setting higher goals and seeking more challenging tasks. Self-regulatory processes can be adjusted both proactively and reactively in order to reach personal goals.

When we compare the Zimmerman’s SRL model to the four other models presented in preceding section 3.2.1, we see that it differs most strikingly from two models. Winne argues that metacognitive monitoring produces internal feedback during any SRL process phase. Boekaerts mainly focuses on the preparatory phase of the SRL process treating the performance and appraisal phases more unprofoundly. Niemivirta (2004) has later addressed all three phases thoroughly in his more elaborative version of the aforementioned SRL model.

3.2.3 Self-Regulation of Learning

Figure 4 describes self-regulation of learning tasks as a cyclical, three-phase process (Zimmerman, 1998). The phases in this learning cycle are forethought, performance or volitional control, and self-reflection. Forethought precedes a learning performance. The determinants of forethought are goal setting, strategic planning and personal beliefs (self-efficacy, goal orientation and intrinsic interest). They reinforce commitment to act and prepare the learner for the actual
effort required by learning. Performance or volitional control guides the learning process and regulates concentration and learning performance.

Self-reflection refers to examining and making meaning of the learning experience. This process occurs throughout the course of learning. Some traditional views of reflection describe it in a way that is passive, contrived, after the fact, time-consuming, narrow and context independent. Others, however, such as Seibert (1996), describe reflection as an active mental process of conscious involvement with experience that requires deliberately bringing one’s thinking to the level of conscious awareness. Therefore, it is a matter of understanding learning experiences not only on the level of action but also on the level of conscious thinking. In this view, reflection becomes a natural adaptive response to turbulent learning conditions. In rapidly changing conditions, the right moment for reflection is not retrospectively after the learning experience; instead, reflection has to take place during an experience so that it can have an impact on the experience.

Figure 4. Cyclical Phases and Sub processes of Self-regulation (Ruohotie, 2002a, p. 40; adapted from Zimmerman, 1998, 2002)
Each phase of the learning cycle can be further subdivided into components or factors. Learning depends on the learner’s ability to manage these different aspects of self-regulation.

**Forethought phase** is described in the model in terms of goal setting, strategic planning and self-efficacy beliefs. Goal setting refers to specific learning outcome decisions and strategic planning refers to the selection of those learning strategies and methods with which the learner tries to attain a desired goal. A number of personal beliefs, such as the learner’s self-efficacy, outcome expectations and goal orientations, as well as intrinsic interest in the subject or skill to be learned, affect goal-setting and strategic planning. Goals represent objects, events, states, or experiences one seeks to attain, and the concept of goal orientation refers to a personal factor that contributes to the individual’s selection of different goals (Niemivirta 2004, p. 31). Niemivirta thus defines goal orientations as individual’s tendencies to select or favor certain goals and outcomes over some others. Nurmi, Salmela-Aro and Koivisto (2002) studied, amongst other things, importance of goals in the transition from vocational school to work. Their sample consisted of 250 young adults studying last year at a business-oriented vocational school and a vocational institute of technology. They found that more the young adults emphasized the importance of work-related goals while they were still at school the more likely they were to find a relevant job to their education.

Strategies in the **performance or volitional control phase** help the learner to focus on the given task and to optimize performance. By focusing and concentrating his/her attention, the learner is protected from distractions and competing interests. The learner also chooses strategic applications and learning methods him/herself; for example, learning by using his/her imagination or by repeating previously learned information out loud. Time management, environmental structuring, and help seeking are also widely used by skillful self-regulators. Self-monitoring gives individual information about his/her progress. However, self-monitoring can also break one’s concentration and the learning process can suffer as a consequence. An acquired skill requires less and less intentional monitoring as the skill becomes routine and automatic so that self-
monitoring moves to a more general level focusing to the learning environment and the outcome of action.

**Self-reflection phase** begins with self-evaluation, which is the process whereby individual compares information attained through self-monitoring to extrinsic standards or goals. She wants to have fast and accurate feedback on his/her performance as compared to others. Self-evaluation leads to attribution interpretations; the learner interprets the reasons for success or failure. For example, she might blame failure on a lack of abilities or on low effort. Attribution interpretations can lead to positive self-reactions. The learner might interpret the failure of a strategy as the result of too little effort and then increase his/her efforts, but if she interprets the reason for failure as being a lack of ability, the reaction is liable to be negative. Attribution interpretations reveal the possible reasons for learning mistakes and help the learner to find those learning strategies which best suit the given situation. They also develop or promote the adaptation process. Self-regulated learners are more adaptive and evaluate their performance appropriately. Positive reactions (e.g., self-satisfaction) reinforce positive interpretations of oneself as a learner and intrinsic interest in the task.

The self-regulated learning cycle can, in summary, be described in terms of forethought, performance or volitional control and self-reflection. Forethought prepares the student for the performance phase where the control strategies the learner chooses play an important role. Applying these strategies then leads to the self-reflection phase and then back to performance-guiding self-interpretations and forethought. Thus, the self-reflection phase may lead to evaluation of current methods and replacing them with new, more developed ones.

Theoretical framework of self-regulation in professional learning is summarized in Figure 5. The figure represents self-regulation (Zimmerman, 1998, 2000) as a system concept (Boekaerts & Niemivirta, 2000) managing leadership behavior through interactive processes between motivation, volition, emotion, attention, metacognition and action control systems. As Hannula (2006) points out, self-regulation should be seen to be much more than mere metacognition. For example, Malmivuori states that within self-system processes, emotions activate various self-regulatory processes at different levels.
of self-awareness, including self-reflection (2006). She contrasts automatic affective regulation (low level of control) to active regulation of affective responses (high level of control).

3.2.4 Empirical Support for Self-Regulation

Several researchers have devised hypothetical models of the mechanism of self-regulation and attempted to demonstrate their accuracy with research results (for example Snow, Corno & Jackson, 1996; Pintrich, 2000a; McCombs, 2001; Schunk, 2001; Winne, 2001; Zimmerman, 2001, 2002). The following six conclusions can be drawn from their research. First, the self-monitoring and self-evaluation processes that people use direct their cognitive and meta-cognitive abilities to regulate and control their own development. The design of learning tasks and experiences affects the learner’s opportunities for the direction and control of learning processes and activities. Second, important preconditions for self-regulation are: 1) Understanding the role of self-beliefs in thinking processes, emotions, motivation and the regulation of one’s actions; 2)
Awareness of one’s own facilities for self-regulation and their utilization in learning; 3) Intrinsic understanding that the learning assignments and experiences are meaningful and relevant. Third, the essential precondition for efficient self-regulation is that the learner has an optimal (not too strong or too weak) faith in his/her own abilities and possibilities. Fourth, the intrinsic motivation of students is higher if they can do things that they want to do. They need to have the opportunity to exert an influence and make choices, experience successes and perform assignments that serve their purposes and goals. Fifth, feedback concerning competence enhances intrinsic motivation by strengthening the learners’ beliefs in their abilities to control the reaching of the goal. Feelings of success enhance intrinsic motivation more than belief in one’s competence. Sixth, commitment to the learning process is a necessary, but not sufficient, condition for the development of skills and knowledge. Goal-orientation in behavior and the self-regulative processes that learners use to implement their intentions are also important. They have to observe and assess their own actions. These processes give one the possibility of adjusting one’s behavior, enhancing learning and strengthening desire for the continuous development of one’s skills and abilities.

Learners from different cultures tend to use different strategies to regulate their learning. In addition, the beliefs inherent in the culture influence the learners’ own perceptions of their learning needs and how their goals can best be achieved. However, there are wide variations within cultures so such generalizations cannot be applied to individuals.

McCombs and Whistler (1997) list various characteristics of learner-centered environments. Two of these involve individual differences in learning: Firstly, despite the numerous basic principles of learning, motivation and efficient teaching, learners have differing abilities and they prefer different learning strategies. Secondly, personal beliefs, thoughts and understanding are all based on prior learning and interpretations and they form a framework for all later construction of reality and interpretation of life experiences.

These principles focus on the individual background factors and abilities that affect learning. They explain why people learn different things at different times and in different ways. The same principles can be applied to all human
beings when it comes to learning, thinking, emotions and development. Despite this, the things that are learned, and how this learning is communicated, is different in different environments (e.g., various cultures or social groups) and is also affected by hereditary factors. From this starting point, learners create their individual thoughts and beliefs, and their concept of themselves and the world around them. Appreciating and understanding the differences between learners that are revealed in a learning context help develop efficient learning environments for diverse learners.

3.2.5 Self-Regulatory Abilities as Metacompetencies

Self-regulatory abilities that enhance professional development, which are sometimes referred to as self-management skills, may be understood as conative constructs (see Ruohotie, 2000a, 2002a). Conation, which is intermediate between an individual’s cognitive and affective attributes, is sub-divided into two parts: motivation and volition. Motivation is related to pre-decisional attributes and behaviors, and volition to the attributes and behaviors that affect carry through on decisions once they are made. An individual’s motivation to act is affected by such things as his/her values, beliefs, self-concept, outcome expectations and level of goal orientation. Once the decision is made, such things as the individual’s ability to focus his/her attention, choose effective strategies, monitor and maintain effort, manage time and seek help when necessary determine his/her success in achieving the goal. This is referred to as volitional control.

Another way of thinking about self-regulation (Zimmerman, 2002) is in terms of the three sub-processes of forethought, performance control and self-reflection explained previously (see Figure 4). In this model, motivation is a primary aspect of forethought (along with planning) and volition is a synonym for performance control. For the purposes of this discussion, the terms motivation and volition will be used.

Sub-components of motivation include achievement orientations (ego-orientation, task-orientation, and the need for achievement) and self-directed orientations (self-efficacy beliefs, control beliefs, and self-esteem). Volitional constructs include activity-related control strategies (metacognitive skills, critical
thinking, and management of the resources) and orientation to others (social ability, empathy, and persuasibility). Individual interests and learning styles (independent, competitive or cooperative, see Grasha & Yangarber-Hicks, 2000) also guide development, but it is difficult to classify them as either motivational or volitional attributes (Ruohotie, 2000a, 2002a; Ruohotie & Koiranen, 2000).

People are generally reasonably aware of what they know and what they do not know, their intellectual strengths and weaknesses, how they should apply their skills and knowledge to solve various problems, how they can acquire missing skills, and their chances of succeeding at a particular task. This self-related knowledge and assessment guides our actions, and can be a significant factor in determining performance. An increase in knowledge or training does not necessarily guarantee better learning and performance. Often it is those who know more about themselves and are able to apply this knowledge in practice that perform best in difficult problem solving situations.

‘Metacognition’ encompasses knowledge of how to learn best, of the strategic use of knowledge, and of personal factors influencing the effectiveness of learning and application (Maudsley & Strivens, 2000). Driscoll (2005, p. 107) provides a comprehensive definition of this term by stating “metacognition refers to one’s awareness of thinking and the self-regulatory behavior that accompanies this awareness.” Thus metacognition comprises higher level monitoring and controlling of cognitive functions, linking decision-making and memory, learning and motivation, and learning and cognitive development (Nelson & Narens, 1994).

Knowledge about one’s own knowledge and reasoning ability is called ‘metaknowledge’. However, mere knowledge of self is not sufficient to improve performance. One also needs to understand the possibilities and limits to one’s competencies in specific situations and be able to judge when and how to use them, which Nelson and Narens (1990) have termed ‘metacompetence’. Thus blending knowledge, performance and occupation illustrates metacompetence (Maudsley & Strivens, 2000). On the other hand, Hyland (1993) prefers concept of vocational or occupational expertise over metacompetence.

A basic prerequisite for the acquisition of metaknowledge is the ability to observe one’s own cognitive processes and products. With the help of
introspection a learner becomes aware of his/her own learning processes and
tinking, gains experience in learning and applying knowledge and skills, and
develops metaknowledge about the areas of his/her competence as well as his/her
areas of relative incompetence. (Weinert, 2001.)

Metacognition that pertains to self-regulation may be sub-divided into
metacognitive knowledge, metacognitive awareness and metacognitive control;
that is, into declarative knowledge about situations and one’s behavior, and the
awareness and control that allows for application of that knowledge as
procedural metacompetence. (id.)

Declarative metaknowledge includes experience and knowledge about the
difficulty of various tasks; knowledge about one’s own abilities, skills, and
cognitive deficits; knowledge about learning, problem solving, and action
regulation; knowledge about efficient strategies for learning, memory, problem
solving, and rectification; and knowledge about techniques for mastering various
tasks, compensating for missing knowledge, and setting realistic goals. (Pintrich
& Ruohotie, 2000; Ruohotie, 2003; Weinert, 2001.)

Procedural metacompetencies are necessary for applying metacognitive
knowledge to optimize task-directed behavior. They include all those strategies
that help to organize tasks/problems and facilitate problem solving (for example,
organizing a task to make memorizing easier, underlining and marking important
points in a text, and searching for memory cues). Strategic measures also include
the use of effective cognitive tools (for example, graphics and analogies),
purposeful application of cognitive resources and continuous observation and
evaluation of learning/performance. (Pintrich & Ruohotie, 2000; Ruohotie, 2003;
Weinert, 2001.)

The quantity and quality of relevant metaknowledge affects learning,
memory, and problem solving more than chronological age; and may be taken as
a general indicator of mental development in childhood and general intelligence
in adulthood. (Pintrich & Ruohotie, 2000.)

Self-regulation of learning requires knowledge based on facts and
experience and application of various strategies and problem solving heuristics.
It also requires self-reliance. Although metaknowledge can be distinguished
from motivational and volitional processes analytically, it is not possible to
separate them empirically. What we observe is always an inseparable combination of the three.

3.2.6 The Development of Self-Regulation
It is important to develop professionals’ concept of and skill in self-regulation to enable them to be effective lifelong learners (McCombs, 1988; Paris & Newman, 1990; Pintrich, 1990; Pintrich & DeGroot, 1990; Schunk & Zimmerman, 1994; Pintrich & Ruohotie, 2000). Improving perceptions of personal control leads to the strengthening of intrinsic motivation, the improvement of learning outcomes, and the development of responsibility and a sense of self-efficacy. It also increases the likelihood that skills and strategies that are learned will be used subsequently in new learning situations.

Self-regulation of learning is not always conscious (Vohs & Baumeister, 2004). Like other skills, it may develop to a level where it is automatic and seems natural. The roots of this automatic behavior are, however, still deep since such subconscious self-regulation is based on knowledge, skills and beliefs that have been integrated through learning experiences over a long period of time.

Pintrich (1995) suggests five things that may improve learner’s self-regulation: 1) Learners have to be more conscious of their activities, motivation and cognition; 2) They have to adopt positive motivational beliefs; 3) They have to have self-regulatory models; 4) They should be allowed to practice adapting to different learning strategies; 5) Their assigned tasks should support the use of self-regulation. In various self-regulation development programs learners have been trained to monitor and control internal and external conditions; for example, time management, study environment, different distractions, concentration on the task, emotions and motivation (Trawick & Corno, 1995).

Zimmerman (2000) identifies four different stages in the development of self-regulatory skill. The first level involves the observation of social models of self-regulation. One acquires new skills and strategies by observing the performance of skilled individuals and listening to their verbal explanations and self-expressed beliefs. In addition to strategic skills, models develop performance standards, motivational orientations, and values. The second, emulation level of self-regulatory skill is attained when the learner’s
performance resembles that provided by a model. The learner does not copy exactly, but rather adopts the general pattern or style of functioning. The learner’s approximation of the desired self-regulation can be improved by providing guidance, feedback, and social reinforcement during practice. The third, a self-controlled level of self-regulatory skill is attained when the learner employs skills of self-regulation in a structured setting even in the absence of a model by using an internalized model. The fourth, full self-regulatory skill level is achieved when the learner can systematically adapt his/her self-regulation to changing personal and contextual conditions by altering the application of strategies and making adjustments based on the conditions. She is no longer dependent on fidelity to the model. The mastery of self-regulatory skill means that the learner need not pay overt attention to the learning process, but can shift the focus of his/her attention to the content and desired outcomes. This is the ultimate goal for all learners, and an increasingly important one for professionals in a world of knowledge work and lifelong learning (Ruohotie, 2000a).

Learning is most natural and effective when it is experienced as meaningful and relevant and when the learning environment supports the learning and encourages self-regulation and self-control. According to Zimmerman (1994), self-regulation is only possible in a context that allows people to make their own choices and be in control. If learners are not given the opportunity to exercise control and the freedom to choose their own strategies, they will not learn to regulate their own behavior and they will not develop enthusiasm for taking initiative. Therefore, professional learning environments should be designed in a way that creates an appetite for learning and makes it possible to practice self-regulation. In other words, they should offer possibilities for choice, challenge, collaboration, making meaning, taking initiative in one’s actions, and receiving constructive feedback and rewards.

Effective self-regulation is dependent on the development of self-image and self-processes, such as self-awareness, self-monitoring and self-evaluation. Learners have to be able to define their learning goals and the meaning that they have for them. They have to know (and accept) themselves and have realistic expectations in relation to their own capacities. They have to understand their responsibility for their self-improvement and take an active role in it, and they
have to understand their own ability to control their cognition, affections, motivations and behavior.

Learners also need to have the option to work independently, to make personal plans and to choose strategies with which they feel comfortable. They have to experience the learning goals as meaningful and relevant. The efficient planning of studies and choice of strategies are connected to meta-cognitive knowledge and processes. While studying, professionals have to direct and maintain their attentiveness, assess their own progress in relation to their goals, regulate and control their feelings and decrease the discrepancy between their real and ideal goals. The critical abilities are, self-monitoring, self-reflection and self-evaluation.

In intervention programs designed to develop independent study skills, professionals have been helped to alter their views of themselves as learners and to understand their own ability to affect thinking processes and their contents. Alongside the development of the learning environments, attention has been paid to consciously affecting their observations, self-assessment, interpretations, emotions, motivations and self-regulative processes. Professionals have to learn to take advantage of their learning experiences and see them as a chance to develop their own unique skills and abilities. They also have to understand their responsibility for themselves and for others in social interaction. They have to see their varied possibilities to grow as human beings with intellectual, physical, social and spiritual dimensions.

### 3.3 Attribution Theory

Attributions are those reasons individuals give for an outcome, such as success or failure in a task (Heider, 1958). Factors involved in attributional thinking, such as specific reason for success and failure, have shown to be related in achievement settings (Weiner, 1974, 1980, 1986, 1994, 2000). Weiner found in his studies that the four most frequent reasons for success and failure are ability, effort, task difficulty and luck.

Dai, Moon and Feldhusen (1998) classify attribution constructs into three groups. First, *attribution appraisals* are online explanations assessed following actual or manipulated success or failure in performing a specific task. Second,


attribution beliefs are domain-specific or domain-general beliefs about the causes of success or failure. Third, attribution styles are generalized, stereotypical patterns of attributions and dispositional beliefs. Attribution styles are assessed in a similar way to attribution beliefs, except that a certain typology is imposed on the data using predetermined criteria.

In this dissertation, attribution styles are examined using Weiner’s (1992) classification of reasons for success and failure: 1) Internal and external attributions, referring to within or outside individual causes; 2) Stable and unstable attributions, referring to consistent or inconsistent causes over time; 3) Controllable and uncontrollable attributions, referring to the extent an individual believes he or she has control over the cause of an outcome.

In Figure 4, the self-reflection phase referred to examining and making meaning of the learning experience. Zimmerman states, based on Bandura’s (1986) work, that self-reflection begins with self-judgment, which is the process where an individual compares information attained through self-monitoring to extrinsic standards or goals. The goal is to have fast and accurate feedback on his/her performance as compared to others. Self-judgment leads to attribution interpretations where an individual interprets the reasons for success or failure.

The most widely applied theoretical viewpoint to attribution interpretations is Weiner’s attribution theory (1974), which is based on a principle that an individual is in constant search for understanding why an event has occurred.

Attribution interpretations can lead to both positive and negative self-reactions. The individual might interpret the failure of a strategy as the result of too little effort and then increase his/her efforts, but if he or she interprets the reason for failure as being a lack of ability, the reaction is most likely to be a negative one.

Attribution interpretations reveal the possible reasons for learning mistakes and help the learner to find those learning strategies which best suit the given situation. They also develop or promote the adaptation process; self-regulated individuals are more adaptive and evaluate their performance appropriately. Positive reactions (e.g., self-satisfaction) reinforce positive interpretations of oneself as an individual and enhance intrinsic interest in the task.
3.4 Self-confidence Attitude Attribute Scales (SaaS)

As noted earlier, there is a large research body on attribution interpretations, especially on attributonal properties in achievement settings, for example, Weiner’s work. This is mainly due to the fact that role of motivation in academic achievement has proven to be a popular topic over the years. The interest is apparent as we examine the structures of the most well known measurement instruments.

Biggs’ (1985) 42-item *Study Process Questionnaire* (SPQ) consists of two scales (motive and strategy) with three approaches: 1) Surface; 2) Deep; 3) Achieving. The questionnaire contains six subscales (surface motive, deep motive, achieving motive, surface strategy, deep strategy, and achieving strategy).

Entwistle’s *Approaches to Studying Inventory* (ASI) contains subscales including such factors as fear of failure, extrinsic motivation, and achieving orientation (Ramsden & Entwistle, 1981).

Marsh’s *Self-description Questionnaire* (SDQ) consists of a set of scales for different age groups measuring self-concept with a multifaceted view. The SDQ includes mathematical, verbal, academic and physical self-concepts (Marsh & O’Neill, 1984).

Campbell’s (1996a) *Self-confidence attitude attribute Scales* (SaaS) measures individual’s attributions based on self-attribution theory (Weiner, 1974). Although Weiner’s original conceptualization contained four attributions (ability, effort, difficulty, luck), the statistical analysis based on numerous empirical samples produced only two distinct scales, namely effort and ability (Campbell, 1996a, 1996b). SaaS is applied in the second study of this dissertation as it concentrates solely on the operationalization of the attribution theory.

3.4.1 Operationalization of the Second Research Question

The *second research question* of this dissertation is formulated as follows: Are the theoretical dimensions of the Self-confidence attitude attribute Scales (SaaS) questionnaire identified in the domain of three groups of mathematically gifted
participants: Academic mathematics Olympians, polytechnic institute of higher education students, and elementary school students who have participated in mathematical competitions?

As defined earlier in the second chapter, *competence* is the potential capacity of an individual to successfully complete a certain task according to certain criteria set by someone else (Ellström, 1994). In this second study setting, an interesting point is that a competence may also be seen as an attribute of the individual, for example, referring to a human resource that the person brings to a mathematical problem solving situation (Nokelainen, Tirri & Merenti-Välimäki, 2007). Thus, attributions may emphasize potential competence as indicated by the capacity of an individual to successfully complete tasks and face new challenges on the basis of demonstrated personal attributes and abilities (other than those obtained through formal training).

Ellström (2001) has noticed that potential competence may vary greatly between individuals with the same formal qualifications, because they may possess very different levels of inherent ability and may have learned different things outside of school or studies through their working life and recreational activities. Thus, ability attributions affect later performance expectations and, in negative cases, the development or continuation of learned helplessness (Ruohotie & Nokelainen, 2000a).

In the second original publication, the focus was on participants’ self-evaluations on the basis of mathematics achievement and academic ability. My major goal was to investigate if the theoretical dimensions of the SaaS were identified with the sample, and further, to study if the attribution styles of mathematically highly gifted individuals differ from the other participant group’s attribution styles.

Figure 6 illustrates self-regulation as a system concept and shows its operationalization with the SaaS questionnaire. Details of the study design and the questionnaire are described in section 5.2 and in the original publication (Study II).
3.5 Learning Strategies

As noted in the previous chapter, Weiner (1974) found in his studies that the four most frequent reasons for success and failure are ability, effort, task difficulty and luck. However, subsequent research identified learning strategies as a fifth possible reason as it makes no sense to try harder if you do not know in the first place how to try (Alderman, 2004).

Many taxonomies of learning strategies have been proposed (e.g., Weinstein & Mayer, 1986; Pintrich, 1988). One approach is to characterize learning strategies as relating to cognition, metacognition or resource
management. Cognitive strategies help the learner to classify new material and to structure knowledge. Metacognitive strategies help the learner to plan, to regulate, to verify and to shape his/her own cognitive processes. Resource management strategies help the learner to control available resources—such as time, effort and outside help—in order to cope with a task.

One of the most extensive strategy analyses is the classification presented by Pintrich (1988). The strategies involved in his analysis are discussed below.

3.5.1 Metacognitive and Cognitive Strategies
The term ‘metacognition’ is often treated synonymous to the conscious selection and assessment of strategies. It can be subdivided into knowledge and skills. Metacognitive knowledge includes an individual’s knowledge about his/her own schemas, strategies and processes, and the conscious awareness of one’s own learning abilities. Awareness of the difficulty of various tasks and their demands are also part of metacognitive knowledge. Awareness of one’s own learning abilities is closely related to motivational components, such as self-efficacy, control beliefs, and expectancy for success. Awareness of learning tasks is linked to task value and goal orientation.

Pintrich and McKeachie (2000) argue that metacognitive strategies involve the control and regulation aspect of metacognition more than the knowledge aspect. Metacognitive control activities include planning, monitoring and self-regulation. Planning activities include setting goals for studying, skimming, generating questions before reading the text, and doing a task analysis of the problem. All these activities help the learner plan the use of strategies and the processing of information. In addition, they help to activate relevant aspects of prior knowledge that make organizing and comprehending the material easier. Good learners engage in more planning and more metacognitive activities than poor learners.

Monitoring activities include tracking of attention in the learning situation, self-evaluation that is aimed at understanding the material during the learning process, monitoring comprehension of a lecture, and certain kinds of test-taking strategies (e.g., monitoring speed and adjusting to time available). A learner can monitor difficulties in understanding the content. An example of this type of a
task is self-questioning, which ensures understanding the content. In terms of reading a book, one can ask questions at the beginning of each chapter, which directs reading and helps in getting deeply into one’s subject. These various monitoring activities assist the learner in understanding the material and integrating it with prior knowledge.

Self-regulation strategies are related to monitoring strategies. For example, as learners monitor the comprehension of a text, they can regulate their reading speed according to the difficulty of the material, or they can review the material as a whole and answer the problems in a flexible way (e.g., skipping questions and coming back to them later). These self-regulation activities make learning more effective and improve performance because the learners are continuously monitoring their behavior, and they are able to make corrections as they proceed.

Pintrich (1988) classifies cognitive strategies into rehearsal, elaboration, organization and critical thinking strategies. Rehearsal strategies have an effect on attention and the processes involved in acquiring knowledge, but they do not appear to assist learners’ internal connections among the information or integrate the information with prior knowledge. Elaboration strategies help learners store information in long-term memory by building internal connections between new information and prior knowledge. Organization strategies help learners in selecting the information to be learned and in building internal connections within it. Weinstein and Mayer (1986) make distinctions among cognitive strategies used in learning situations (rehearsal, elaboration and organization strategies) between simple and complex versions.

Pintrich and McKeachie (2000) include the students’ ability to apply previous knowledge to new situations in the cognitive strategies that direct the learning process. Application of acquired knowledge often requires critical thinking. However, it appears that critical thinking is somewhat domain or discipline specific. The nature of critical thinking in literary analysis may be very different from the type of critical thinking promoted in psychology, which fosters a critical analysis of research methodology and theory.
3.5.2 Resource Management Strategies

Resource management strategies assist learners in managing the environment and the resources available (e.g., time and study management, effort regulation and help-seeking). These strategies can be seen as both cognitive and metacognitive, but they are different enough to warrant a separate category. Resource management strategies help learners in adapting to environment as well as changing the environment to suit their needs.

**Time Management:** Both long-term and short-term studying requires time management. Long-term studying may include study periods that last for several weeks or even months. Scheduling such studies effectively involves planning and regulation activities that are metacognitive in nature. Studying should be flexible, and learners should be prepared to allow adaptations in their study plans as the course proceeds. Time management is equally important for short-term studies. For example, if a student has set aside three hours one evening for studying, she must be able to schedule the use of that time efficiently. Time management directs the choice of activities, and therefore it is related to various motivational components, such as goal orientation and task value.

**Study Environment Management:** Learners may work in many different environments (e.g., library, own work room, or kitchen table). It is important that the location chosen is free of distractions. The learner needs to organize the study environment in such a way that she is able to concentrate on work.

**Effort Regulation:** According to Pintrich and McKeachie (ibid.), effort management is one of the most important learning strategies, and it is also at the center of the interaction between motivation and cognition. A good learner knows when to increase effort and persist on the task as well as when maximal effort is not required for success. She also knows that different learning strategies may be required depending on the task.

**Peer Learning and Help-Seeking:** Peer learning is based on social interaction: the learner constructs meaning and understanding by actively participating in a discussion and sharing knowledge (Dart, 1998). The learner has to explain, specify, and argue his/her position to others. Collaborative learning groups provide an opportunity to examine and refine understanding: opposing
views are alternatives worth exploring together rather than competitors to be eliminated.

A learner needs to know when support of others is needed, and when/how it can be obtained. Most often this type of help comes from the teacher or peers (e.g., tutors). Sternberg (1985) describes “practical intelligence,” which refers to good students knowing when they do not know something well enough, and being able to find someone to provide assistance. For several reasons, many students are unable to benefit from the support of others, or they seek it seldom. According to Ruohotie (2000a), help seeking is related to learning motivation.

### 3.6 Motivation and Volition in Learning

In order to understand differences in learning styles and strategies, it is important to differentiate between cognitive, affective and conative constructs. *Cognition* is a generic term for those processes through which an organism recognizes and obtains information about a certain object. Cognitive constructs include the following concepts: perceiving, recognizing, conceiving, judging and reasoning. *Affect* is the feeling response to a certain object or idea. Sometimes it means the energy resulting from an emotion or a general reaction to something that one likes or dislikes. Affective constructs include feeling, emotion, mood and temperament. *Conation* refers to those mental processes that help an organism to develop. It is a kind of intrinsic unrest (the opposite of intrinsic balance or homeostasis) or a conscious tendency to act or strive for something. Conative constructs include impulse, desire, volition, and purposive striving. (Snow, Corno & Jackson, 1996; Ruohotie, 2000a.)

Snow and his colleagues (1996) developed a taxonomy that describes affective, cognitive and conative constructs. A construct is a hypothetical, psychological state; in other words, an inferred system, structure, process, force or activity that is seen in the regular patterns of observed behavior. Personality and intelligence are found at the top of the taxonomy. These are very broad terms that are difficult to define. *Personality* is rarely limited to just personal characteristics, temperament or emotions. In general, it includes all those factors that make a person an individual. Many personality traits are also linked to cognitive factors. *Intelligence* is also a complex construct. It means,
firstly, the ability to undertake activities which are difficult, complex, abstract, demanding, goal-oriented, socially prestigious and original; and, secondly, the ability to accomplish these activities in situations which demand concentration and the control of one’s emotions (Barrow & Milburn, 1990, p. 157.) Clearly, there is some overlap between these two constructs.

The concepts at the next level—affect, conation and cognition—are each further divided into two subcategories. Once again, however, the constructs are not entirely distinct. For example, cognitive theories make a distinction between declarative and procedural knowledge. Declarative knowledge is a sort of knowledge network in which concepts and facts are linked together. New knowledge develops as a result of the construction of knowledge; in other words, analyzing the interdependencies of different pieces of knowledge. Procedural knowledge can be expressed as a set of procedures or rules that help in remembering and applying knowledge. This, however, presupposes and depends upon declarative knowledge.

Similarly, affect can be subdivided into temperament and emotion, although personality factor interpretations rely on aspects of both. Temperament refers to biological traits that are not dependent on situational factors, whereas an emotion may be strongly linked to the given situation.

The conative constructs of learning, which include motivational and volitional aspects of human behavior, were not appreciated until the 1980’s (e.g., Snow & Farr, 1987; Snow & Jackson, 1994). Motivational aspects of conation include, among others, internal and external goal orientation, fear of failure, need for achievement, self-esteem, belief in one’s own abilities and possibilities (efficacy beliefs), value of incentive (valence), and different attribution interpretations. Volitional structures include, among others, persistence, the will to learn, endeavor/effort, mindfulness in learning, intrinsic regulation and evaluation processes, as well as different control strategies (e.g., allocation and control of resources as well as emotional attentiveness and motivational control strategies) and styles of processing knowledge (Ruohotie, 1999, 2000a).

Constructs in all parts of the taxonomy—for example, traits of temperament, skills, styles, strategies, orientations and tactics—directly affect learning outcomes and can, conversely, be interpreted as learning outcomes
themselves. Learning experiences have consequences that are not exclusively confined to any one part of the taxonomy. For example, as a learner learns changes take place not only in his/her knowledge base but also in his/her metacognitive skills, motivation, beliefs and self-esteem. As learning experiences accumulate, styles, strategies and techniques develop which are derived from these experiences. The learner also begins to look for new ways and opportunities to apply his/her learning styles. A career-oriented learner will focus on those styles, values, beliefs, skills and knowledge that are respected in his/her field.

Although the concepts in the taxonomy can be distinguished from each other theoretically, it is impossible to maintain these ‘pure’ distinctions when doing research. For example, many personality attributes, such as, self-esteem, anxiety in examinations, flexibility and authoritarianism are related to attitudinal and belief structures. Attitudes and beliefs, either positive or negative, can easily lead to patterns of behavior and thus to personality developments. A positive attitude towards studying results in engagement, success and positive self-concept as a learner, whereas a negative attitude may lead to dropping out.

3.6.1 Control of Motivation and Volition

The conative element of the preceding taxonomy of individual differences in intelligence and personality is particularly important in the consideration of self-regulation. According to Corno (1989), motivational processes help to formulate decisions and promote decision-making whereas volitional processes guide the subsequent enactment of the decision. It is useful to distinguish between pre-decisional processes of motivation and post-decisional processes of volition because even highly motivated students may have problems setting clear goals and enacting their intentions.

Volition and motivation (i.e., conation) can be explained using various dynamic cycles (Snow et al., 1996) that connect volitional constructs or processes, motivational factors and learning outcomes or results. These connections are strongly influenced by the context (Corno, 1993). One cannot assume that all learners function methodically and deliberately and that different processes/constructs affect everyone in the same way. Volition and motivation
are dependent on the learning task and on such individual traits as cognitive-intellectual abilities and reinforcement experiences (how learning has been earlier reinforced).

According to Ruohotie (2002a), several personal determinants of motivation have been identified which are linked to decision-making and to the willingness to participate in learning and performance tasks. These determinants include a personal need for achievement, fear of failure, various intrinsic and extrinsic goals and perspectives on the time required in order to achieve a goal. These factors are called achievement orientations. Two other goal-choice and motivational factors have been identified. The first one includes self-directed orientations such as self-concept, self-worth and self-efficacy. The second one includes values, attitudes and interests, which determine our preference for certain subject matter, tasks and procedures.

Constructs involved in goal implementation are closely linked to volition. They assist the individual in carrying out plans and intentions. Action control is one category of individual difference constructs, which is used when trying to control competing intentions, and attention-distracting factors. Individuals can also control available resources in a timely and efficient way. Students differ in what they consider sensible activities and how long they are prepared to work. Self-regulated learning traits are mindfulness, effort and persistence. Modern research on conation has tried to describe, for example, individual goal-setting, self-regulation that leads to goal-oriented action, effort investment, and the attempt to set realistic goals. Presumably, conative factors have an effect on educational outcomes, cognitive abilities, affect and other personal and situational factors. (id.)

Another area of volition is orientation toward self and others, which refer to an individual’s susceptibility to, acceptance of, and, to some extent, desire for influence by other people. This attribute can be described using, for example, the following concepts: persuasibility, empathy, and social ability or social intelligence. Reactions to different people and new situations vary between individuals. There are also differences in how people try to influence others as well as in their willingness to be influenced by others. (id.)
Personal styles belong in the third volitional category. Many stylistic constructs of learning and studying are dependent on volition. The learner may decide for him/herself, for example, how superficially or in-depth she wants to study. Other-directed constructs and personal styles are not mutually exclusive. Rather, there are many overlapping factors, which can be seen in many cognitive skills, strategies and methods. (ibid.)

3.7 Integrative Model of Motivation

Motivational expectancy model (Pintrich, 1994) categorizes and integrates the central elements of modern motivational theories. This model includes different beliefs or expectancies; for example, perceived competence, test anxiety, perceptions of task difficulty, the learner’s belief in his/her efficacy, and expectancy of success. The learner who has a strong self-image and high expectations will put more effort into his/her task and will persist longer, even on a difficult task, compared to the student who has low expectancy of success.

A value perspective is evident in the evaluations of the task value as well as in the student’s goal orientation. The task value has three facets: attainment value, interest value and benefit value. The attainment value refers to the degree of challenge the learner anticipates. It is high if a learner feels him/herself capable and estimates him/herself to be able to master demanding course work. Interest value refers to a learner’s intrinsic interest in the contents of a certain subject (e.g., she likes chemistry). The utility value refers either to the goal itself or to the instrumental task. For example, a learner who is not interested in applied statistics may nonetheless study it enthusiastically because statistics is a compulsory part of the graduation program. The course then has a high utility value. The material to be learned could well have utility value as well as interest value for the learner but, while it is desirable, it is unrealistic to expect that instruction will always increase a student’s intrinsic interest towards learning.

In his several articles Pintrich has described the components of motivation and their role in learning (e.g., Pintrich, 1988, 1990, 2000a, 2000b; Pintrich & McKeachie, 2000). Next, I will define the most important aspects of his theoretical model.
3.7.1 Value Components

Learner goal orientation and task value are value components affecting the learner’s choice of activities as well as his/her persistence at a task. Value components may refer to the reason why someone is engaging in a task or taking a course.

**Learner Goal Orientation:** Internal goal orientation may be to learn the content in a particular domain; to experience challenge, curiosity or joy through learning; or to increase self-worth. External goal orientation is related to external goals, such as good grades, rewards or acceptance. It is generally understood that learners with strong internal motivation try harder, are more persistent than others, and apply more effective learning strategies in their learning than those with external motivation.

**Task Value of Learning:** Task value refers to the learner’s perceptions of the importance, utility or interest of a task or course. Task value is related to learner goal orientation and to the intensity of behavior.

3.7.2 Expectancy Components

Expectancy components include the learner’s beliefs about his/her own ability to perform a task, beliefs of self-efficacy and control and expectancy for success. They are related to the student’s self-regulation and metacognitive control; that is, planning skills, ability to concentrate and regulation processes. (Weinert, 2001.) Learners who believe that their ability to perform a task also affects the output are more likely to evaluate their own progress, and in conflicting situations they may apply various cognitive strategies.

**Control Beliefs:** Learners who believe that they are in control of their own behavior and that they can influence the environment tend to achieve at higher levels than students who do not believe they have any control.

**Efficacy Beliefs:** Learners’ self-efficacy has been defined as beliefs about performance capabilities in a particular set of tasks or goals. (e.g., Bandura, 1991, 1993; Schunk, 1991) Therefore, self-efficacy refers to the extent to which a learner relies on his/her cognitive abilities. It is important to distinguish these perceptions of efficacy from students’ beliefs about outcome; that is, their belief that the environment is responsive to their actions. Contingency beliefs about
behavior and outcome may lead to high success expectations, and they are also likely to improve persistency in learning. On the other hand, lack of contingency may lead to passivity, restlessness, lost effort and poor achievement.

3.7.3 Affective Components
Affective components include learners’ emotional reactions to the task (e.g., test anxiety and cognitive conflict situations) and their evaluations of themselves in terms of self-worth.

Test Anxiety: Learners’ cognitive capacity to process information is limited. A learner who has been preparing well for a test and who applies relevant strategies in his/her answers has free cognitive capacity to process various disturbances in a test situation. Similarly, learners with poor cognitive skills may suffer from several disturbances (e.g., interfering anxious thoughts and anxiety), which may decrease their cognitive capacity.

Self-Worth: Self-worth is a critical factor. According to Covington (1984), individuals generally tend to establish, maintain, and promote a positive self-image. Learners may develop a variety of coping strategies to maintain self-worth. However, some of these strategies may have a debilitating effect on student performance. For example, learners may choose easy courses, or they may complete and submit assignments at the last minute. If they still succeed despite the last minute tactics they can attribute this success to ability (because they were able to do well even with poor preparation).

Pintrich’s own studies show that different motivational components, such as self-efficacy, internal goal orientation and test anxiety clearly correlate with the use of cognitive and metacognitive strategies (Pintrich, Smith, Garcia & McKeachie, 1993; Pintrich & Schrauben, 1992).

3.8 Metacognition in Motivation and Learning
The motivational processes discussed in the preceding sections are closely related to emotional outcomes, whereas the learning strategies described earlier in Pintrich’s taxonomy are related more to encoding and processing the information to be learned (Garcia, 1995). Metacognition, however, is important to both.
Cognition, or thinking, refers to a learner’s ability to attend to, acquire, represent, and retrieve information. It directs the learner’s perceptions, actions, thinking, and memory. Metacognition, or “thinking about thinking”, refers to the knowledge and regulation of thinking and learning. It directs the learner’s ability to reflect upon, understand, and control his/her learning. (Dart, 1998.) Metacognitive regulation involves metacognitive knowledge, metacognitive awareness and metacognitive control.

Metacognitive knowledge has different components. One way of thinking about these components is to divide them into knowledge about learning, about the matter to be learned, about one’s own learning style and strengths, and about learning strategies (id.).

Firstly, metacognitive knowledge includes the learner’s knowledge about learning (what it is, how it occurs, and how one knows when something has been learned). This knowledge is related to the learner’s conception of learning. The learner may, for example, believe that learning is about constructing meaning through such activities as familiarizing oneself with topic-related literature, having discussions with colleagues, and attempting to relate new information to what is already known, and then testing if one is able to understand the matter and apply it in practice.

Secondly, metacognitive knowledge involves the learner’s perception of the matter to be learned, which in turn determines how the learner will approach the learning activity. The learner may, for example, realize that a critical review essay is more demanding than a descriptive essay and that different learning tasks require different methods of study.

Thirdly, metacognitive knowledge involves the learner’s recognition of his/her strengths and weaknesses, and his/her assets and liabilities in learning. She may, for example, recognize that she is good at identifying the main ideas from text and organizing these in such a way as to facilitate the integration of new information with prior knowledge.

Fourthly, metacognitive knowledge includes information about effective cognitive learning strategies; in other words, knowing what cognitive strategies are (declarative knowledge) and how to apply them (procedural knowledge). The learner may know that paraphrasing helps his/her to develop connections
between new information and prior knowledge. This, in turn, requires identifying the main ideas from text, rewriting the main ideas in one’s own words, and intentionally integrating new information with prior knowledge.

Metacognitive awareness is the second factor in metacognitive regulation. It results from the learner’s conscious questioning in relation to the learning task and activity. For example, one could ask questions like: “What is the purpose of the task?” and “What do I already know about this?” By asking questions in relation to cognitive processes and attempting to answer them, the learner is able to control his/her learning, progress, and the outcomes of learning.

Metacognitive control is the third factor in metacognitive regulation. It is related to the level of awareness developed. The main areas of metacognitive control are planning, self-regulation, and evaluation. Operations related to planning involve task analysis, identifying relevant prior knowledge, goal setting, selecting appropriate cognitive strategies, anticipating possible difficulties in the completion of the task, and finding ways to overcome these. Self-regulation involves the monitoring of learning (i.e., whether learning is proceeding according to the plan) and the assessment of understanding or the lack of it (i.e., whether other strategies are required to facilitate understanding). Evaluation involves the assessment of both the learning processes and the outcomes of these processes.

3.8.1 Metacognitive Abilities

Metacognition and metacompetence are fundamental to learning to learn. Because this knowledge must be context specific to guide concrete actions, it is important to develop a range of specific metaknowledge for various purposes and domains. Therefore, in addition to addressing the products of learning (i.e., knowledge), teaching and learning processes at all levels should include reflection the processes of learning in order to optimize both declarative metacognition and procedural metacompetence for learners in all contexts.

The changes in working life have led to a situation in which all workers are required to have cognitive skills and take part in the decision making processes. Skills and abilities are interconnected with the ability to analyze domain-specific information and understand the basics and the meanings of different work tasks.
Experts’ cognitive processes are characterized by complexity of domain-specific knowledge structures and deep understanding of concepts. (Pillay, 1998.)

Knowledge structures are different from declarative knowledge. The former relate to the analysis or parsing of information, whereas the latter describes the amount of knowledge or learned facts. The structure of knowledge stored in memory may be more important from the point of view of learning than the amount of it. Knowledge structures affect subsequent knowledge parsing and memory retrieval processes. The retrieval of information from memory speeds up and deepens understanding, helps in decision-making and the anticipation of future events, and makes it easier to find optimal solutions to problems. (Day, Arthur & Gettman, 2001.)

Experts also have the ability to apply their knowledge and skills to new tasks and situations. They are able to transfer information, such as knowledge of terminology and processes to new problem solving processes, for example, in the area of digital communication technology.

Also, thinking skills are needed in working life. Those who master their work have the ability to analyze problems, are proactive and able to anticipate the development of their field, and take responsibility for the effectiveness of their work practices. Higher level thinking skills are related to critical thinking, problem solving and creative thinking, which are all cognitive processes that advance professional knowledge, deepen the understanding of knowledge and increase the transferability of knowledge and skills (Pillay, 1998). Trishman, Jap and Perkins (1993) found seven factors that promote higher level thinking: 1) Open-heartedness and broadmindedness; 2) Intellectual curiosity; 3) Inquisitiveness for connections and explanations; 4) Ability to anticipate outcomes and to make plans; 5) Ability to process information; 6) Ability to assess frameworks and reasons; 7) Ability to monitor one’s own thoughts.

According to Ruohotie (2002a), individuals may have strong professional knowledge and still have extensive shortcomings in their thinking skills.

My conclusion at this point is that work life requires experts to own a strong professional knowledge, ability to transfer their skills and knowledge, and high metacognitive skills. In other words, they need to be both competent and qualified. Success also requires self-regulation skills and motivational
competencies that support self-regulation, especially efficiency beliefs: Trust in one’s abilities to plan and execute the activities that lead towards a skilful accomplishment.

3.9 Abilities for Professional Learning Questionnaire (APLQ)

Finnish vocational education research projects have mostly utilized the *Motivated Strategies for Learning Questionnaire* (MSLQ) developed by Pintrich and his colleagues (Pintrich, Smith, Garcia & McKeachie, 1991). For example, Kivinen (2003) studied reliability of MSLQ with an international sample of 198 secondary level students (median age 17 years) from Finland and Luxembourg. He concluded that the instrument appears to be a reliable and valid for the assessment of the student’s motivational beliefs and strategy use in different cultural environments.

However, self-regulation researchers have for years called for the improvement in the measurement of this construct (Hofer, Yu & Pintrich, 1998). Dugan (2005) states that their main concern has been the area of construct validity. According to Trochim (2006), construct validity is the degree to which inferences can legitimately be made from the operationalizations in the study to the theoretical constructs on which those operationalizations were based. Improvement is most likely needed, as Kivinen (2003, p. 163) concludes about the MSLQ that it “offers a solid instrument for the assessment of student’s motivational beliefs and strategy use” — without mentioning that the scale reliabilities reported on page 85 ranged from .45 (peer learning) to .91 (self-efficacy) with the mean of .68! According to the author (p. 84), those “scale reliabilities are good.” In a recent study, Dugan (2005) reports internal consistency values for the MSLQ from .60 (help seeking) to .91 (self-efficacy) with the mean of .76. His convenience sample was more representative than Kivinen’s as it consisted of 491 higher educational institute students (mean age of 22.5 years) from the U.S. Puklej and Peklaja (in press) reported congruent results with a sample of 245 second-year undergraduate student teachers as their internal consistency values ranged from .60 (control beliefs) to .85 (self-efficacy) with the mean of .74.
At this point it is reasonable to ask what is an adequate level of scale reliability? I refer to Nunnally’s (1978, p. 245) seminal work where he states that “increasing reliabilities much beyond .80 is often wasteful of time.” He carries on (p. 246) with the exception of “... applied settings where important decisions are made with respect to specific test scores.” Using this judgment, most of the aforementioned scale reliabilities are satisfactory at best.

The original 81–item MSLQ consists of two categories, one measuring motivation with 31 items and the other measuring learning strategies with 50 items (Pintrich et al., 1991). The Finnish version, Abilities for Professional Learning Questionnaire (APLQ), retains the same basic structure (Ruohotie, 2000c), but with fewer items — 28 and 40, respectively. MSLQ is strictly limited to measure domain-specific motivation (state), for example, on a class or course level. APLQ also measures domain-specific (state) motivation (e.g., meaningfulness of study, but also contains at least partially domain-general (trait) sections (e.g., goal orientations).

The focus of this dissertation is on the motivation category of the APLQ, which has three sections: A value section, an expectancy section, and an affective section. The value section has three subscales: 1) Intrinsic goal orientation; 2) Extrinsic goal orientation; 3) Meaningfulness of study. The expectancy section consists of two subscales: 4) Control beliefs; 5) Self-efficacy (the correlation between perceived competence, expectancy for success and self-efficacy having been found to be so strong that these two factors have been collapsed into one). The affective section includes one scale, 6) Test anxiety. (Table 8.)

In some sense it is natural to assume that, for example, meaningfulness of study is a domain-specific state (e.g., APLQ item A24: “I believe that my vocational studies will be of practical benefit for me”), and personal goal orientations are more like a domain-general trait (e.g., APLQ item A6: “I expect to get excellent grades in my vocational studies”). Test anxiety is somewhere between domain-general and domain-specific orientations (e.g., APLQ item A21: “Nervousness in exams affects my performance”). However, when considering the latter case we might argue that in certain testing situations person’s anxiety level is exceptionally low due to an unusual high level of
expertise: A person who feels himself incompetent and tense during mathematics test, may feel competent and relaxed while playing musical composition for the examination committee. Dugan (2005) concludes that there seems to be both a state and trait aspect to academic self-regulation.

The learning strategies category is divided into four sections: metacognition in learning, metacognition in practice, learning by doing, and resource management. The metacognition scales have been adapted from the original instrument on the basis of research results. Similarly, the resource management strategies section has been collapsed to a single subscale: time and effort. A new learning by doing subscale has also been added on the basis of theoretical considerations. Many professional skills are not taught directly. They are inculcated through practice and work experiences to the extent that the learning environment provides opportunities for learning and rewards the acquisition of new skills. Thus, intentional engagement in activities that are inherently stimulating and instructive, or which consolidate academic learning through application, is an important strategy for learning.

The revised APLQ instrument has been used extensively by Finnish researchers in the field of professional education with different target groups and subjects (Ruohotie, 1994a, 1996c, 2000a; Nokelainen & Ruohotie, 2002; Kaartinen, 2005). The reported internal consistency values for the motivational dimension of the APLQ are congruent with those of MSLQ, showing highest reliability values for the self-efficacy scale (Nokelainen & Ruohotie, 2002, $\alpha = .79 - .82$; Kaartinen, 2005, $\alpha = .87$) and lowest for the control beliefs (Nokelainen & Ruohotie, 2002, $\alpha = .58 - .60$; Kaartinen, 2005, $\alpha = .60$). The rationale of using APLQ instead of MSLQ in the research field of professional learning is three-fold. Firstly, APLQ is designed to meet the needs of Finnish vocational education research on both theoretical and operationalizational level. Secondly, APLQ is tested with Finnish empirical data. Thirdly, APLQ’s motivational section is considerably shorter (21 items) than MSLQ’s corresponding section (31 items).
Table 8. Comparison of the MSLQ and APLQ Frameworks

<table>
<thead>
<tr>
<th>Motivated Strategies for Learning (Pintrich et al., 1991)</th>
<th>Abilities for Professional Learning (Ruohotie, 2000c)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Value Components</strong></td>
<td><strong>Value Components</strong></td>
</tr>
<tr>
<td>Intrinsic Goal Orientation</td>
<td>Intrinsic Goal Orientation</td>
</tr>
<tr>
<td>Extrinsic Goal Orientation</td>
<td>Extrinsic Goal Orientation</td>
</tr>
<tr>
<td>Task Value</td>
<td>Meaningfulness of Study</td>
</tr>
<tr>
<td>Expectancy Components</td>
<td>Expectancy Components</td>
</tr>
<tr>
<td>Control Beliefs</td>
<td>Control Beliefs</td>
</tr>
<tr>
<td>Perceived Competence</td>
<td>Self-Efficacy</td>
</tr>
<tr>
<td>Self-Efficacy</td>
<td></td>
</tr>
<tr>
<td>Expectancy for Success</td>
<td></td>
</tr>
<tr>
<td>Affective Components</td>
<td>Affective Components</td>
</tr>
<tr>
<td>Test Anxiety</td>
<td>Test Anxiety</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Learning Strategies</strong></th>
<th><strong>Learning Strategies</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive Strategies</td>
<td>Metacognition in Learning</td>
</tr>
<tr>
<td>Rehearsal</td>
<td>Planning</td>
</tr>
<tr>
<td>Selection</td>
<td>Self-Monitoring</td>
</tr>
<tr>
<td>Organization</td>
<td>Self-Evaluation</td>
</tr>
<tr>
<td>Elaboration</td>
<td>Critical Thinking</td>
</tr>
<tr>
<td>Metacognition</td>
<td>Metacognition in Practice</td>
</tr>
<tr>
<td>Surface Processing</td>
<td>Planning</td>
</tr>
<tr>
<td>Critical Thinking</td>
<td>Self-Monitoring</td>
</tr>
<tr>
<td>Original Thinking</td>
<td>Self-Evaluation</td>
</tr>
<tr>
<td>Resource Management</td>
<td>Learning by Doing</td>
</tr>
<tr>
<td>Time and Study Environment</td>
<td>Resource Management</td>
</tr>
<tr>
<td>Effort</td>
<td>Time and Effort</td>
</tr>
<tr>
<td>Help-Seeking Behavior</td>
<td></td>
</tr>
</tbody>
</table>

3.9.1 Operationalization of the Third Research Question

The third research question is formulated as follows: What is the optimal number of dimensions and items in the Abilities for Professional Learning Questionnaire (APLQ) to describe the theoretical model of learning experiences and motivation? Dimensionality of the motivation construct was examined in the
third original publication with empirical samples. The SRL design applied here is an adaptation of Zimmerman’s (1998) three-phase cyclical process: 1) Forethought; 2) Performance or volitional control; 3) Self-reflection. Pintrich’s framework sits in well, as noted earlier in section 3.2.1 (Models of self-regulation). Figure 7 illustrates how the motivational part of the APLQ is related to the SRL model and shows its operationalization with the APLQ questionnaire. Details of the study design and the questionnaire are later described in section 5.3 and in the original publication (Study III).

![Figure 7. Operationalization of the Self-regulated Learning Model Relating to the Third Research Question](image)

Note. Study III = Nokelainen, P., & Ruohotie, P. (2002). Modeling Student’s Motivational Profile for Learning in Vocational Higher Education. RQ 3 = What is the optimal number of dimensions and items in the Abilities for Professional Learning Questionnaire (APLQ) to describe the theoretical model of learning experiences and motivation?
4 BAYESIAN MODELING

Probability statements do not imply the existence of [the hypothetical] population in the real world. All that they assert is that the exact nature and degree of our uncertainty is just as if we knew [the sample] to have been one chosen at random from such a population.

(Fisher, 1956, p. 22)

The problem of rational inference under uncertainty has been the subject of considerable attention since the systematic study of the probability theory began in the eighteenth century. Many different theories of inference have been proposed, and there has hardly been a time when inference under uncertainty was not a matter of real controversy. It seems that educational research community (among many other fields using applied statistics) has been largely unaware of such controversies, and used what is known as classical, frequentistic or Gaussian inference (Hastings, 1997).

Since 1960’s there has been a steady revival of interest in an alternative way of reasoning with probabilities called Bayesian inference (Berger, 1985, Bernardo & Smith, 2000). Many applied fields including astrophysics (Loredo, 1990), medicine (Smith, Spiegelhalter & Parmar, 1996), econometrics (Zellner, 1971), archaeology (Buck, Cavanagh & Litton, 1996) and political sciences have adopted Bayesian techniques, but the penetration of Bayesian inference into educational research has not been particularly influential.

According to Tirri (1999), this is somewhat surprising as quantitative analysis in education exhibits all the features where Bayesian approaches excel: small data sets with many measured issues, emphasis on hierarchical models, models involving latent structures, and data sets with discrete (nominal) values. Thus, when an organizational researcher wants to study dependencies between observed and/or latent variables, for example, the factors of growth-oriented atmosphere (Ruohotie, 1996b), the assumptions for the data may become quite challenging in traditional frequentistic statistical analysis. Examples of such assumptions are the assumption of continuous measurement level, multivariate normality and linearity of both the data and phenomena under investigation.
The result of violating multivariate normality assumption is that chi-square becomes too large (too many models are rejected) and standard errors become too small (significance tests have too much power). Marini, Li and Fan (1996) state that in situations where a latent construct cannot be appropriately represented as a continuous variable, or where ordinal or discrete indicators do not reflect underlying continuous variables, or where the latent variables cannot be assumed normally distributed, traditional Gaussian modeling is not appropriate. In addition, normal distribution analysis sets minimum requirements for the number of observations (i.e., sample size).

In this chapter, I will first discuss the typical problems of using parametric frequentistic statistical techniques, such as Pearson product moment correlation or Student’s t-test, to answer professional growth and learning research questions. Next I will present the Bayesian modeling approach and discuss its essential benefits for professional growth research. Then I will show some practical applications of Bayesian modeling implemented in the B-Course system (Myllymäki, Silander, Tirri & Uronen, 2001, 2002). In the following chapter 5 (Summary), I will show how Bayesian modeling is applied in the research field of professional growth and learning.

### 4.1 Typical Problems of Using Parametric Frequentistic Techniques

First problem of using parametric frequentistic statistical techniques in the research field of professional growth and learning is that they are based on frequencies produced by repeated measures. In practice, this means that widely applied multivariate techniques, such as exploratory factor analysis or multivariate analysis of variance, require a researcher to collect about ten informants for each question that is to be analyzed in the same process. For example, if we use Growth-oriented Atmosphere Questionnaire (Nokelainen & Ruohotie, 2008) that has 62 items on its current version, we need to have about 600 respondents if we want to study its variable structure and compare it with the theoretical model. This figure is not including data that is needed to conduct split analysis in order to convince also the reader that the results are true in the target population (the validation of generalizability).
Secondly, all parametric frequentistic analysis is based on the concept of normal distribution. This means that both the phenomenon modeled and its reflection (data) should follow so-called “Gaussian” (bell-curve) distribution. If this assumption is met, it leads to a desirable state of multivariate normality allowing linear inter-item analysis. However, if the assumption is not met, for example, when the variables in the analysis have different shapes of distributions, the models based on normal distribution are not working correctly and produce false results. This is the case quite often with empirical samples for various reasons. The most common violation of the rules set by the developers of frequentistic parametric statistics is to collect something else than a randomized sample (RS, sometimes called a probability sample). Of course, any non-probability sample that approaches the size of target population will do as well, but quite often the samples we see in the social science research field are ‘convenience samples’ by nature. In such case, even if the phenomenon under investigation is normally distributed, its representative in the data (item, value, etc.) may have a totally different distribution that violates normality assumption. The second common normality assumption violation is related to the first point discussed earlier: Having too small data, that is, too few observations. Even if we have a RS, it does not help us to meet the normality assumption if the data is not ‘saturated,’ there are too few observations to produce a univariate normal distribution. There are numerous other reasons leading to non-normal distributions, for example, unexplained variance in responses due to poorly constructed or misunderstood questions in a questionnaire.

Thirdly, indicators should be continuous (at least an interval level measures) in order to meet the aforementioned multivariate normality assumption (leading to linear dependencies between variables). Although this assumption does not relate to grouping variables that are allowed to have nominal or ordinal measurement levels (for example, linear discriminant analysis), most of the questionnaire items have typically a five-point ordinal ‘Likert-scale’ (de Vellis, 2003, p. 78-80), named after Rensis Likert (Wainer et al., 2000, p. 243). He intended it as a summated scale, which was assumed to have interval scale properties (Likert, 1932). According to Albaum (1997), the Likert-scale is not used that way by academic researchers; individual scale item
scores are used for analysis, as are average scale values for the scales comprising a construct. Statistical dependencies between such non-continuous items are not necessarily linear, calling for non-parametric techniques instead of parametric ones (Bradley & Schaefer, 1998).

4.2 Typical Solutions to Problems of Using Parametric Frequentistic Techniques

Typical solution when assumptions of parametric frequentistic statistical techniques are not met is to use frequentistic non-parametric techniques. Such techniques provide opportunity to apply, for example, Spearman’s rank order correlation ($r_s$) instead of Pearson product moment correlation ($r_p$) or Mann-Whitney U-test instead of Student’s t-test. However, such techniques are, although allowing smaller sample sizes than their parametric counterparts, still frequentistic techniques. Non-parametric techniques are sometimes called ‘distribution free’, but they are not ‘assumption free’ as most of them expects simultaneously analyzed distributions to have symmetrical and similar shapes.

During last two decades there have been numerous attempts to address these important limitations within the circle of traditional frequentistic parametric analysis techniques. As a first response to the continuous measurement level and multivariate normality assumptions, Muthén and Kaplan (1985) suggest that in treating the ordinal variables as continuous does produce viable results as long as the frequency distributions are unimodal with an internal mode. The second response, that is quite similar to the first one, continues by using maximum likelihood (ML) estimates, but in addition uses bootstrapping (see e.g., Yung & Bentler, 1994) to study the effects of non-normality. The third response is so called “robust methods” (see e.g., Filzmoser, 2002) that aim at yielding reliable results via reducing the influence of “unusual” observations on statistical estimates in cases where the classical assumptions are violated. This approach is available in LISREL (ML robust), EQS (ML robust) and Mplus (MLM, maximum likelihood mean adjusted). The fourth response that is presented theoretically by Muthén (1983) and applied in practice in LISREL by Jöreskog (2003) is to estimate tetrachoric (for binary variables) or polychoric (for categorical variables) correlations among the ordinal variables and use these
correlations to estimate the model using asymptotic distribution free function (ADF) by Browne (1984). The fifth response to address modeling problems with ordinal non-normal data is the categorical variable model (CVM) developed by Muthén (1993). The model that is implemented in the Mplus program, uses the general ADF function but without aforementioned limitations.

However, none of the above-described techniques, with the exception of the fifth approach, address the problem of non-linear dependencies between observed variables. In this dissertation, I argue that the Bayesian modeling approach, named after English reverend Thomas Bayes (1701-1761) for his contributions (Bayes, 1763), is a viable alternative to frequentistic statistical techniques addressing all the abovementioned modeling problems. Although the subject domain of this dissertation is mostly focused on professional growth (vocational education), according to the expert field of Pekka Ruohotie, this discussion fully applies to all the other research areas of social sciences where the measurement level of the indicators is categorical.

4.3 Introduction to Bayesian Modeling

Probability is a mathematical construct that behaves in accordance with certain rules (Berry, 1996) and can be used to represent uncertainty. The classical statistical Gaussian inference is based on a frequency interpretation of probability, and the Bayesian inference is based on the “degree of belief” interpretation (Bernardo & Smith, 2000).

Bayesian theory of probability (e.g., Bernardo & Smith, 2000; Myllymäki & Tirri, 1998) is interested in probability of certainty that a given fact or proposition is true. Bayesian way of calculating the probability is often labeled as "subjective probability” or “inverse probability”, as its probability values ranging from zero (proposition is false) to one (proposition is true) are dependent on how much weight we are willing to lay on both the evidence and prior information available.

By the middle of the 18th century, thanks to the weak law of large numbers by Jakob Bernoulli (1687-1759) and the normal approximation to the binomial by Abraham de Moivre (1667-1754), is was known that if we have $n$
independent observations, the chance of success \( \theta \) has the same value and the probability of exactly \( r \) success is given by the binomial distribution

\[
P(r \mid \theta, n) = \binom{n}{r} \theta^r (1 - \theta)^{n-r}
\]  

(1)

According to Lindley (2001), the passage from a known value of \( \theta \) to the empirical observation of \( r \) was greatly appreciated. For example, if we know that about 10 per cent of doctoral students are not going to pass the quantitative analysis course, and if we take a randomized sample of five students, we are able to estimate the probability of one student in the sample to fail the course:

\[
P(r \mid \theta, n) = \binom{5}{1} 1^1 (1 - .1)^{5-1}
\]

\[
= \frac{5!}{1!(5-1)!} (.1)(.66)
\]

\[
= \frac{120}{24} (.066)
\]

\[=.33
\]

Later, Pierre-Simon de Laplace (1749-1827) showed the central limit theorem that assured an approximate normal distribution for practically all sums of independent random variables, if only the number of the summands is large (Fischer, 2001). Thus, the normal distribution is an essential concept in frequentistic statistics having the following assumptions: 1) There is enough data (normal minimum frequency is 30 observations); 2) Measurement scale is continuous (although some researchers treat also ordinal scale as continuous); 3) The phenomena under investigation is normally distributed; 4) Due to central limit theorem, the observed distributions in the data are approximately normal.

Thomas Bayes studied the inverse problem by asking what did the data \( (r, n) \) tell about the probability \( \theta \):

\[P(\theta \mid r, n) \propto P(r \mid \theta, n)P(\theta \mid n)
\]

(2)

In the Equation 2, \( P(\theta \mid n) \) is a priori distribution of probabilities before we know \( r \). It represents a subjective view of our understanding before we have seen any evidence.
Bayesian inference uses conditional probabilities to represent uncertainty (Congdon, 2001). Conditional probability refers to a probability that one event will occur given that another has occurred (Hastings, 1997, p. 23). Therefore, researchers’ interest lies in $P(H|D, I)$ — the probability of unknown things ($H = \text{hypothesis}$) given the evidence ($D = \text{data}$) and background information ($I = \text{information}$). The initial uncertainty about $H$ is also represented as a conditional probability $P(H|I)$. For example, we could have some initial belief that some answers are more likely than others. The essence of Bayesian inference is in the rule, known as Bayes’ theorem (1763), that tells us how to update our initial probabilities $P(H|I)$ if we see data $D$, in order to find out $P(H|D, I)$.

Consequently Bayesian inference briefly comprises the following three principal steps: 1) Obtain the initial probabilities $P(H|I)$ for the unknown things. These probabilities are called the prior (distribution); 2) Calculate the probabilities of the data $D$ given different values for the unknown things, that is, $P(D|H,I)$. This function of the unknowns is called the likelihood; 3) Calculate the probability distribution of interest, $P(H|D,I)$ using Bayes’ theorem given above. This so called posterior (distribution) will then express what is known about our hypothesis ($H$) after observing the data.

Bayes’ theorem can be used sequentially. If we first receive some data $D$, and calculate the posterior $P(H|D, I)$, and at some later point in time receive more data $D'$, the calculated posterior can be used in the role of prior to calculate a new posterior $P(H|D, D', I)$ and so on. The posterior $P(H|D, I)$ expresses all the necessary information to perform predictions. The more data we get, the more certain we will become of the unknowns, until all but one value combination for the unknowns have probabilities so close to zero that they can be neglected. Bayes’ theorem could be expressed formally as
As an example of application of Bayes’ theorem, I will present a case of a business organization ‘B’ that is employing workers for a short-term jobs that are well paid. In order to be successful in the job, an applicant needs to have a fluent writing skill in a specific language X. In the old days, all the applicants were interviewed, but nowadays it has become an impossible task as the number of both open vacancies and applicants has increased enormously. Organization’s psychometrician was given a task to develop a questionnaire that would pre-screen the most suitable applicants for the interview. Psychometrician, who developed the screening instrument, is a bit concerned about the multicultural issues of the user interface. However, he estimates that it would work correctly with 95 per cent of the applicants despite of their ethnic background. We know on the basis of publicly available demographic statistics that the language criterion is valid for one per 100 persons living in the target population. The question is: If an applicant gets enough points to participate in the interview, what is the probability that he/she is hired for the job (after an interview)?

A priori probability \( P(H) \) is described by the number of those people in the target population who are able to meet the requirements of the job (1 out of 100 = .01). More specifically, H refers to the fact that a person is able to write fluently the specific language X. Counter assumption of the a priori is \( P(\neg H) \) that equals to \( 1 - P(H) = .99 \). Psychometrician’s belief about how the instrument really works is called conditional probability, \( P(D|H) = .95 \). More specifically, \( D \) refers to the fact that the instrument reports that a person meets the specific language X writing skill requirement. Instrument’s failure to indicate any non-valid applicants, that is, those person’s who are not able to meet the criterion, is stated to be \( P(D|\neg H) \) that equals to .10. The probability \( P(H|D) \) that a person who passes the psychometrician’s test will actually get the job is .09, according to Bayes’s theorem presented in Equation 3. Why the number is so low? There are two plausible reasons: Firstly, although the screening instrument is valid, it is not perfect (10% error rate), and secondly, there are not that many valid persons in the target population (only 1%). What if we make the screening instrument...
better, resulting to only one per cent of measurement error? Then the posterior probability will increase to .49. If we further assume that there are ten out of 100 persons that have the desired linguistic ability, then the posterior probability will become as high as .91. According to Anderson (1995, p. 326), most people fail to update the estimations of even the most trivial probabilities when the situations are chancing dynamically.

4.4 The Null Hypothesis Significance Testing Procedure

One of the most important rules educational science scientific journals apply to judge the scientific merits of any submitted manuscript is that all the reported results should be based on so called ‘null hypothesis significance testing procedure’ (NHSTP) and its featured product, \( p \)-value. Gigerenzer, Krauss and Vitouch (2004, p. 392) describe ‘the null ritual’ as follows: 1) Set up a statistical null hypothesis of “no mean difference” or “zero correlation.” Don’t specify the predictions of your research or of any alternative substantive hypotheses; 2) Use 5 per cent as a convention for rejecting the null. If significant, accept your research hypothesis; 3) Always perform this procedure. I am blushing every time writing or reading this procedure as it sounds so familiar to me. This is the way I was educated, and I believe that many of my colleagues, too. Next I will elaborate the NHSTP a bit further, discuss its limitations and provide some thoughts how it should be applied in a more meaningful way in educational studies.

A \( p \)-value is the probability of the observed data (or of more extreme data points), given that the null hypothesis \( H_0 \) is true, \( P(D|H_0) \) (id.). The first common misunderstanding is that the \( p \)-value of, say \( t \)-test, would describe how probable it is to have the same result if the study is repeated many times (Thompson, 1994). Gerd Gigerenzer and his colleagues (id., p. 393) call this replication fallacy as “\( P(D|H_0) \) is confused with 1—\( P(D) \)” The second misunderstanding, shared by both applied statistics teachers and the students (Haller & Krauss, 2002), is that the \( p \)-value would prove or disprove \( H_0 \). However, a significance test can only provide probabilities, not prove or disprove null hypothesis. Gigerenzer (id., p. 393) calls this fallacy an illusion of certainty: “Despite wishful thinking, \( p(D|H_0) \) is not the same as \( P(H_0|D) \), and a significance test does
not and cannot provide a probability for a hypothesis.” A Bayesian statistics provide a way of calculating a probability of a hypothesis (discussed later in this section).

When we elaborate the use of NHSTP a bit further in the empirical educational science research context, it becomes clear that the concept of ‘a perfectly true null hypothesis’ is an impossible one. Only few concepts under investigation are totally independent by nature during the study (e.g., age, gender) and are not affected by other factors. This fact leads to a conclusion that as null hypothesis is never 100 per cent true (i.e., it can only approach the value 1.0 that indicates the full certainty), we may always reject it by increasing our sample sizes. According to Abelson (1995, p. 9), ”a null hypothesis test is a ritualized exercise of devil’s advocacy.” Murphy and Myors (1998, p. 33) explain this a bit further by saying that “The traditional null hypothesis testing is virtually always wrong because is infinitely precise, and none of the real world phenomena it is designed to test can possibly measured with that level of precision.” This statement makes sense, as one must assume as a basis for argument that there is no systematic difference between the experimental and control tests (except for errors of sampling and measurement the two groups’ performances are indistinguishable). In practice, this will never be the case, at least when participants are human beings.

Although large samples make it easier to reject any null hypothesis, they have also been proven to be a desirable feature of many rigorous scientific studies. However, it is the research problem that should be the king of a research design selection process. It should be remembered that classics, such as Sir Frederic Bartlett (1886-1969), Ivan Pavlov (1849-1936), Jean Piaget (1896-1980) and Burrhus Skinner (1904-1990), have developed their theories by observing individual participants – and without NHSTP.

It was Sir Ronald Fisher (1890-1962) who introduced the null hypothesis testing to a large social science research audience in the 1940’s. He is considered to be one of the most influential statisticians of the last century as he was able to combine various statistical techniques into a systematic theoretical model that he generously demonstrated in practice in his numerous studies (e.g., Fisher, 1935, 1956).
Fisher convinced before the Second World War most of the statisticians (and contemporary philosophers like Karl Popper) with his publicly celebrated studies showing that null hypothesis could only be rejected, not verified (e.g., Fisher, 1935). According to Fisher, probability that the data is 'true' is calculated given the hypothesis \( P(D|H_0) \), whereas Bayesians think about the inverse probability, that is the probability of a hypothesis given the data, \( P(H_0|D) \). Fisher (ibid.) suggests that a null hypothesis should be rejected when an index value of, say \( t \) or \( F \) test, is statistically significant. What is more seldom referred, is his statement that the only valid instance to set the rejection limit (usually .1%, 1% or 5%) is the researcher him/herself, not a publishing house or general opinion.

As a capable mathematician, Fisher used mainly formal symbolic language to express his thoughts. This caused a need for more concrete and practical ‘method books’, written by applied statisticians in 1950's and 1960's. The major problem with these works is that they sometimes mix seamlessly Jerzy Neyman's (1894-1981) and Egon Pearson's (1885-1980) contradictory thoughts into Fisherian models. For example, Fisher himself did only speak about a null hypothesis and nothing about an alternative hypothesis, as the latter was not present in the measurement model. The alternative hypothesis concept was taken from Neyman and Pearson’s alternative hypothesis testing model that was originally a response to Fisher’s ‘ad hoc’ approach. The single most important limit of NHSTP is that there is only one statistical hypothesis (\( H_0 \)), which does not allow for comparative hypothesis testing (Gigerenzer et al., 2004). It should be remembered that Fisher originally developed NHSTP as an ‘ad hoc’ method for those situations where extremely little about the topic was known (ibid.).

According to Gigerenzer (2000, p. 272), early Fisher suggested setting the fixed statistical significance level before the analysis: “It is usual and convenient for experimenters to take 5 per cent as a standard level of significance, in the sense that they are prepared to ignore all results which fail to reach this standard”. However, after a conflict with the Neyman-Pearson hypothesis testing model, and able critique especially from Neyman, late Fisher recommended reporting the exact value of the test: “… no scientific worker has affixed level of significance at which from year to year, in all circumstances, he rejects hypotheses …” (1956, p. 45).
Neyman and Pearson insisted that one has to specify the level of significance before the experiment as it described to them a long-run frequency of error. They argued that 5 per cent level of significance means that if the $H_0$ is correct and the experiment is repeated many times, then the experimenter will reject the hypothesis in 5 out of 100 cases. According to Gigerenzer (2000), to reject the $H_0$ when it is correct is called ‘Type I error’ and its probability is called alpha ($\alpha$). ‘Type II error’ occurs when an alternative hypothesis $H_1$ (not included in Fisher’s NHSTP) is falsely rejected. Its probability is called beta ($\beta$). Alpha is called the size of the test and 1-beta is called its power. It is obvious that we aim at high power level as it indicates the long-run frequency of accepting $H_1$ if it is true. Neyman-Pearson model is based on an assumption that one of the hypotheses is true. Murphy and Myers (1998) argue that if the null hypothesis is virtually always wrong, like we discussed earlier, it is impossible to make Type I error (rejection of $H_0$ when it is correct), and the only error that is relevant is Type II error (failing to reject $H_0$ when it is false). This leads to the fact that we should be more concerned with the statistical power than with the control of Type I errors: “Power often translates into the probability that the test will lead to a correct conclusion about the null hypothesis” (ibid., p. vii).

Gigerenzer and his colleagues (2004) provide a comparison of the three ‘tools’ mentioned above: 1) $p(D|H_0)$ is obtained from null hypothesis testing (late Fisher); 2) $p(D|H_1)/p(D|H_2)$ is obtained from Neyman-Pearson hypothesis testing; 3) $p(H_0|D)$ is obtained by Bayes’s rule. This comparison reveals nicely earlier ‘illusion of certainty’ as Bayesian approach is the only one to provide a way of calculating a probability of a hypothesis that is often the most interesting value for real-life research purposes.

Gigerenzer (2000, pp. 267-288) proposes a Freudian analogy model for the unconscious conflicts in the minds of researchers using NHSTP. The model has three components: Superego (Neyman-Pearson), Ego (Fisher) and Id (Bayes). A sample researcher, Dr. Publish-Perish, would like to get published efficiently by a) having only null hypothesis, b) significance level computed after the experiment, c) beta ignored and d) sample size determined by the rule of thumb (Fisherian thoughts rule his Ego). However, his Superego keeps whispering into his ear that he might need a) two or more hypotheses, b) both alpha and beta
determined before the experiment, c) sample size computation and d) no statements about the truthness of hypotheses. To make this conflict perfect, his Bayesian Id wants probabilities of hypotheses, which Ego and Superego can not provide.

Sir Karl Popper (1902-1994) advocated Fisher’s view that a hypothesis could only be disproved. In contrast, Harold Jeffreys (1891-1989) saw the advantage of broader view of probability (i.e., allowing the resulting numbers to vary between zero and one) in real-life decision making (Lindley, 2001). Fisher fought most time of his active career against competing opinions of both frequentists (e.g., Neyman) and bayesians (e.g., Harold Jeffreys). Due to his great influence (e.g., member of the Royal Society from 1929, a Knight Bachelor from 1952), many of his criticized thoughts (e.g., null hypothesis testing, fiducial inference) have managed to survive in applied statistical books and scientific journal policies. Later he took few steps towards bayesianism by admitting that inverse probability has some merit, but still noting that “… question, however, is whether in scientific research, and especially in the interpretation of experiments, there is cogent reason for inserting a priori” (1956, pp. 8-17). To concretize the difference between NHSTP and Bayesian approach, I will present an example that relies extensively on Berger and Wolpert (1988) and Lavine (1999).

I have been teaching elementary level statistics for educational science students for the last 12 years. I grade each student on a scale from one to five. My latest statistics course grades (Autumn 2006, n = 12) ranged from one to five as follows: 1) n = 3; 2) n = 2; 3) n = 4; 4) n = 2; 5) n = 1, showing that the lowest grade frequency (1) from the course is three (25.0%). Earlier data from the same course (2000-2005) shows that only five students out of 107 (4.7%) had the lowest grade. Next, I will use the classical statistical approach (the likelihood principle) and Bayesian statistics to calculate if the number of the lowest course grades is exceptionally high on my latest course when compared to my earlier stat courses. There are numerous possible reasons behind such development, for example, I have become more critical on my assessment or the students are less motivated in learning quantitative techniques. However, I believe that the most important difference between the last and preceding courses is that the
assessment was based on a computer exercise with statistical computations. The preceding courses were assessed only with essay answers.

I assume that the 12 students earned their grade independently (independent observations) of each other as the computer exercise was conducted under my or my assistant’s supervision. I further assume that the chance of getting the lowest grade ($\theta$), is the same for each student. Therefore $X$, the number of lowest grades (1) in the scale from 1 to 5 among the 12 students in the latest stat course, has a binomial ($12, \theta$) distribution: $X \sim \text{Bin}(12, \theta)$.

According to Equation 1, for any integer $r$ between 0 and 12,

$$P(r | \theta, n) = \binom{12}{r} \theta^r (1-\theta)^{12-r}$$

The expected number of lowest grades is $12(5/107) = 0.561$. Theta is obtained by dividing the expected number of lowest grades with the number of students: $0.561 / 12 \approx 0.05$. The null hypothesis is formulated as follows: $H_0$: $\theta = 0.05$, stating that the rate of the lowest grades from the current stat course is not a big thing and compares to the previous courses rates. Three alternative hypotheses are formulated to address the concern of the increased number of lowest grades: $H_1$: $\theta = 0.06$; $H_2$: $\theta = 0.07$; $H_3$: $\theta = 0.08$. The alternative hypotheses assume that there are six $[12(6/107)/12 \approx 0.056]$, seven $[12(7/107)/12 \approx 0.065]$ and eight $[12(8/107)/12 \approx 0.075]$ cases of the lowest grades out of 107 possible cases.

To compare the hypotheses, we calculate binomial distributions for each value of $\theta$. For example, the null hypothesis ($H_0$) calculation yields

$$P(r | \theta, n) = \binom{12}{r} 0.05^r (1-0.05)^{12-r}$$

$$= \binom{12!}{(12-3)!} 0.05^3 (1-0.05)^{12-3}$$

$$= \frac{479001600}{2177280} 0.05^3 (1-0.05)^{12-3}$$

$$= .017$$

The results for the alternative hypotheses are as follows: $P_{H1}(3, 0.06, 12) \approx .027$; $P_{H2}(3, 0.07, 12) \approx .039$; $P_{H3}(3, 0.08, 12) \approx .053$. The ratio of the hypotheses is
roughly 1:2:2:3 and could be verbally interpreted with statements like “the second and third hypothesis explain the data about equally well”, or “the fourth hypothesis explains the data about three times as well as the first hypothesis”.

Lavine (1999) reminds that \( P(r|\theta, n) \), as a function of \( r \) (3) and \( \theta \{.05; .06; .07; .08\} \), describes only how well each hypotheses explains the data; no value of \( r \) other than 3 is relevant. For example, \( P(4|0.05, 12) \) is irrelevant as it does not describe how well any hypothesis explains the data. This likelihood principle, that is, to base statistical inference only on the observed data and not on a data that might have been observed, is essential feature to Bayesian approach.

The Fisherian, so called ‘classical approach’ to test the null hypothesis (\( H_0 : \theta = .05 \)) against the alternative hypothesis (\( H_1 : \theta > .05 \)) is to calculate the \( p \)-value that defines the probability under \( H_0 \) of observing an outcome at least as extreme as the outcome actually observed:

\[
p-value = P(r = 3 | \theta = .05) + P(r = 4 | \theta = .05) + \ldots + P(r = 12 | \theta = .05)
\]

After calculations, the \( p \)-value becomes \( \approx 0.02 \) and would suggest \( H_0 \) rejection, if the rejection level of significance is set at five per cent. Another problem with the \( p \)-value is that it violates the likelihood principle by using \( P(r|\theta, n) \) for values of \( r \) other than the observed value of \( r = 3 \) (Lavine, 1999): The summands of \( P(4|0.05, 12), P(5|0.05, 12), \ldots, P(12|0.05, 12) \) do not describe how well any hypothesis explains observed data.

A Bayesian approach will continue from the same point as the classical approach, namely probabilities given by the binomial distributions, but also make use of other relevant sources of \textit{a priori} information. In this domain, it is plausible to think that the computerized test would make the number of total failures more probable than in the previous times when the evaluation was based solely on the essays. On the other hand, the computer test has only 40 per cent weight in the equation that defines the final stat course grade: 

\[
\frac{.3\text{(Essay}_1) + .3\text{(Essay}_2) + .4\text{(Computer test})}{3} = \text{Final grade.}
\]

Another aspect is to consider the nature of the aforementioned tasks, as the essays are distance work assignments while the computer test is to be performed under observation. Perhaps the course grades of my earlier stat courses have a narrower dispersion due to violence of the independent observation assumption? For example, some
students may have copy-pasted text from other sources or collaborated without a permission. As we see, there are many sources of \textit{a priori} information that I judge to be inconclusive and, thus, define that null hypothesis is as likely to be true or false. This \textit{a priori} judgment is expressed mathematically as \( P(H_0) \approx \frac{1}{2} \approx P(H_1) + P(H_2) + P(H_3) \). If I further assume that the alternative hypotheses \( H_1, H_2 \) or \( H_3 \) share the same likelihood \( P(H_1) \approx P(H_2) \approx P(H_3) \approx \frac{1}{6} \).

These prior distributions summarize the knowledge about \( \theta \) prior to incorporating the information from my course grades.

An application of Bayes' theorem (Equation 3) yields

\[
P(H_0 | r = 3) = \frac{P(r = 3 | H_0)P(H_0)}{P(r = 3 | H_0)P(H_0) + P(r = 3 | H_1)P(H_1) + P(r = 3 | H_2)P(H_2) + P(r = 3 | H_3)P(H_3)} \\
\approx \frac{P(r = 3 | 0.017)P\left(\frac{1}{2}\right)}{P(r = 3 | 0.017)P\left(\frac{1}{2}\right) + P(r = 3 | 0.027)P\left(\frac{1}{6}\right) + P(r = 3 | 0.039)P\left(\frac{1}{6}\right) + P(r = 3 | 0.053)P\left(\frac{1}{6}\right)} \\
\approx 0.30
\]

Similar calculations for the alternative hypotheses yields \( P(H_1 | r = 3) \approx .16; P(H_2 | r = 3) \approx .23; P(H_3 | r = 3) \approx .31 \). These \textit{posterior} distributions summarize the knowledge about \( \theta \) after incorporating the grade information. The four hypotheses seem to be about equally likely (.30 vs. .16, .23, .31). The odds are about 2 to 1 (.30 vs .70) that the latest stat course had higher rate of lowest grades than 0.05. The difference between the classical and Bayesian statistics would be only philosophical (probability vs. inverse probability) if they would always lead to similar conclusions. However, in this case the \( p \)-value would suggest rejection of \( H_0 \) (\( p = .02 \)), but the Bayesian analysis indicate not very strong evidence against \( \theta = .05 \), only about 2 to 1.

What if the number of the lowest grades is two? The classical approach would not anymore suggest \( H_0 \) rejection (\( p = .12 \)), the Bayesian result would stay quite the same (.39 vs .17, .20, .24), saying that there is not much evidence against the \( H_0 \). For more technical approach to the topic, reader is encouraged to see Hoijtink and Klugkist’s (2007) excellent work.
4.5 Essential Benefits of Bayesian Modeling

The essential benefits of using Bayesian modeling with empirical samples in professional growth research are summarized as follows.

*Theoretical minimum sample size is zero.* This is due to fact, that the phenomenon, not the data, is modeled. The latter is the case in traditional Gaussian modeling approach. However, Bayesian modeling is also capable of scaling up to meet the requirements of large data modeling tasks.

*Bayesian modeling is based on probabilities* allowing prediction with the model. For example, researcher is able to “fix” interesting values of variables in the Bayesian Network model, and further investigate values of other variables’ conditional distributions in the model.

*Bayesian modeling is inductive*, as the model is constructed from the data $P(M|D)$. Traditional frequentistic analysis is testing with sampling distribution the opposite assumption, if the data fits the model $P(D|M)$, that is, the probability of observing the particular data given that the null hypothesis is true. However, rejecting the null hypothesis says nothing or very little about the likelihood that the null is true. In practice this means that Bayesian statisticians are not interested in testing hypotheses in a traditional way with the $p$-value.

*Researcher is able to input a priori information to the model.* The source of subjective information could be, for example, an interview with an expert of certain topic, or previously collected data. Nokelainen, Miettinen, Kurhila, Silander and Tirri (2002) have implemented an adaptive online questionnaire that profiles users online with Bayesian probabilistic modeling. A priori profile information is used to reduce the number of questions. Investigations with numerous empirical samples suggest that if a priori information (i.e., “learning data”) is collected from the same or suchlike population, only approximately 35 per cent of the questions are needed to achieve 99 per cent accuracy in all the remaining (i.e., unasked) responses.

*Bayesian modeling is designed to analyze discrete categorical variables.* Vocational education researchers collect vast part of their comparative data with paper and pencil or web-based online surveys. The most typical question types in survey research are dichotomous and multiple-choice questions. In both cases the
categories are discrete (i.e., have no overlap and are mutually exclusive) and exhaust the possible range of responses (Cohen, Manion & Morrison, 2000, p. 251). One of the major differences between traditional Gaussian (frequentistic) and Bayesian models lies in the fact that the latter does not require multivariate normal distribution of the indicators (i.e., observed variables) or underlying phenomena. This feature is especially useful for a social science researcher who collects his/her data with, for example, Likert–scale type questions as the response options from 1 to 7 produce data that is more qualitative than quantitative in nature. Measurement level of such items is ordinal and it is not advisable to model it with traditional statistical analysis that rely on the concept of normal distribution, and require calculation of mean and standard deviation.

Bayesian modeling is able to analyze both linear and non-linear dependencies between variables. Phenomena under investigation are seldom purely linear or continuous in nature. Unfortunately most commonly applied traditional linear Gaussian models (e.g., regression and factor analysis) are statistically inadequate for understanding non-linear dependencies between variables. However, Bayesian dependency models for discrete data allow also description of non-linearities. Bayesian theory gives a simple criterion, that is, probability of the model, to select among such models (Nokelainen, Silander, Ruohotie & Tirri, 2003). Nokelainen, Tirri, Campbell and Walberg (2004) analyzed factors that account for adult productivity with three samples that came from U.S. \( n = 239 \), Germany \( n = 229 \) and Finland \( n = 159 \). Nokelainen and Tirri (2004) further investigated the number of non-linear and multi-modal relationships between variables in the three data sets. The results presented in Table 9 show that 64 per cent of all dependencies were purely linear (linear mode, linear mean, unimodal). This is the most appropriate data for traditional linear analysis as no information is lost due to non-linearity. The rest of the dependencies, 37 per cent, were to some extent non-linear. These dependencies are missed or falsely modeled using linear techniques, such as exploratory factor analysis.
Table 9. Comparison of Linear and Non-linear Dependencies in Three Empirical Samples (Nokelainen & Tirri, 2004, p. 144)

<table>
<thead>
<tr>
<th>Data</th>
<th>U.S. (% n = 239)</th>
<th>Germany (% n = 229)</th>
<th>Finland (% n = 159)</th>
<th>Total (% n = 627)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear mode, linear mean, unimodal</td>
<td>59.4</td>
<td>71.4</td>
<td>59.1</td>
<td>63.4</td>
</tr>
<tr>
<td>Linear mode, linear mean, multimodal</td>
<td>3.1</td>
<td>0.0</td>
<td>4.5</td>
<td>2.4</td>
</tr>
<tr>
<td>Linear mode, non-linear mean, unimodal</td>
<td>21.9</td>
<td>21.4</td>
<td>22.7</td>
<td>22.0</td>
</tr>
<tr>
<td>Linear mode, non-linear mean, multimodal</td>
<td>9.4</td>
<td>7.1</td>
<td>4.5</td>
<td>7.3</td>
</tr>
<tr>
<td>Non-linear mode, linear mean, unimodal</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Non-linear mode, linear mean, multimodal</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Non-linear mode, non-linear mean, unimodal</td>
<td>3.1</td>
<td>0.0</td>
<td>9.1</td>
<td>3.7</td>
</tr>
<tr>
<td>Non-linear mode, non-linear mean, multimodal</td>
<td>3.1</td>
<td>0.0</td>
<td>0.0</td>
<td>1.2</td>
</tr>
</tbody>
</table>

4.6 Relation of Bayesian Modeling to the Traditional Parametric and Non-parametric Frequentistic and Non-frequentistic Techniques

Figure 8 shows an illustration of how Bayesian modeling is related to the traditional parametric and non-parametric frequentistic techniques. Reader should notice that the Bayesian application domain is further elaborated in Table 11 in section “4.11 Operationalization of the Fourth and Fifth Research Questions”.

The left-hand side of the figure shows two main categories of data collection: Probability sample (PS) and Non-probability sample (NPS). Both methods aim to produce a scientific, representative sample from the target population. According to Jackson (2006), a representative sample is “like” the
population. Thus, we can be confident that the results we find based on the sample also hold for the population. This is not a problem with PS, which is based on random, stratified or cluster sampling. In random sampling each member of the population has an equal likelihood of being selected into the sample. Stratified random sampling allows taking into account different subgroups in the population. If the population is too large for random sampling of any sort, cluster sampling is applied. Problems arise with NPS as the individual members of the population do not have an equal likelihood of being selected to be a member of the sample. The most commonly applied NPS technique is convenience sampling (CS) in which participants are obtained wherever they can be found and wherever is convenient for the researcher (Hair, Anderson, Tatham & Black, 1998). Why, then, educational scientists use NPS, typically CS? Simply because it “tends to be less expensive [than RS] and it is easier to generate samples using this technique” (Jackson, 2006, p. 84). However, on the lower left-hand part of the Figure 8 I show that when researcher ensures that the CS is like the population on certain characteristics (location and dispersion descriptive statistics about, for example, age and job title), it becomes a quota sample (QS). A quota sample is better than a CS as it allows us to ensure that the results we find based on the sample also hold for the population.

The upper part of the Figure 8 contains two sections, namely “parametric” and “non-parametric” divided into eight sub-sections (“DNIMMOCS OLD”). As stated earlier in section 4.1, parametric approach is viable only if 1) Sample size is large enough (at least 30 observations); 2) Both the phenomenon modeled and the sample follow normal distribution; 3) Continuous indicators are used. Otherwise non-parametric techniques should be applied.

The eight sub-sections are the next steps in the path leading from the previously described sample to the statistical analyses and, finally, to the statistical dependencies (i.e., results).

First, study design (D) is made on the basis of the research question and major goal. According to de Vaus (2004, p. 9), “… research design is to ensure that the evidence obtained enables us to answer the initial question as unambiguously as possible.” Thus, in order to obtain relevant evidence, we need to specify the type of evidence needed to answer the research question. More
specifically, we need to ask: Given this research question, what type of evidence (data) is needed to answer the question in a convincing way? Sometimes we proceed with the so-called qualitative designs, sometimes a quantitative orientation is more appropriate, and sometimes we work both qualitatively and quantitatively (for a thorough discussion, see Brannen, 2004). Controlled experiment (a.k.a. ‘experimental study’) is the most recommended approach in psychological research, but only possible with a randomized sample. Far more popular approach in educational research, correlational technique (a.k.a. ‘descriptive study’), allows the use of non-probability sample. Using experimental design, both reliability and validity are maximized via random sampling and control in the given experiment (de Vaus, 2004). Most correlational designs are missing control, and thus loose some of their scientific power (Jackson, 2006). Why do, then, educational scholars use correlational designs over controlled experiments? The first answer is simple: Correlational designs are far easier, faster and inexpensive to conduct than experimental designs. The second answer is more complex as we need to ask if the controlled experiment approach is at all viable method to study educational research questions. In science and psychology, most areas of interest are quite easily quantifiable and replicable (like, for example, freezing point of chocolate or systolic blood pressure). However, in educational research we study, for example, topics like ‘pedagogical aspects of digital learning material’ (Nokelainen, 2006) or compare pre-existing characteristics of interest (e.g., gender, age, educational level). In those cases, random assignment is impossible and inappropriate. In such situations researchers do apply correlational designs, but still aim to employ different types of data in the analysis with a complementary way (quasi-experimental study). At this point Abelson’s (1995) concept of statistics as principled argument becomes useful: Data analysis should not be pointlessly formal, but instead “... it should make an interesting claim; it should tell a story that an informed audience will care about and it should do so by intelligent interpretation of appropriate evidence from empirical measurements or observations” (p. 2). There are also other ways of classifying study types (see Caskie & Willis, 2006), but they are out of scope of this discussion. I will just shortly mention the two most common types: cross-
sectional and longitudinal studies. Both aim to gather information about age differences, but the former is considerably cheaper and faster to conduct (although producing less controllable and less powerful results). Causal conclusions are usually out of scope of this research type (ibid.). Longitudinal study produces more convincing results as it allows the understanding of change in a construct over time and variability and predictors of such change over time (ibid.). However, it takes naturally more time to carry out and suffers from participant drop-out.

Second, optimal sample size (N) is divided into two sections in Figure 8: Samples that operate in the optimal area \( n \approx 30 – 250 \) for traditional parametric frequentistic techniques (Black, 1993; Tabachnick & Fidell, 1996), such as t-test or exploratory factor analysis, and the samples that fail to do so \( n < 30 \) or \( n > 250 \). Traditional non-parametric techniques, such as Mann-Whitney U-test, are considered to operate robustly, also with small samples. Bayesian approach, however, is free of such restrictions (discussed later in section 4.5).

Third, independent observations (IO) are always expected, also in time series analysis. Controlled experiment designs, when conducted properly, rule out IO violations quite effectively (Martin, 2004), but correlational designs usually lack such control (e.g., to rule out employee’s co-operation when they respond to the survey questions). On the other hand, some qualitative techniques, like focus group analysis (Macnaghten & Myers, 2004), are heavily based on non-independent observations as informants are asked to talk to each other as an important part of the data collection.

Fourth, parametric techniques assume continuous \( (c) \) measurement level (ML) of indicators (i.e., so called ‘quantitative’ variables). Non-parametric analysis is based on ordering of values and thus discrete \( (d) \) or, when applicable, nominal \( (n) \) values are expected (i.e., so called ‘qualitative’ variables). A respondent’s income level (euros) or age (years or months) is a representative example of the first indicator type. A Likert-scale from 1 to 5 is an example of the second indicator type (ordered discrete values). Respondent’s gender is an example of the third indicator type (nominal discrete values). It is important to note that the central limit theorem, discovered by Pierre-Simon Laplace (1749 - 1827), assures an approximate normal distribution for practically all sums of
independent random variables, for example, allowing the use of parametric t-test with binominal or ordinal indicators (as the sample of normally distributed group means are compared, not the indicator values themselves). Bayesian analysis is based on discrete values, and thus, continuous values must be disceticized (automatically or manually) before the analysis. This issue is further discussed in section 6.2.

_Fifth_, parametric techniques are technically based on the assumption of the multivariate distribution (MD) that is normal (n) by nature. Non-parametric techniques expect any shaped similar distributions (s). This is a great news to anyone who has collected real-life educational science empirical data and checked both univariate and multivariate variable distributions as usually almost all variables violate quite heavily against the normal distribution assumption with small sample sizes (e.g., below \( n = 100 \)). Some researchers try to force their indicators to follow multivariate normal distribution by applying various transformation techniques (e.g., logarithmic, square), but with varying success. The motivation for transformations lies behind the fact that in order to enable parametric analysis (i.e., based on, e.g., normal distribution) the bivariate or multivariate statistical dependencies (S) must be linear (l). It is important to note that this assumption does not hold for the Bayesian techniques.

_Sixth_, extreme values, outliers (O), affect the results and, thus, the conclusions, of some parametric techniques severely (e.g., regression and discriminant analysis) and should be recognized and removed (see, e.g., Tabachnick & Fidell, 1996). Non-parametric analysis techniques are not affected by such values as their analysis is not based on multivariate normal assumption (i.e., linear dependencies between variables).

_Seventh_, when calculating correlations (C), Pearson product moment correlation \( (r_P) \) should be applied with continuous indicators, and Spearman rank-order correlation \( (r_S) \) with ordinal indicators. Both techniques are valid to detect linear dependencies.

The last point in Figure 8 is to discuss about the two types of statistical dependencies (S) among the variables under analysis, namely linear (l) and non-linear (nl). It is natural to assume, that both parametric and non-parametric techniques designed to detect linear dependencies work best with samples that
contain linear dependencies. However, there are non-linear techniques, such as Bayesian analysis and neural networks that also allow the investigation of both dependency types.

Figure 8 contains also a reference to the qualitative analysis techniques, referring here to the empirical textual evidence based approach (e.g., individual or focus group interviews, narrative stories). My goal is to show how these techniques are located in the Figure 8, and to point out their quite close (at least in philosophical level) relativity to the Bayesian techniques. Firstly, it is obvious that qualitative research operates with small samples (usually \( n < 30 \)). I think that there is nothing suspicious working with small samples, as I have given earlier in section 4.4 numerous examples of famous ‘small sample scholars’ (Bartlett, Pavlov, Piaget and Skinner). Secondly, probability samples could also be used by qualitative researchers (as I have stated in Figure 8), but not as the only way to produce scientifically important findings. Gobo (2004) illustrates this by listing important qualitative research studies based solely on non-probability samples: Alvin Gouldner (1920-1980) observed one small gypsum extraction and refining factory located close to his university (haphazard sample); Howard Becker (1928-) studied dance musicians; Ernest De Martino (1908-1965) observed 21 people suffering from tarantism disease; ethnomethodologists David Sudnow (1938-2007) and Aaron Cicourel (1928-) observed two hospitals and two police districts respectively. Gobo defines a new concept of generalizability for qualitative research by arguing that the concept of generalizability is based on the idea of social representativeness, which allows the generalizability to become a function of the invariance (regularities) of the phenomenon. Thus, “The ethnographer does not generalize one case or event … but its main structural aspects that can be noticed in other cases or events of the same kind or class.” (id., p. 453.) Thirdly, both qualitative and Bayesian analysis techniques allow researcher to apply a priori input to the modeling process and update the model on the basis of increased level of knowledge.
Figure 8. Relation of Bayesian Modeling to the Traditional Parametric and Non-parametric Frequentistic and Non-frequentistic Techniques

Note. Abbreviations on the left hand side of the figure: PS = Probability sample, based on randomized sampling. NPS = Non-probability sample. PP = Population known parameters (location and dispersion descriptive statistics about, e.g., age and job title) are compared to sample parameters. QS = Quota sample, when population and NPS parameters are comparable.

Abbreviations on the upper part of the figure: D = Design (ce = controlled experiment, co = correlational study). N = Numerus (n = optimal sample size). IO = Independent observations. ML = Measurement level (c = continuous, d = discrete, n = nominal). MD = Multivariate Distribution (n = all normal, s = all similar). O = Outliers (i.e., extreme values) removed. C = Correlation ($r_p$ = continuous; $r_s$ = discrete). S = Statistical dependency (l = linear, nl = non-linear).
4.7 Bayesian Dependency Modeling

Bayesian dependency modeling (BDM) predicts the most probable statistical dependency structure between the observed variables. It visualizes the result in a form of Bayesian network allowing user to probe the model by adjusting the values of all variables and examining the effects to other variables included in the best model. Examples of typical research questions in the field of professional growth research utilizing this technique are as follows: “Are Encouraging Leadership and Team Spirit leading to Commitment Towards Work?”, “Are the Abilities for Professional Growth Questionnaire items measuring all the six dimensions of Learning Motivation in this domain?” and “What is the most probable Growth Orientation variable structure in this domain?” (Nokelainen & Ruohotie, 2002, 2003a, 2003b; Nokelainen, Ruohotie, Miettinen, Kurhila & Tirri, 2003; Ruohotie & Nokelainen, 2002, 2003.)

A Bayesian network (BN) is a viable way to examine dependencies between variables by both their visual representation and probability ratio of each dependency. It is a representation of a probability distribution over a set of random variables, consisting of a directed acyclic graph (DAG), where the nodes correspond to domain variables, and the arcs define a set of independence assumptions which allow the joint probability distribution for a data vector to be factorized as a product of simple conditional probabilities.

A graphical visualization of BN (Heckerman, Geiger & Chickering, 1995) contains two components: 1) Observed variables visualized as ellipses; 2) Dependences visualized as lines between nodes. The darker the line, the more stronger statistical dependency between the two variables, and more important the dependency is for the model. Variable is considered as independent of all other variables if there is no line attached to it. In Figure 9, we see a typical BN with two observed variables.
Next, I will present how Bayesian score (BS), that is, the probability of the model $P(M|D)$, is firstly calculated and secondly compared for the two models presented in Figure 9. Let us assume that we have the following data:

\[
\begin{array}{c|c}
\text{x1} & \text{x2} \\
1 & 1 \\
1 & 1 \\
2 & 2 \\
1 & 2 \\
1 & 1 \\
\end{array}
\]

Model 1 ($M_1$) represents the two variables, x1 and x2 respectively, without statistical dependency, and the model 2 ($M_2$) represents the two variables with a dependency (i.e., with a connecting arc). The binomial data might be a result of an experiment, where the five participants have solved a job related task before (x1) and after (x2) a vocational training period.

In order to calculate $P(M_1,2|D)$, we need to solve $P(D|M_{1,2})$ for the two models $M_1$ and $M_2$. Probability of the data given the model is solved by using the following marginal likelihood equation (Congdon, 2001, p. 473; Myllymäki, Silander, Tirri, & Uronen, 2001; Myllymäki & Tirri, 1998, p. 63):

\[
P(D | M) = \prod_{i=1}^{n} \prod_{j=1}^{q_i} \frac{\Gamma(N'_{ij})}{\Gamma(N'_{ij} + N_{ij})} \prod_{k=1}^{r_i} \frac{\Gamma(N'_{ijk} + N_{ijk})}{\Gamma(N'_{ijk})}
\]

(4)

In the Equation 4, following symbols are used: $n$ is the number of variables (i indexes variables from 1 to $n$); $r_i$ is the number of values in $i$:th variable ($k$ indexes these values from 1 to $r_i$; $q_i$ is the number of possible configurations of parents of $i$:th variable; $N_{ij}$ describes the number of rows in the data that have $j$:th configuration for parents of $i$:th variable; $N_{ijk}$ describes how many rows in the data have $k$:th value for the $i$:th variable also have $j$:th configuration for parents.
of I:th variable; \( \hat{N} \) is the equivalent sample size (ESS) set to be the average number of values divided by two:

\[
ESS = \frac{|V_1| + |V_2| + \ldots + |V_N|}{2N}
\]

(5)

In the sample data we have two possible values \{1,2\} resulting in the following ESS:

\[
ESS = \frac{2 + 2}{2 \times 2} = 1
\]

The marginal likelihood equation produces a Bayesian Dirichlet score that allows model comparison (Heckerman et al., 1995; Tirri, 1997; Neapolitan & Morris, 2004).

First, I will calculate \( P(D|M_1) \) given the values of variable \( x_1 \):

\[
P(D_{x_1} | M_1) = \frac{\Gamma(\hat{N}^{'}_{q_i})}{\Gamma(\hat{N}^{'}_{q_i} + N_{ij})} \frac{\Gamma(\hat{N}^{'}_{r*q} + N_{jik_1})}{\Gamma(\hat{N}^{'}_{r*q})} \frac{\Gamma(\hat{N}^{'}_{r*q} + N_{jik_2})}{\Gamma(\hat{N}^{'}_{r*q})}
\]

\[
= \frac{\Gamma(1.00)}{\Gamma(1.00 + 5)} \frac{\Gamma(0.50 + 4)}{\Gamma(0.50)} \frac{\Gamma(0.50 + 1)}{\Gamma(0.50)}
\]

\[
\approx 0.008 \times 6.563 \times 0.500
\]

\[
\approx 0.027
\]

Second, the values for the \( x_2 \) are calculated:

\[
P(D_{x_2} | M_1) = \frac{\Gamma(\hat{N}^{'}_{q_i})}{\Gamma(\hat{N}^{'}_{q_i} + N_{ij})} \frac{\Gamma(\hat{N}^{'}_{r*q} + N_{jik_1})}{\Gamma(\hat{N}^{'}_{r*q})} \frac{\Gamma(\hat{N}^{'}_{r*q} + N_{jik_2})}{\Gamma(\hat{N}^{'}_{r*q})}
\]

\[
= \frac{\Gamma(1.00)}{\Gamma(1.00 + 3)} \frac{\Gamma(0.50 + 4)}{\Gamma(0.50)} \frac{\Gamma(0.50 + 2)}{\Gamma(0.50)}
\]

\[
\approx 0.008 \times 1.875 \times 0.750
\]

\[
\approx 0.012
\]

The BS, probability for the first model \( P(M_1|D) \), is 0.027 \* 0.012 \approx 0.000324.

Third, \( P(D|M_2) \) is calculated given the values of variable \( x_1 \):
\[ P(D_{x1} \mid M_2) = \frac{\Gamma\left(\frac{N}{q_i}\right) \Gamma\left(\frac{N_{i,j,k}}{r \ast q}ight) \Gamma\left(\frac{N_{i,j}}{r \ast q}ight)}{\Gamma\left(\frac{N}{q_i} + N_{y}\right) \Gamma\left(\frac{N_{i,j,k}}{r \ast q} + N_{y,c}\right) \Gamma\left(\frac{N_{i,j}}{r \ast q} + N_{y,d}\right)} \]

\[ \approx \frac{\Gamma(1.00) \Gamma(0.50 + 4) \Gamma(0.50 + 1)}{\Gamma(1.00 + 5) \Gamma(0.50) \Gamma(0.50)} \]

\[ \approx 0.008 \ast 6.563 \ast 0.500 \]

\[ \approx 0.027 \]

Fourth, the values for the first parent configuration (x1 = 1) are calculated:

\[ \approx \frac{\Gamma(0.50) \Gamma(0.25 + 3) \Gamma(0.25 + 1)}{\Gamma(0.50 + 4) \Gamma(0.25) \Gamma(0.25)} \]

\[ \approx 0.152 \ast 0.703 \ast 0.250 \]

\[ \approx 0.027 \]

Fifth, the values for the second parent configuration (x1 = 2) are calculated:

\[ \approx \frac{\Gamma(0.50) \Gamma(0.25 + 0) \Gamma(0.25 + 1)}{\Gamma(0.50 + 1) \Gamma(0.25) \Gamma(0.25)} \]

\[ \approx 2.000 \ast 1.000 \ast 0.250 \]

\[ \approx 0.500 \]

The BS, probability for the second model P(M2|D), is 0.027 * 0.027 * 0.500 \approx 0.000365.

Bayes’ theorem enables the calculation of the ratio of the two models, M_1 and M_2. As both models share the same a priori probability, P(M_1) = P(M_2), both probabilities are canceled out. Also the probability of the data P(D) is canceled out in the following equation as it appears in both formulas in the same position:

\[ \frac{P(M_1 \mid D)}{P(M_2 \mid D)} = \left[ \frac{P(D \mid M_1) P(M_1)}{P(D \mid M_2) P(M_2)} \right] \approx \frac{0.000324}{0.000365} \approx 0.88 \]

The result of model comparison shows that since the ratio is less than 1, the M_2 is more probable than M_1. This result becomes explicit when we investigate the sample data more closely. Even a sample this small (n = 5) shows
that there is a clear tendency between the values of \( x_1 \) and \( x_2 \) (four out of five value pairs are identical).

However, we need to acknowledge that an interpretation of a BN depends mostly of the study design. If we have a causal dependency design (e.g., drug effect study), the interpretation is naturally causal. Quite often educational measurements are based on self-report measures and, thus, do not allow causal interpretations. If the nodes in the BN are connected with a solid arc (like in our sample model 2), researcher is allowed to believe that their dependency is not affected by a latent cause. If the arc is marked with a dashed line, a possible third factor is causing at least some of the change in the dependency. This holds true, of course, only with causal designs as correlational designs allow only naïve interpretation of the BN (i.e., assumption that there are no latent causes in the network). For further discussion on this topic, see Pearl (2000b).

4.8 Bayesian Classification Modeling

In the professional growth research field a typical research task is to predict effectively a certain interesting group membership. For example, researcher might want to solve from an empirical data “how workers with different backgrounds differ from each other in their growth-orientations?” Further, he or she might project the investigation towards the common characteristics (growth-orientations that are similar, i.e., do not predict differences) between the company’s workers or predictive characteristics (growth-orientations that point out differences) between the company’s employees. (Nokelainen, Tirri, Campbell & Walberg, 2004; Nokelainen, Tirri & Merenti-Välimäki, 2007; Ruohotie, Nokelainen & Tirri, 1999.)

Linear discriminant analysis (LDA) is a tool for classifying cases into different groups with a better than chance accuracy (Huberty, 1994, p. 118-126). LDA is technically close to both analysis of variance (ANOVA) and multivariate analysis of variance (MANOVA). With ANOVA, one may ask whether or not two or more groups are significantly different from each other with respect to the mean of a particular variable. With MANOVA, the question is whether group
membership is associated with reliable mean differences on a combination of dependent variables.

The linear combination of variables (discriminant function) is similar to the right side of a multiple regression equation because it sums the products of variables multiplied by coefficients. Then procedure estimates the coefficients and the resulting function can be used to classify new cases. Next I will shortly describe the four main applications of LDA. The most common application of discriminant function analysis is to let selection technique to determine the subset of variables that constitutes relevant model. Entry or removal decision can be made with forward, backward or stepwise technique.

Forward selection first removes all variables from the model and after that starts enter them one by one. The first variable entered at step one is the one with the strongest correlation with the dependant (classification) variable. At each subsequent step the variable with the strongest partial correlation enters the model. The hypothesis that the coefficient of the entered variable is zero is tested using its $F$ statistic. Stepping stops when an established criterion for the $F$ no longer holds.

For each candidate predictor variable, $F$ statistic is computed that measures the change in Wilks’ Lambda when variable is added to the list of accepted variables. The variable with the largest $F$ enters the list. It was just announced that stepping stops when an established criterion for the $F$ no longer holds. Now it is time to check what exactly is that established criterion.

The $F$ value for the change in Wilks’ Lambda when a variable is added to a model that contains $p$ independent variables is

$$F_{\text{change}} = \left[ \frac{n-g-p}{g-1} \right] \frac{(1-\hat{\lambda}_{p+1}/\hat{\lambda}_p)}{\hat{\lambda}_{p+1}/\hat{\lambda}_p}$$  \hspace{1cm} (6)$$

where $n$ is the total number of cases, $g$ is the number of groups, $\lambda_p$ is Wilks’ Lambda before adding variable, and $\lambda_{p+1}$ is Wilks’ Lambda after inclusion (SPSS, 1997).

Backward technique is opposite to forward selection. The idea is to start from situation where all variables are in the model and after that, step by step,
removing least useful predictor, end up in a situation where only the strongest predictors exist.

Stepwise selection is identical to forward selection except for one point; at each step already entered variables are tested for removal. This makes sense because entry of third variable can diminish the importance of an already entered variable.

Even though mathematically MANOVA and LDA are the same, classification is a major extension of LDA over MANOVA. The point is to find out how well we can predict to which group a particular case belongs.

Importance of variable selection becomes more important when we discuss about problem of over fitting. In that case one has to be aware not to include too many variables in the model. The problem seems to be the fact that such a model will not predict correctly when applied to a new sample.

We might draw a conclusion about variable selection that selection techniques are useful tools to find the most important variables describing the phenomenon we are studying. These tools also help us to drop down the number of variables to help us avoid the problem of over fitting. As we analyze the results of variable selection we might consider the technique that comes along with least number of variables as the best predictor. This all comes down to fact that it is a lot easier to report common factors of five than twenty variables.

Primary goal for many researchers that use discriminant analysis is to find discriminant function to predict group membership. The most important issues we can ask are: Can group membership be predicted reliably from the set of predictors? What is the number of significant discriminant functions? What are the dimensions of Discrimination?

Criteria for evaluating overall statistical reliability is based on multivariate tests e.g., Wilks’ Lambda, Hotelling’s trace criterion and Pillai’s criterion. Here I concentrate on describing major features of Wilks’ Lambda (see equation below). Wilks’ Lambda is a likelihood ratio statistic that tests the likelihood of the data under the assumption of equal population mean vectors for all groups against the likelihood under the assumption that population mean vectors are identical to those of the sample mean vectors for the different groups. When
testing equality of group centroids it varies between 0 and 1. Small values indicate that the group means differ.

\[
\Lambda = \frac{S_{\text{error}}}{S_{\text{effect}} + S_{\text{error}}} \tag{7}
\]

That makes Wilks’ Lambda as the pooled ratio of effect variance to error variance to effect variance plus error variance (SPSS, 1997; Tabachnick & Fidell, 1996). On comparison, Pillai’s criterion is simply the pooled effect variances. When separation of groups is distributed over dimensions, Pillai’s criterion is more adequate. Most research reports use Wilks’ Lambda unless there is reason to use Pillai’s criterion.

For two groups one discriminant function (often called Fisher discriminant function after Sir R. A. Fisher, see section 4.4 for more details) is computed, for three groups two discriminant functions are possible.

As stated earlier, computationally linear discriminant function analysis is analogous to MANOVA. For two variables, discriminant function is an equation of a plane where we fit following linear equation between two groups:

\[
D = d_1z_1 + d_2z_2 + \ldots + d_mz_m \tag{8}
\]

Discriminant function score \(D\) is found by multiplying the standardized score on each predictor \(z\) by its standardized discriminant function coefficient \(d\) and the adding the products for all predictors (Tabachnick & Fidell, 1996).

When analyzing more than two groups, one has to focus on canonical variables. The first canonical variable is the linear combination of the variables that maximizes the differences between the means of the \(n\) groups in one dimension. The second canonical variable represents the maximum dispersion of the means in a direction orthogonal to the first direction.

Interpreting canonical functions is somewhat similar to factor analysis, comparing canonical variables as factors that discriminate optimally among the group centroids relative to the dispersion within the groups. There are two main paths to examine canonical functions; first one can compare (e.g., in a table) the second canonical variable against the first. This provides an easy way to display group differences. Second path is to look factor structure for which variables define best a particular discriminant function. The factor structure coefficients
are the correlations between the variables in the model and the canonical functions.

After discriminant functions have been derived we can predict with classification functions to which group particular case belongs. Classification functions are not to be confused with the discriminant functions. There are as many classification functions as there are groups. Each function allows us to compute classification scores (which can be stored as new variables) for each case for each group with the following formula:

\[ C = c_0 + c_1X_1 + c_2X_2 + \ldots + c_mX_m \]  

(9)

A score on the classification function group \(C\) is found by multiplying the score on each predictor \(X\) by its associated classification function coefficient \(c\), summing over all predictors, and adding a constant \(c_0\). (ibid.)

When considering predictive classification of cases one must ask: What is adequacy of classification? This means in the first place that researcher must evaluate the classification results based on his/her previous knowledge on subject. Another way to approach this problem is to compare cases correctly classified by classification procedure to those obtained by chance alone. This is done simply by making allegation that 50 per cent of the cases are correctly classified by chance alone when there are two groups and 33 per cent with three groups. This conclusion assumes that all groups are equal by size.

Bayesian classification modeling (BCM) resembles the traditional LDA described earlier, but the implementation is totally different in several ways. A variable selection problem that is addressed with forward, backward or stepwise selection procedure in LDA is replaced in BCM with a genetic algorithm approach (e.g., Hilario, Kalousisa, Pradosa & Binzb, 2004; Hsu, 2004). The genetic algorithm approach means that variable selection is not limited to one (or two or three) specific approach; instead many approaches and their combinations are exploited (e.g., Cormen, Leiserson & Rivest, 1996). One possible approach is to begin with the presumption that the models (i.e., possible predictor variable combinations) that resemble each other a lot (i.e., have almost the same variables and discretizations) are likely to be almost equally good. This leads to a search strategy in which models that resemble the current best model are selected for
comparison, instead of picking models randomly. Another approach is to abandon the habit of always rejecting the weakest model and instead collect a set of relatively good models. The next step is to combine the best parts of these models so that the resulting combined model is better than any of the original models. B-Course is capable of mobilizing many more viable approaches, for example, rejecting the better model (algorithms like hill climbing, simulated annealing) or trying to avoid picking similar model twice (tabu search). BCM is depicted in detail in Silander and Tirri (1999).

Nokelainen, Ruohotie and Tirri (1999, p. 113) summarized the assumptions of linear and non-linear classification analysis in Table 10. I would like to stress that both LDA and BCM approaches allow class prediction of a new data vector, but only latter technique is capable of analyzing both linear and non-linear dependencies between variables during model-building phase. In practice this means that Bayesian approach is also able to model a phenomenon that is non-linear by nature.

Table 10. Comparing the Assumptions of Linear Discriminant Analysis and Non-linear Bayesian Classification Modeling

<table>
<thead>
<tr>
<th></th>
<th>LDA</th>
<th>BCM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size</td>
<td>At least 20 cases in smallest group.</td>
<td>At least two cases.</td>
</tr>
<tr>
<td>Unequal Sample Sizes</td>
<td>No effect. Highly unequal sample sizes are not recommended for classification.</td>
<td>No effect.</td>
</tr>
<tr>
<td>Missing Data</td>
<td>Reflects as a problem of unequal n.</td>
<td>No effect.</td>
</tr>
<tr>
<td>Multivariate Normality</td>
<td>Normal distribution. No skewness allowed.</td>
<td>No effect.</td>
</tr>
<tr>
<td>Outliers</td>
<td>Major effect. Test for univariate and multivariate outliers for each group separately must be performed.</td>
<td>No effect.</td>
</tr>
<tr>
<td>Linearity</td>
<td>Linear relationships assumed. Violation leads to reduced power.</td>
<td>No effect.</td>
</tr>
</tbody>
</table>
4.9 Bayesian Unsupervised Model-based Visualization

Unsupervised techniques (e.g., exploratory factor analysis) discover variable structure from the evidence of the data matrix as opposite to supervised techniques (e.g., discriminant analysis) that assume a given structure (Venables & Ripley, 2002, p. 301). Unsupervised techniques are further divided into four subcategories: 1) Visualization techniques; 2) Cluster analysis; 3) Factor analysis; 4) Discrete multivariate analysis.

According to Venables and Ripley (id.), visualization techniques are often more effective than clustering techniques discovering interesting groupings in the data, and they avoid the danger of over-interpretation of the results as researcher is not allowed to input number of expected latent dimensions. In cluster analysis the centroids that represent the clusters are still high-dimensional, and some additional illustration techniques are needed for visualization (Kaski, 1997), for example MDS (Kim, Kwon & Cook, 2000).

Several graphical means have been proposed for visualizing high-dimensional data items directly, by letting each dimension govern some aspect of the visualization and then integrating the results into one figure. These techniques can be used to visualize any kinds of high-dimensional data vectors, either the data items themselves or vectors formed of some descriptors of the data set like the five-number summaries (Tukey, 1977).

Simplest technique to visualize a data set is to plot a “profile” of each item, that is, a two-dimensional graph in which the dimensions are enumerated on the x-axis and the corresponding values on the y-axis. An alternative is a scatter plot where two original dimensions of the data are chosen to be portrayed as the location of an icon, and the rest of the dimensions are depicted as properties of the icon. For example the lengths of rays emanating from the center of the icon may visualize the values of the rest of the components. Also the familiar pie diagrams can be used. The major drawback that applies to all these techniques is that they do not reduce the amount of data. If the data set is large, the display consisting of all the data items portrayed separately will be incomprehensible. (Kaski, 1997.)
Techniques reducing the dimensionality of the data items are called projection techniques. The goal of the projection is to represent the input data items in a lower-dimensional space in such a way that certain properties of the structure of the data set are preserved as faithfully as possible. The projection can be used to visualize the data set if a sufficiently small output dimensionality is chosen. (id.)

Next, I shall review briefly two seminal subtypes of projection techniques: 1) Multidimensional scaling (MDS); 2) Bayesian unsupervised model-based visualization (BUMV).

Projection techniques are divided into two major groups, linear and non-linear projection techniques. Linear projection techniques consist of principal component analysis (PCA) and projection pursuit. In exploratory projection pursuit (Friedman, 1987) the data is projected linearly, but this time a projection, which reveals as much of the non-normally distributed structure of the data set as possible is sought. This is done by assigning a numerical “interestingness” index to each possible projection, and by maximizing the index. The definition of interestingness is based on how much the projected data deviates from normally distributed data in the main body of its distribution. Non-linear projection techniques consist of multidimensional scaling, principal curves and various other techniques including SOM, neural networks and Bayesian supervised (Kontkanen, Lahtinen, Myllymäki, Silander & Tirri, 2000) and unsupervised (Kontkanen, Lahtinen, Myllymäki & Tirri, 2000) networks.

Aforementioned PCA technique, despite its popularity, cannot take into account non-linear structures, structures consisting of arbitrarily shaped clusters or curved manifolds since it describes the data in terms of a linear subspace. Projection pursuit tries to express some non-linearities, but if the data set is high-dimensional and highly non-linear it may be difficult to visualize it with linear projections onto a low-dimensional display even if the “projection angle” is chosen carefully (Friedman, 1987).

Several approaches have been proposed for reproducing non-linear higher-dimensional structures on a lower-dimensional display. The most common techniques allocate a representation for each data point in the lower-dimensional space and try to optimize these representations so that the distances between
them would be as similar as possible to the original distances of the corresponding data items. The techniques differ in how the different distances are weighted and how the representations are optimized. (Kaski, 1997.)

Multidimensional scaling (MDS) is not one specific tool, instead it refers to a group of techniques that is widely used especially in behavioral, econometric, and social sciences to analyze subjective evaluations of pairwise similarities of entities. The starting point of MDS is a matrix consisting of the pairwise dissimilarities of the entities. The basic idea of the MDS technique is to approximate the original set of distances with distances corresponding to a configuration of points in a Euclidean space.

If each item \( X_k \) is represented with a lower-dimensional two-dimensional data vector \( X_k \), then the goal of the projection is to optimize the representations so that the distances between the items in the two-dimensional space will be as close to the original distances as possible. If the distance between \( X_k \) and \( X_i \) is denoted by \( d(k,l) \) and the distance between \( X_k \) and \( X_i \) in the two-dimensional space by \( d'(k,l) \), the metric MDS tries to approximate \( d(k,l) \) by \( d'(k,l) \). If a square-error cost is used, the objective function to be minimized can be written as

\[
E_M = \sum_{k=1}^{X} \left[d(k,l) - d'(k,l)\right]^2. \tag{10}
\]

SPSS employ ALSCAL algorithm to perform MDS analyses. The ALSCAL program begins by estimating an additive constant \( c \) to ensure that the triangle inequality holds for all triples among the original dissimilarities \( \delta_s \) (Jobson, 1992, p. 595). The scaling matrix is computed with Equation 11

\[
a_{rs} = -\frac{1}{2} \left[\delta_r^2 - \delta_s^2 - \delta_r^2 - \delta_s^2\right], \quad r,s = 1,2,\ldots,n, \tag{11}
\]

where

\[
\delta_r^2 = \frac{1}{n} \sum_{s=1}^{n} \delta_{sr}^2,
\]

\[
\delta_s^2 = \frac{1}{n} \sum_{r=1}^{n} \delta_{rs}^2,
\]

and
A measure used to determine goodness of fit is named STRESS and is given by

$$\delta^2 = \frac{1}{n^2} \sum_{r=1}^{n} \sum_{s=1}^{n} \delta_{rs}^2.$$ (id., p. 570.)

$$\text{STRESS} = \sum_{r<s} \left[ \left( d_{rs}^{(0)} - \hat{d}_{rs}^{(0)} \right)^2 \right] / \sum_{r<s} \sum \hat{d}_{rs}^{(0)^2}, \quad (12)$$

which is the sum-of-squared deviations normalized by $\sum_{r<s} \sum d_{rs}^{(0)^2}$ (id., p. 590). Ideally the stress value of the solution should be less than 0.10; however, the stress value is also affected by the number of fitted measures and the amount of error or noise in the data. Two dimensional solution requires $(4p-1, \ p = \text{dimensionality})$ seven data points and three dimensional solution 11 data points. Ruohotie and Nokelainen (2000a) provide a practical example with 14 data points.

MDS can be considered to be an alternative to factor analysis. In general, the goal of the analysis is to detect meaningful underlying dimensions that allow the researcher to explain observed similarities or dissimilarities (distances) between the investigated objects. In factor analysis, the similarities between objects (e.g., variables) are expressed in the correlation matrix.

With MDS we may analyze any kind of similarity or dissimilarity matrix, in addition to correlation matrices, specifying that we want to reproduce the distances based on $n$ dimensions. After formation of matrix MDS attempts to arrange “objects” (e.g., factors of growth-oriented atmosphere) in a space with a particular number of dimensions so as to reproduce the observed distances. As a result, the distances are explained in terms of underlying dimensions.

MDS based on Euclidean distance do not generally reflect properly to the properties of complex problem domains. In real-world situations the similarity of two vectors is not a universal property; in different points of view they in the end may appear quite dissimilar (Kontkanen, Lahtinen, Myllymäki, Silander & Tirri, 2000). Another problem with the MDS techniques is that they are computationally very intensive for large data sets.

Bayesian unsupervised model-based visualization (BUMV) is based on Bayesian Networks (BN). BN is a representation of a probability distribution.
over a set of random variables, consisting of a directed acyclic graph (DAG),
where the nodes correspond to domain variables, and the arcs define a set of
independence assumptions which allow the joint probability distribution for a
data vector to be factorized as a product of simple conditional probabilities. Two
vectors are considered similar if they lead to similar predictions, when given as
input to the same Bayesian network model. (Kontkanen, Lahtinen, Myllymäki,
Silander & Tirri, 2000.)

When BUMV is compared to MDS, following major differences should be
considered (Kontkanen, Lahtinen, Myllymäki, Silander & Tirri, 2000;  
Kontkanen, Lahtinen, Myllymäki & Tirri, 2000): 1) Bayesian approach is
parameter-free and the user input is not required, instead, prior distributions of
the model offer a theoretically justifiable technique for affecting the model
construction; 2) Bayesian techniques work with probabilities and can hence be
expected to produce smooth and robust visualizations with discrete data
containing nominal and ordinal attributes; 3) Bayesian approach has no limit for
minimum sample size; 4) Bayesian techniques do not assume multivariate
normal model.

Naturally, there are numerous viable options to BUMV, such as Self-
Organizing Map (SOM) and Independent Component Analysis (ICA). SOM is a
neural network algorithm that has been used for a wide variety of applications,
mostly for engineering problems but also for data analysis (Kohonen, 1995).
SOM is based on neighborhood preserving topological map tuned according to
geometric properties of sample vectors. ICA minimizes the statistical
dependence of the components trying to find a transformation in which the
components are as statistically independent as possible (Hyvärinen & Oja, 2000).
The usage of ICA is comparable to PCA where the aim is to present the data in a
manner that facilitates further analysis. The major difference between Bayesian
and neural network approaches for educational science researcher is that the
former operates with a familiar symmetrical probability range from 0 to 1 while
the upper limit of asymmetrical probability scale in the latter approach is
unknown.
4.10 Bayesian Modeling in Practice

Most of the analysis in this dissertation is conducted with web-based online data analysis tools, namely B-Course\(^1\) (Myllymäki, Silander, Tirri & Uronen, 2001, 2002) and BayMiner\(^2\) (Kontkanen, Lahtinen, Myllymäki, & Tirri, 2000). Both programs and their variations are developed by the researchers of the Complex Systems Computation Group (CoSCo, http://cosco.hiit.fi) in various projects, such as Finnish Academy funded Computational Intelligence Techniques for Non-linear Modeling in Social Sciences (NONE)\(^3\). Next, I will present the essential features of the B-Course software as it is the major analysis tool in three out of four original publications.

4.10.1 B-Course

B-Course allows the users to analyze their data for multivariate probabilistic dependencies. These dependencies are represented as graphical models known as Bayesian networks. Although the analysis techniques, modeling assumptions and restrictions are totally transparent to the user, this transparency is not achieved at the expense of analysis power. With the restrictions stated in the online material, B-Course is a powerful analysis tool exploiting several theoretically elaborated results developed recently in the fields of Bayesian and causal modeling.

B-Course can be used with most web-browsers, and the facilities include features such as automatic missing data handling and discretization, a flexible graphical interface for probabilistic inference on the constructed Bayesian network models, automatic printed layout for the networks (PNG and PostScript formats), exportation of the constructed models, and analysis of the importance of the derived dependencies. (Tirri & Silander, in press.)

The advances of B-Course user interface design aspects, when compared to traditional statistical modeling software such as SPSS, are listed as follows (Nokelainen & Tirri, 2004; Tirri & Silander, in press):

No parameters: B-Course is aimed to researchers or students that either are taking (or have taken) an accompanying course in Bayesian modeling, or have

---

\(^1\) B-Course is available at http://b-course.hiit.fi
\(^2\) BayMiner is available at http://bayminer.com
\(^3\) http://cosco.hiit.fi/Projects/NONE
some background in the topic. The user cannot be expected to be able to enter complex technical parameters or make decisions on selection of the mathematical techniques used. Consequently, B-Course has no user definable technical parameters; all the data pre-processing (discretization, missing data handling etc.) and search related decisions (search criteria, search bias etc.) are handled automatically.

Ease of access: There are no problems of installation to various environments, as Application Service Provider (ASP) allows a thin client at the user end for “non-power” users, and the computational load for searching models can be allocated to a server farm. B-Course can be used with most web-browsers and their early versions (even Lynx), and only requires the user data to be presented in tabular text format.

One resource — many trails: The B-Course is arranged around the notion of “trails”; it currently supports the “Dependency trail” and “Classification trail”. Also the ready-made examples are arranged in “trails” in order to simplify things.

Exporting results: In many cases a resource such as B-Course will be used for coursework, demonstrations or scientific work. For such purposes it is important that the results of the analysis can be easily exported in order to be used in reports, term papers and so on. This does not only mean that systems like B-Course have to be able to store the results using some standard formats, it also forces the software to include additional features such as pretty-printing and sometimes verbose explanations about the results.

Interactivity: B-Course allows the user to study the inferred model interactively by providing an inference engine as an “Amazing Bayes-browser” applet. Implementation of the inference engine in B-Course has been the most time-consuming and error-prone interface task in the whole design. However, offering both model construction and inference with the model in the same service is quite essential for making the learning, teaching and research use easier. Integrated environment such as B-Course is not only simpler to use, it also eliminates many of the errors caused by switching between several tools.

B-Course allows researcher to model his/her own data with two different techniques, Bayesian dependency and classification modeling. The tabular
delimited data might come from a Likert-scale questionnaire (ordinal discrete measurement level) or it might be nominal by nature (for example, list of all kinds of attributes that people relate to the concept of growth orientation).

4.11 Operationalization of the Fourth and Fifth Research Questions

The *fourth research question* of this dissertation is formulated as follows: Is there a difference between substantive interpretations of the results of Bayesian dependency modeling and linear bivariate correlational analysis with professional growth data that has both linear and non-linear dependencies? Results of both parametric and non-parametric correlations are compared to the BDM.

Further, the *fifth research question* is formulated as follows: Is there a difference between substantive interpretations of the results of Bayesian dependency modeling and linear confirmatory factor analysis with professional growth data that has both linear and non-linear dependencies? Results of BDM are here compared to the results of EFA and CFA.

Table 11 illustrates how the two research questions are related to the statistical techniques and research tasks. Details of the study design are described in section 5.4 and in the original publication (Study IV).
Table 11. Research Questions Four and Five Grouped by Statistical Techniques and Research Tasks

<table>
<thead>
<tr>
<th>Research task</th>
<th>Parametric techniques</th>
<th>Bayesian techniques</th>
<th>Non-parametric techniques</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I Strength of dependencies between variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>II Significance of group differences</td>
<td>Independent-samples t-test</td>
<td>BDM</td>
<td>Mann-Whitney U-test</td>
</tr>
<tr>
<td></td>
<td>Paired-samples t-test</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>One-way between-groups ANOVA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Two-way repeated-measures ANOVA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ANCOVA, MANOVA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>III Prediction of group membership</td>
<td>LDA</td>
<td>BCM</td>
<td>Logit form of MFA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IV Modeling of both observed and latent variable structures</td>
<td>EFA, PCA</td>
<td>BDM</td>
<td>Categorical EFA</td>
</tr>
<tr>
<td></td>
<td>CFA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BUMV</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Study IV = Nokelainen, Silander, Ruohotie & Tirri, (2007). Investigating the Number of Non-linear and Multi-modal Relationships Between Observed Variables Measuring Growth-oriented Atmosphere. RQ 4 = Is there a difference between substantive interpretations of the results of Bayesian dependency modeling and linear bivariate correlational analysis with professional growth data that has both linear and non-linear dependencies? RQ 5 = Is there a difference between substantive interpretations of the results of Bayesian dependency modeling and linear confirmatory factor analysis with professional growth data that has both linear and non-linear dependencies?

ANOVA = Analysis of variance. ANCOVA = Analysis of covariance. BDM = Bayesian dependency modeling. BCM = Bayesian classification modeling. BUMV = Bayesian unsupervised model-based visualization. CFA = Confirmatory factor analysis. EFA = Exploratory factor analysis. LDA = Linear discriminant analysis. MANOVA = Multivariate analysis of variance. MANCOVA = Multivariate analysis of covariance. MDS = Multidimensional scaling. MFA = Multiway frequency analysis. PCA = Principal components analysis.
5 OVERVIEW OF THE ORIGINAL STUDIES

To generalize is to be a finite mind.
(Lewis, 1964, p. 259)

5.1 Study I

5.1.1 Goals
The study examined the factors of growth-oriented atmosphere in a Finnish polytechnic institution of higher education with categorical exploratory factor analysis, multidimensional scaling and Bayesian unsupervised model-based visualization. The following four research questions were considered: 1) Is the thirteen-factor model of the growth-oriented atmosphere relevant to describe organisational growth prerequisites?; 2) Does the empirical sample support the four domain model of growth-oriented atmosphere?; 3) To what extent employees’ position is connected to growth motivation and commitment to the organization?; 4) Is employee's nature of contract connected to growth motivation and commitment to the organization?

5.1.2 Participants and Procedure
The sample included employees that worked in a Finnish polytechnic institution of higher education during the year 2002. The sample size is 87 per cent of the survey population of 512 workers, indicating 13 per cent non-response rate. The target population of Finnish polytechnic institution of higher education personnel in 2002 was 9,661. Target population is the group of units about which information is wanted, and a survey population is the group of units that we are able to survey.

The sample was obtained with a non-probability sampling. Each employee of the organization was personally invited via email to complete an online version of the Growth-oriented Atmosphere Questionnaire. The online questionnaire presented one to five questions at the same page allowing respondents to attach an open comment to each question.
5.1.3 Measures

The Growth-oriented Atmosphere Questionnaire (GOAQ) contains 67 items describing the growth-oriented atmosphere model in thirteen dimensions: 1) Encouraging leadership; 2) Strategic leadership; 3) Know-how rewarding; 4) Know-how developing; 5) Incentive value of the job; 6) Clarity of the job; 7) Valuation of the job; 8) Community spirit; 9) Team spirit; 10) Psychic stress of the job; 11) Build-up of work requirements; 12) Commitment to work and organization; 13) Growth motivation. A demographics sheet was attached to the instrument enquiring respondents’ position in the organization and nature of the contract. The response options varied from 1 (strongly disagree) to 5 (strongly agree). Details of the instrument are described in detail in section 2.4 and in the original publication in section 3.2.

5.1.4 Data analysis

Research questions were addressed with unsupervised multivariate data analysis techniques that allow ordinal indicators. The first research question was investigated with exploratory factor analysis for categorical indicators (CEFA), that is implemented in Mplus (Muthén & Muthén, 2001), and Spearman nonparametric rank-order correlations. The other three research questions were investigated with non-linear visualization techniques. Non-linear visualization techniques were applied as they are capable of analyzing both linear and non-linear dependencies between variables under investigation.

5.1.5 Results

Thirteen-dimension Varimax-rotated solution in the categorical factor analysis was found to be interpretable in terms of meaningful clusters and correspondence to both theoretical and empirical findings of previous research (Ruohotie & Nokelainen, 2000b). Results of two-dimensional scaling showed that the components on the negative end of the first dimension represent operational capacity of the team. Components on the positive end of the first dimension are related to supporting and rewarding management. Second dimension visualized work-related stress; the most components with the most negative coordinates were psychical stress of the job and increase in the demands of the work. Rewarding for know-how, clarity of the job assignments, and encouraging
leadership represented the positive polarity of the second dimension. Research evidence suggests that the psychic stress caused by the work affects increasingly to the build-up of work requirements.

5.1.6 Discussion
Multidimensional scaling and Bayesian unsupervised model-based visualization both provided evidence to conclude that encouraging rather than strategic leadership increases work satisfaction. Results further showed that managers and teachers had the highest growth motivation and the level of commitment towards work. Employees across all job titles in the organization with temporary or part-time contracts, had higher self-reported growth motivation and commitment to work and organization than their established colleagues.

A recent study among the employees of a U.S. restaurant chain showed that conscientiousness was the best predictor of job performance against work experience, psychological atmosphere and work effort (Byrne, Stoner, Thompson & Hochwarter, 2005). Results indicated that being conscientious might not be enough to secure the highest levels of performance unless the individual is concurrently willing to work hard, and is a member of a psychologically secure work setting. In the current study, most of the employees were working with established contracts and the work performance was not measured, but still the results are comparable at least to some extent. Findings of both studies underline the importance of willingness to work hard (i.e., high growth motivation and valuation of the job) and psychologically secure work setting (i.e., low level of psychic stress, strong team and community spirit).

According to Nokelainen and Ruohotie (2003), encouraging leadership could be used as an indicator for empowerment (Walsh, Bartunek & Lacey, 1998; Spreitzer, DeJanasz & Quinn, 1999). In the future, relationship between Commitment to work and Strategic leadership, Incentive value of the job, Valuation of the job, and Psychic stress of the job should be studied more intensively.
5.2 Study II

5.2.1 Goals
The purpose of this study was to explore the attribution styles (i.e., personal explanations for success and failure) of Finnish adolescents and adults \((n = 203)\) with varying levels of mathematical giftedness to discover what attributions contribute to or impede the development of mathematical talent. The research questions were formulated as follows: 1) Are the four dimensions of the SaaS instrument (success due to ability, failure due to lack of ability, success due to effort, failure due to lack of effort) identified in this domain?; 2) What are the best predictors for the level of mathematical giftedness (highly = “Olympians,” moderately = “Prefinalists” and mildly = “Polytechnics”) and gender among the SaaS variables?; 3) Do the attribution styles differ by the level of mathematical giftedness or gender?

5.2.2 Participants and Procedure
Influence of attribution styles on the development of mathematical talent was investigated in three groups of mathematically gifted Finnish adolescents and adults \((n = 203)\). The first group, “Olympians,” consisted of mathematically highly gifted adults who have participated in the international Olympics for mathematics \((n = 77)\). The second group, “Prefinalists”, consisted of secondary school students, who have taken part in national competitions in mathematics \((n = 52)\). The group represents the top level of Finnish 15-year old students that participated in the international PISA 2000 study. The third group, “Polytechnics,” consisted of adolescent students from a technical polytechnic institution of higher education who study mathematics as their major subject \((n = 74)\). In Finland most of the polytechnic institutions of higher education are highly specialized regional institutions training professionals for expert and development careers. This particular institute is the top-rated technically oriented polytechnic institution of higher education in Finland.

The sample was obtained with a non-probability sampling. Each participant in the three target group was personally invited via email, letter or group leader (teacher) to complete a traditional paper and pencil version of the Self-confidence attitude attribute Scales (SaaS) questionnaire (Campbell, 1996a).
5.2.3 Measures
SaaS included 18 items measuring ability and effort attributions, based on Weiner’s properties of attributional thinking (1974; 1980; 1986; 1994; 2000), on four dimensions: 1) Success due to ability; 2) Failure due to lack of ability; 3) Success due to effort; 4) Failure due to lack of effort. The response options varied from 1 (completely disagree) to 5 (completely agree). Details of the instrument are described in detail in section 3.4 and in the original publication in section “Measurement Instrument”.

5.2.4 Data analysis
The data analysis consisted of four stages. In the first stage all the SaaS items were examined to see if they were technically applicable for linear statistic computations based on multivariate normality assumptions, such as exploratory factor analysis (EFA) and multivariate analysis of variance (MANOVA). In the second stage, an exploratory factor analysis was conducted to answer the first research question. In the third stage, Bayesian classification modeling (Silander & Tirri, 1999; 2000) was conducted to answer the second research question. In the fourth stage, MANOVA with the Roy-Bargman stepdown analysis and Bonferroni post hoc test was conducted to answer the third research question.

5.2.5 Results
The first research question was addressed with EFA. The results showed that the four dimensions of the SaaS were present in the empirical domain. A four-factor solution with eight items grouped the variables in all three sub samples and the combined sample as expected. The eight items operationalizing four SaaS factors were as follows: Factor 1, success due to ability (“5. Being smart is more important than working hard”); Factor 2, failure due to a lack of ability (“3. There are some things you can not do no matter how hard you try”, “13. When I did poorly in school it was because I did not have the needed ability”); Factor 3, success due to effort (“9. Self-discipline is the key to school success”, “12. I had to work hard to get good grades”, “17. Hard work is the key to get good grades”); Factor 4, failure due to a lack of effort (“8. My achievement would have been better if I tried harder”, “16. I could have done better in mathematics if I have tried harder”). The overall alpha values ranged from .62 to .82 with the
combined sample. Although only one item was found to be a valid measure of the first SaaS dimension, success due to ability, correlations between factors behaved as expected: Ability and effort factors correlated negatively with each other and both effort factors, as well as both ability factors, correlated positively.

The second research question was analyzed using Bayesian classification modeling. The classification variables were the level of mathematical giftedness and gender. Eighteen SaaS items were predictors in all the analyses. The results showed that both Polytechnic students and females think that they had to work hard to get good grades. When we further examined the female’s preference for effort as a cause for success, we learned that the result was true only for the Polytechnics and Prefinalists samples as there was no difference between female and male Olympians’ responses. Failure due to lack of ability was the only self-attribute scale that was able to predict respondent’s age. The youngest students (15-28 years) believed more in their abilities than the older ones (29-41 and 42-55 years).

The third research question was analyzed with $3 \times 2$ factorial design MANOVA. Dependent variables were four SaaS factors (success due to ability, failure due to lack of ability, success due to effort, failure due to lack of effort). Independent variables were the level of mathematical giftedness and gender. Results showed that the level of mathematical giftedness multivariate main effect on the SaaS factors was found to be significant. The gender multivariate main effect on the SaaS factors and the level of mathematical giftedness and gender multivariate interaction were not found to be significant. Highly and moderately mathematically gifted individuals felt that ability is more important to success than effort.

5.2.6 Discussion
This study provided understanding how mathematically highly, moderately and mildly gifted adolescents and adults differ in their specific reasons for success and failure. Differences in attribution styles between the three groups of mathematically gifted, measured with the SaaS questionnaire, indicated that it is important to know if the attributions for success or failure are stable or unstable, external or internal. Knowledge of how learners or trainees use attributions to
account for success and failure can help educators and parents to gain a deeper awareness of the mathematically gifted and, thus, predict their expectancies and plan intervention strategies when needed. The information is also applicable to courses concerning the needs of the gifted. Furthermore, the information can be presented directly to mathematically gifted in order to help them develop more insight into their own behavior.

In this study, we measured attribution styles with a questionnaire. Such self-reporting allows us to study, as opposed to attribution appraisals or causal beliefs, hypothetical success and failure situations without clear reference to who is the performer. The first possible source of error is the SaaS instrument translation from English to Finnish. To control the error variance due to translation, all the items were re-translated back to English and compared with the original items. However, no pilot study with correlational analyses was conducted. A second possible source of error is cross-cultural differences between the U.S. and Finnish mathematicians since the original instrument was developed for studies among the U.S. mathematics Olympians. Fortunately, the language of the SaaS items is free of cultural references. A third possible source of error is the psychometric properties of the SaaS instrument itself as no alpha values were reported in the original study (Campbell, 1996a).

5.3 Study III

5.3.1 Goals
The Abilities for Professional Learning Questionnaire (APLQ) is a self-report instrument designed to measure both of motivation and learning strategies. APLQ is mainly targeted for the needs of a higher level vocational education research, for example, to measure motivational level of professionally oriented university and polytechnic institution of higher education students. This study concentrated on the motivational part of the APLQ with the following three research goals: 1) Investigation of the psychometric properties of Abilities for Professional Learning Questionnaire (APLQ); 2) Optimization of the number of items from the original 28 to a lower number to allow the use of APLQ in limited time tasks; 3) Investigation of the relationships between the empirically
derived motivational dimensions and their relation to the original theoretical framework.

5.3.2 Participants and Procedure

The sample included 459 second (n = 281, 61%) and third (n = 164, 35%) year students of a Finnish polytechnic institution of higher education. The sample represented 32 per cent of the total population (n = 1436) and consisted of 323 (70%) female and 134 (29%) male participants. Most of the students (n = 443, 96%) were adolescents, only ten respondents (2%) studied in the adult degree program.

The sample was obtained with a non-probability sampling. Each participant was personally invited via email and group leader (teacher) to complete a traditional paper and pencil version of the APLQ.

5.3.3 Measures

Polytechnic institution of higher education students responded to 28 APLQ items measuring learning experiences and motivation on six dimensions: 1) Intrinsic goal orientation; 2) Extrinsic goal orientation; 3) Meaningfulness of study; 4) Control beliefs; 5) Self-efficacy; 6) Test anxiety. The response options varied from 1 (completely disagree) to 5 (completely agree). Students also completed a short background information questionnaire containing following sections: age, gender, starting year of studies, degree program, and degree mode (adolescent or adult). Details of the instrument are described in detail in section 3.9 and in the original publication in section “Measurement Instrument”.

5.3.4 Data analysis

The statistical procedures were conducted in five stages: 1) Variable selection based on descriptive statistics; 2) Analysis of communalities; 3) Bayesian dependency modeling (BDM); 4) Exploratory factor analysis (EFA) with construct validity test (Cronbach’s Alpha) and generalizability test (confirmatory factor analysis, CFA); 5) Bayesian unsupervised model-based visualization (BUMV). First, all the twenty-eight items were analyzed to see if they were technically applicable for linear statistic computations, such as exploratory and confirmatory factor analysis, based on multivariate normality assumption.
Second, the data was investigated with EFA to see if the six dimensions (i.e., motivational factors) were identified in this domain. In the third stage, probabilistic dependences between all variables were examined with Bayesian search algorithm in order to find a model with the highest probability. The fourth stage was to conduct the traditional exploratory factor analysis with construct validity test (Alpha) and generalizability test (CFA). The task for the last analysis technique, Bayesian unsupervised model-based visualization, was to provide information how the derived factors are interrelated in individual level.

5.3.5 Results
To address the first research goal, items psychometric properties were examined against the following criteria: 1) Standard deviation maximum half the mean; 2) Skewness less than +/- .3; 3) correlation between +/- .3-.7. After comparing the results of the rejection criteria, four variables out of 28 were omitted from further linear analysis. In the second stage, communalities were examined in the original data \((n = 459)\) and in three sub samples \((n = 200, n = 150, n = 100)\) derived from it. Communalities in the comparison data \((n = 512)\) were also inspected. The lower bound of initial estimates and total variance explained had notable better values as the sample size decreased, but the upper bound of initial estimates remained at the same level. The results suggested the use of 21-item solution as the basis for further analysis.

To address the second research goal, the third stage of the analysis involved building of a Bayesian network out of the 28-item solution measuring self-evaluated motivation. The rationale for this procedure was to examine dependencies between variables by both their visual representation and the probability ratio of each dependency. BDM identified all six motivational factors, as expected. The variable rejection based on visual examination of the network model was in parallel with the results obtained from the first and second phases of the analysis. In the fourth stage, EFA was conducted with Maximum Likelihood (due to adequate sample size and the first phase variable selection) and Direct Oblimin rotation (motivational factors were allowed to correlate with each other). The results of factor analysis and Bayesian dependency modeling showed that the six factor 21-item model has similar descriptive power than the
original 28-item version. In addition, CFA was conducted in order to compare the original 28-item solution to the two optimized (24 and 21 item) solutions. The results of model testing using $\chi^2$, information criteria, root mean square error of approximation measures, and residual measures showed that the optimized 21-item solution was satisfactory when compared to baseline model, although the optimized 24-item solution was slightly better than 21-item solution when compared by residual-based fit indices.

The third research goal was addressed in the fifth stage of the analysis by examining the dispersion of single data vectors ($n = 469$) in three-dimensional space. The data was mapped into six dimensions according to the 21 item optimized solution, from which the Bayesian algorithm produced one optimal model. The three-dimensional model was then plotted into series of two-dimensional figures each presenting one motivational dimension. Results showed that 1) Students who had the weakest intrinsic goal orientation were most likely to suffer from test anxiety; 2) Intrinsically motivated students found their studies more meaningful than their extrinsically motivated peers; 3) Students who had both low intrinsic and extrinsic goal orientation, were according to the model more prone to accuse lack of effort or ability if they would fail on learning task. These BUMV results seemed plausible against the theoretical assumptions.

5.3.6 Discussion
Development process of the 21-item solution has been carried out since publication of this study and it has resulted a shorter version of the APLQ instrument, namely Abilities for Computer Assisted Learning Questionnaire (ACALQ, see Nokelainen & Ruohotie, 2005). A technically interesting point in this study was to investigate the polarity change in the test anxiety scale with BUMV. With this technique, it was easy to profile each respondent, or group, by its motivational level on each of the six dimensions. In addition, profiles of two polar sub samples were examined in detail with the same technique. The first sub sample represented students with high, and the second sub sample with low motivational level. Conclusion of the visual inspection of the profiles was that the students with low motivation were likely to accuse lack of effort or ability if they would fail in learning tasks, as opposite to highly motivated students who
self-reported to believe that they would succeed in their studies due to high level of ability and effort.

The problem with extensively used self-report questionnaires is that they do not provide insight into students’ personal motivation traits in actual classroom learning situations as they are mostly developed for measuring traits of domain-specific motivation (Boekaerts, 2001). However, there is a shift in the way the research methodology is applied in the context of motivational research, as different forms of distance learning are gaining more popularity. In order to study complex phenomenon such as student motivation, both multi-faceted theoretical and methodological approach is required. Unfortunately, severe limitations for data collection exist in the context of virtual university learning. Teacher or tutor has few face-to-face discussions with the students. In fact, as an expert in certain area, she has to deal each time with a new group of attendees’. Furthermore, study processes including collaborative actions, such as group works are conducted through the distance learning platforms. As the volume in virtual studies increases, the personal face to face tutoring time for each student decreases and leads to situation where the guidance is conducted via synchronous or asynchronous tools. In this situation, any teacher or tutor would benefit of seeing the personal profile of his/her students and, thus, determining the correct pedagogical approach for each student (Niemi, 2002; Nokelainen, Ruohotie, Miettinen, Kurhila & Tirri, 2003).

5.4 Study IV

5.4.1 Goals

The major goal of this paper was to investigate the number of non-linear and multi-modal relationships in real-life organizational research data sets. The following five research questions were formulated: 1) What kind of and how many non-linearities are captured by discrete Bayesian networks?; 2) Is there a difference between the results of linear bivariate correlations and Bayesian dependency modeling?; 3) Does an empirical sample containing pure linear dependencies have better overall fit indices in CFA than a sample containing less linear dependencies?; 4) Does an empirical sample containing pure linear dependencies have higher CFA parameter estimates than a sample containing
less linear dependencies?; 5) Is there a difference between the substantive interpretations of the results of CFA and BDM with linear and non-linear samples?

5.4.2 Participants and Procedure
The sample \( (n = 726) \) consisted of adult employees from three Finnish universities of applied sciences (D21, \( n = 447 \); D22, \( n = 71 \); D23, \( n = 208 \)). Respondents had three different kinds of job profiles (with 4% missing data, \( n = 31 \)): Managers (6%, \( n = 46 \)), teachers (61%, \( n = 462 \)), and administrative personnel (29%, \( n = 223 \)). The nature of a respondent’s contract was categorized into three classes (with 3%, \( n = 26 \) missing data): Established (70%, \( n = 533 \)), temporary (22%, \( n = 169 \)), and part-time (5%, \( n = 34 \)) employees.

The non-probability sample was collected during the year 2001. Each employee of the organization was personally invited via email to complete an online version of the GOAQ. The online questionnaire presented one to five questions at the same page allowing respondents to attach an open comment to each question.

5.4.3 Measures
The measurement instrument applied in this study was the same 67-item web-based self-report questionnaire (GOAQ) than in the first original publication (see sections 5.1.3 and 2.4). The online instrument had a five-point Likert-scale ranging from 1 (strongly disagree) to 5 (strongly agree).

5.4.4 Data analysis
To measure non-linear dependencies captured by Bayesian networks, every variable was tested in each network by conditioning it one by one with its immediate neighbors in the network. It was observed whether the modes and means of the conditional distributions were linear and whether the conditional distributions were unimodal. Linearity of modes and means was tested by recording whether the means and modes were increasing or decreasing functions of the conditioning variable. Even clear departures from line-like behavior were accepted as linear as long as the direction of correlation (positive, negative) did not change. Therefore, in these experiments, a ‘linear’ relationship is one that
can be more or less adequately modeled by line describing how the central tendency of the dependent variable varies as a function of the independent variable. In measuring the unimodality of conditional distributions, I judged the dependence to be unimodal if (and only if) none of the conditional distributions \( P(Y|X) \) were clearly multimodal.

The first research question was studied with BDM. The second research question was investigated with linear correlations and BDM. The third and fourth research questions were studied with CFA. The fifth research question was answered by comparing the BDM and CFA results.

5.4.5 Results
Investigation of empirical data (\( n = 726 \)) showed that only 22 per cent of all dependencies between variables were purely linear (linear mode, linear mean, unimodal). Sixteen per cent of dependencies were purely non-linear (non-linear mode, non-linear mean, multimodal). Multimodality was the most common reason for the violation of linearity in both data sets.

Investigations were continued with two sub samples of the vocational high school data, namely D21 (\( n = 447 \)) and D23 (\( n = 208 \)). The D21 sample represents in this study linear empirical data with 23.9 per cent of pure linear and 15.0 per cent of pure non-linear dependencies and the D23 sample represents non-linear data with only 16.2 per cent of pure linear dependencies and 18.3 per cent of pure non-linear dependencies.

The subject domain interpretations of linear correlational analysis and non-linear Bayesian dependency modeling (BDM) were compared to learn if the results would lead to different subjective interpretations. The results showed that in general Bayesian network models were congruent with the correlation matrixes as both techniques found the same variables independent of all the other variables. However, BDM found with both linear and non-linear samples a greater number of strong dependencies between the GOA factors. Comparison of the correlations and dependencies in Bayesian networks showed, that in both samples linear correlations indicated a direct connection between know-how rewarding, know-how developing and valuation of the job, whereas Bayesian
models indicated indirect connections between the variables with encouraging leadership acting as a mediator between them.

Further, we focused on the following four aspects of the GOA theory because our motivation was to investigate if there were differences between the results of linear (CFA) and non-linear (BDM) analysis with the linear and non-linear samples: 1) Support and rewards from the management have a positive influence on how an employee experiences his/her know-how as being rewarded and developed and his/her job as being valued; 2) The incentive value of the job has a positive influence on the development of know-how developing and valuation of the job; 3) The operational capacity of the team and team spirit correlate positively; and 4) Work-related psychic stress hinders the development of the GOA. Results showed that both Bayesian dependency models do not support the second theoretical assumption about the relationship between Incentive value of the job (INV) and Know-how developing (DEV) and Valuation of the job (VAL). However, INV and VAL have a statistical dependence also in BDM derived from the linear sample. The second observation is that the fourth theoretical assumption about the negative influence of Psychic stress (PSY) on all the other factors is only partially supported in both Bayesian models.

5.4.6 Discussion
The results showed that the analysis techniques produced similar results for two out of four theoretical aspects, namely the first and third. Different results leading to different substantive interpretations were considered for the second and fourth theoretical aspect as follows. The BDM was able to find only partial support from the linear sample, and no support at all from the non-linear, sample for the assumption that the incentive value of the job would have a positive influence on the development of know-how developing and valuation of the job. Both Bayesian dependency models suggested that the components under investigation are not directly related, but instead indirectly connected to each other via encouraging leadership. The fourth theoretical aspect was supported in both linear analyses, as the Psychic stress of the job factor was negatively related all the other factors. However, in both Bayesian dependency models the PSY
factor was related only to commitment to work and organization. Finally, linear techniques (i.e., bivariate $r$ and CFA) found stronger statistical relationships between factors measuring the Growth-oriented Atmosphere from a linear than non-linear sample. The weak point of this design is obviously the fact that to answer the fourth research question, we need to compare the results of two philosophically (and technically) different statistical techniques, namely observed variable (BDM) and latent variable (CFA) analysis as the former technique analyzes inter-item dependencies without latent trait assumption. However, the experience with various data sets has shown that when the operationalization of the theoretical model is well implemented, the results of the two models are comparable, as the items operationalizing the same theoretical dimension tend to cluster together (i.e., load to the same factor) with both statistical approaches.
6 DISCUSSION

Me: The model in my thesis consists of just two equations: \( y = bx + e_1 \) and \( z = cy + e_2 \) with \( e_1 \) and \( e_2 \) possibly correlated, but neither is correlated with \( x \). Target of my thesis was to estimate parameter \( c \). I have estimated \( c \) satisfactorily to be \( c = 0.78 \) (±0.05) using the best methods found in the SEM literature, and I have given a causal interpretation to my finding.

Dr. Ex: What do you mean by "\( c \) has a causal interpretation"?

Me: I mean that a unit change in \( y \) will bring about a \( c \) units change in \( E(Z) \).

Dr. Ex: The words "change" and "bring about" sound jargon to me, let’s be scientific. Do you mean \( E(Z|y) = cy + a \)? I can understand this last expression, because the conditional expectation of \( Z \) given \( y \), \( E(Z|y) \), is well defined mathematically, and I know how to measure it. But "change" and "bring about" sound wishy washy.

Me: I actually mean "change", not "conditional expectation". By "change" I mean the following: If we have the physical means of fixing \( y \) at some other constant \( y_1 \), and of changing the constant from \( y_1 \) to \( y_2 \), then the observed change in \( E(Z) \) would be \( c(y_2-y_1) \).

Dr. Ex: Well, well, aren’t we getting a bit metaphysical here? I never heard about "fixing" in my stat classes.

Me: Oh, sorry, I did not realize you have statistics background. In this case, let me rephrase my interpretation a bit, to read as follows: If we have the means of conducting a controlled randomized experiment, with \( y \) randomized, then if we set the control group to \( y_1 \) and the experimental group to \( y_2 \), the observed difference in \( E(Z) \) would be \( E(Z_2)-E(Z_1) = c(y_2-y_1) \) regardless of what values \( y_1 \) and \( y_2 \) we choose.

(Pearl, 2000a)

6.1 Summary of the Research

This article dissertation presented a non-linear statistical, namely Bayesian, approach to the modeling of organizational level differences in professional growth and individual level differences in professional learning with empirical samples.

The major goal of the study was two-fold: First, to contribute to the basic research on professional growth and learning, and second, to contribute to the development and use of quantitative research methodology in the aforementioned research areas.

The research goal was addressed with four empirical studies conducted between 2002 and 2007. Two of the studies addressed the issues of professional growth (organizational level modeling of growth-oriented atmosphere), and the other two, issues of professional learning (individual level modeling of learning motivation and self-attributions). Further, all the four original studies shared
research methodological point of interest by applying both linear and non-linear statistical techniques. To summarize, all the original studies had four macro level common components: First, they all involved modeling of professional growth in an organizational level, or professional learning in an individual level. Second, all the articles draw their conclusions from both theoretical and empirical quantitative evidence. Third, in all the studies a sample of staff or students of university of applied science was present emphasizing interest in the modeling of professional growth or learning. Fourth, a non-linear analysis technique, such as Bayesian Classification Modeling (BCM), Bayesian Dependency Modeling (BDM) or Bayesian Unsupervised Model-based Visualization (BUMV) was applied in all the studies.

Following three research questions concerning professional growth and learning were formulated:

**RQ 1: What is the optimal number of dimensions and items in the Growth-oriented Atmosphere Questionnaire (GOAQ) to describe the theoretical model of growth-oriented atmosphere?** Study I investigated with an empirical sample of Finnish polytechnic institution of higher education personnel the GOA dimensions operationalized by the GOAQ with traditional non-parametric and Bayesian statistical techniques. Results showed that the theoretical four group classification of the growth-oriented atmosphere factors was supported by the empirical evidence: 1) Support and rewards from the management; 2) the incentive value of the job; 3) the operational capacity of the team; 4) Work related stress. Further, the results of categorical factor analysis showed that the 67-item and thirteen-factor solution was the most interpretable in terms of correspondence to the theoretical GOA model. Results of the fourth study were in parallel with the findings of the first study by showing that the most problematic dimensions in terms of low correlations with the other dimensions were “Build-up of work requirements (BUI)” and “Growth motivation (GRM)”. One reason for this result is that the variance of both aforementioned dimensions is to some extent explained by other dimensions: “Psychic stress of the job (PSY)” explains 17 per cent of the variance of BUI, and dimensions like “Encouraging leadership”, “Know-how developing”, “Incentive value of the job”
and “Community spirit” explain 49 per cent of the variance of GRM (measured with multinomial logistic regression).

**RQ 2:** Are the theoretical dimensions of the Self-confidence attitude attribute Scales (SaaS) questionnaire identified in the domain of three groups of mathematically gifted participants: Academic mathematics Olympians, polytechnic institute of higher education students, and elementary school students who have participated in mathematical competitions? Study II investigated the attribution styles of Finnish adolescents and adults with varying levels of mathematical giftedness to discover what attributions contribute to or impede the development of mathematical talent. Results showed that the theoretical four group classification of the self-confidence attitude attribute factors was supported by the empirical evidence: 1) Success due to ability; 2) Failure due to a lack of ability; 3) Success due to effort; 4) Failure due to a lack of effort. Further, the results of exploratory factor analysis showed that the eight item and four factor solution was the most interpretable in terms of the attribution theory.

**RQ 3:** What is the optimal number of dimensions and items in the Abilities for Professional Learning Questionnaire (APLQ) to describe the theoretical model of learning experiences and motivation? Study III investigated with an empirical sample of Finnish polytechnic institute of higher education students the psychometric properties of the APLQ, reduced the number of items in the questionnaire from 28 to 21, and analyzed the relationships between the theoretical and data-based models. Results showed that the theoretical six group classification of the motivational factors was supported by the empirical evidence: 1) Intrinsic goal orientation; 2) Extrinsic goal orientation; 3) Meaningfulness of study; 4) Control beliefs; 5) Self-efficacy; 6) Test anxiety. Further, the results of confirmatory factor analysis showed that the 21-item solution was the most interpretable in terms of correspondence to the baseline model.

Following two research questions concerning Bayesian modeling were formulated: **RQ 4:** Is there a difference between substantive interpretations of the results of Bayesian dependency modeling and linear bivariate correlational
analysis with professional growth data that has both linear and non-linear dependencies?

RQ 5: Is there a difference between substantive interpretations of the results of Bayesian dependency modeling and linear confirmatory factor analysis with professional growth data that has both linear and non-linear dependencies?

Study IV investigated the statistical dependencies between observed and latent variables in various professional growth data sets.

The results regarding the fourth research question showed that in general Bayesian network models were congruent with the correlation matrixes as both techniques found the same variables independent of all the other variables. However, non-linear modeling found with both linear and non-linear samples a greater number of strong dependencies between the GOA factors. Comparison of the correlations and dependencies in Bayesian networks showed, that in both samples linear correlations indicated a direct connection between know-how rewarding, know-how developing and valuation of the job, as Bayesian models indicated indirect connections between the variables with encouraging leadership acting as a mediator between them.

Results regarding the fifth research question showed that the analysis techniques produced similar results for two out of four theoretical aspects, namely the first and third. Different results leading to different substantive interpretations were considered for the second and fourth theoretical aspect as follows. The BDM was able to find only partial support from the linear sample, and no support at all from the non-linear, sample for the assumption that the incentive value of the job would have a positive influence on the development of know-how developing and valuation of the job. Both Bayesian dependency models suggested that the components under investigation are not directly related, but instead indirectly connected to each other via encouraging leadership. The fourth theoretical aspect was supported in both linear analyses, as the Psychic stress of the job factor was negatively related all the other factors. However, in both Bayesian dependency models the PSY factor was related only to commitment to work and organization.

The design of the study to address the five aforementioned research questions is summarized in Figure 10.
Figure 10. Operationalization of the Five Research Questions within Theoretical Models of Professional Growth and Learning
6.2 Evaluation of the Research

As I noted in the previous paragraph (and in the third original publication), there is always a willing and able discussion going on about the use of self-report measures in scientific studies. Zumbo and Rupp (2004, p. 73) state that “... a useful and essential tool such as an automobile, a chainsaw, or a statistical model can be very dangerous if put into the hands of people who do not have sufficient training and handling experience or lack the willingness to be responsible users.” To dig this issue a bit deeper, I will first address how self-report measures challenge reliability and then proceed to discuss their threat towards the validity of any decent scientific work.

According to Martin (2004, p. 351), reliability of scientific study is “the degree to which a measurement can be successfully repeated.” Repeating a measurement can mean variety of components, for example, the same instrument applied on different occasions, two halves of the same instrument, or two scorers using the same observation schedule (Black, 1993). Thus, a measurement instrument is perfectly reliable if we get exactly the same result when we repeat the measurement a number of times under comparable conditions (Martin, 2004). de Vaus (2004) lists numerous sources for unreliability, for example: 1) Poor question wording; 2) Mismatch of the target and sample population; 3) Age, gender, class or ethnicity related mismatch of interviewers and interviewee.

Discussion about the reliability of any self-report instrument is easiest to start with Campbell’s famous statement (1982, p. 692): “One possible exception [to exercise a professional bias] pertains to the use of a self-report questionnaire to measure all the variables in a study.” According to Crampton and Wagner (1994), various studies have been conducted to test the hypothesis that self-report questionnaires, if used as the only data collection methods, artificially elevate measures of covariation, producing so called percept-percept inflation in published correlations. However, when they conducted a large scale meta-analytic research involving 42,934 correlations published in 581 scientific articles, findings challenged the validity of general condemnations of self-report methods showing that percept-percept inflation did not had the broad, comprehensive effects envisioned by so many critics. Perhaps the best
conclusion at this point of discussion is that it does only good for the research validity to underestimate the power of the results produced by the self-report tests.

I tend to link this ‘usefulness of self-report tests’ issue in my mind to the findings of the Johnson and Creech (1983) study, which I reported earlier in the fourth original publication. They announced that when continuous variables are measured by ordinal indicators, researcher’s substantive interpretations might alter with a small sample of 500 cases. As we clearly see, ‘a small sample’ is a flexible, and thus, domain-specific concept!

Two concepts of validity, namely external and internal validity, are fundamental to developing research designs and, thus, theoretically interesting and coherent research results. Internal validity refers to the fact that we need to be confident that the research design can sustain the conclusions that we claim for it, and external validity refers to the extent to which the results from a study can be generalized beyond the particular study (de Vaus, 2004).

After a rigorous meta-analysis of SLR theories, Puustinen and Pulkkinen (2001, p. 283) wished that “in the future we move towards a more integrated conception of SRL, supported by a solid empirical approach.” They are concerned about the use of self-report measures instead of more naturalistic and empirically methods, which they hope to produce a more dynamic and diversified appreciation of the nature of SRL as a phenomenon. I agree that their concern for the decreased level of internal validity is real and directly applicable to my work as I report the findings mostly based on various professional growth and learning self-report measures. “However, all the abovementioned instruments are based on solid theory and research (observation, interview), which I conclude, at least partly, to satisfy Puustinen and Pulkkinen’s “naturalistic and empirically valid methods” cited earlier. Further, all the measures presented here (Growth-oriented Atmosphere Questionnaire, Self-confidence attitude attribute Scales and Abilities for Professional Learning Questionnaire) have lived through numerous reported development phases (e.g., Kautto-Koivula, 1993; Ruohotie, Nokelainen & Tirri, 1999; Ruohotie & Nokelainen, 2000a; Nokelainen & Ruohotie, 2008). All the instruments are also constantly under development and both validity of their theoretical models and
the results are compared to the findings of empirical qualitative research, most recently in the works of Antikainen (2005) and Nokelainen et al. (2007).

If we carry this methodological debate a bit further, we should remark that the recent research has inevitably shown (e.g., Fleeson, 2007; Zautra, Affleck, Davis, Tennen & Fasman, 2007) the benefits of studying both between-person (or across-person) and within-person associations. I admit that I have applied the three measurement instruments in the four original studies to draw within-person inferences from between-person associations. For example, in the second original publication, the (between-person) results showed that highly and moderately mathematically gifted participants reported that ability was more important for success than effort, and mildly mathematically gifted tended to see effort as leading to success, $F(2, 194) = 3.28, p = .040, \eta^2 = .03$. By adding a within-person dimension to the analysis with the Experience-sampling methodology (EMS, e.g., Bolger, Davis & Rafaeli, 2003), I would have been able to enhance the relevance of the analysis, for example, by calculating within-person mean, and further, deviance from the mean of self-rated attributions. Fleeson (2007) lists four main benefits of EMS: 1) Reduction of memory biases or reconstructive processes; 2) Collection of a large amount of information about each individual; 3) Enabling both between-person and within-person investigation; 4) Possibility to control participants’ personalized psychological functions. However, I have personally faced some difficulties of conducting EMS in the research field of professional growth and learning. The first impediment from my point of view is money, as each participant needs a personal data assistant (purchase price) with customized software (development cost). The second issue is the need for additional resources, as at least some of the participants need a personal help desk service. The second issue leads to the third problem of using EMS: How to ensure high data quality with repeated measures? Some participants will experience technical problems or fear of using PDA’s, some will get tired of filling the same data collection sheet. My last critical notion of using this approach is related to the study design and analysis: EMS generates a huge mass of data per participant. I have personally found in the research field of educational technology (e.g., Miettinen, Kurhila, Nokelainen & Tirri, 2006; Kurhila, Miettinen, Nokelainen & Tirri, 2007;
Nokelainen, Miettinen, Kurhila, Floréen & Tirri, 2005) that it is difficult to formulate small enough but still meaningful hypotheses and to keep the data matrix and analyses manageable.

Discussion about the issue of external validity is also important as all the participants in this dissertation share the same nationality, namely Finnish. The critical issue here is, according to de Vaus (2004), whether the results are likely to apply more widely. I admit the existence of cultural differences to some extent, as two out of three questionnaires (GOAQ and APLQ) applied in the four original publications are specially designed for Finnish respondents. However, the theoretical concepts that those measurement instruments operationalize, are expected to be culturally independent as they are based on large international research body. Perhaps the most culturally biased instrument presented here is the GOAQ, as instead of measuring a personal quality, it measures subordinates perceptions of their leader’s and co-workers characteristics – which are under strong socio-cultural influence.

6.2.1 Is There Magic in the Air?

Traditional concepts of reliability and validity, developed originally by Lee Cronbach (1916-2001) are useful, but sometimes I feel more tempted, referring to Robert Abelson’s (1995) more recent framework, to ask myself “Do I have the magic in my work?” The ‘magic’ acronym stands for five properties of scientific work: 1) Magnitude; 2) Articulation; 3) Generality; 4) Interestingness; 5) Credibility.

The first property, magnitude, stands for quantitative magnitude of support for its qualitative claim and is usually measured with an index of magnitude such as the “effect size” (Cohen, 1988) in experimental designs or the “model fit indices” (e.g., Kaplan, 2000) in structural equation model validating. The point here is that the significance level (so called “p-value”) should never be used as the only indicator of the merit of the outcome as it depends not only on the degree of departure from the null hypothesis, but also on the sample size. I have calculated indexes of magnitude where applicable; the effect sizes in the second original study and model comparison indices in the third and fourth original studies.
As described earlier in section 4.4, power of a statistical test is the probability that it will lead researchers to reject the H\(_0\) when that hypothesis is false. Murphy and Myors (1998, p. 17) recommend that power should be above .50, otherwise the study is unlikely to reject the H\(_0\). They have generated a model of translating common statistics (i.e., t, r\(^2\), \(\chi^2\)) into F-equivalent values, and further specified ‘a one stop F table’ for testing significance and estimating power for null hypothesis tests. The table also allows negligible effect testing of significance and power (i.e., treatments account for 1% / 5% or less of the variance). For example, I have communicated three paragraphs earlier that for highly and moderately mathematically gifted participants, ability was more important for success than effort, and for mildly mathematically gifted the effort was more important for success than ability. The proportion of variance (PV) index (id., p. 31) for this finding is about three per cent:

\[
P V = \frac{v1F}{v1F + v2}
\]

\[
= \frac{2 \times 3.28}{2 \times 3.28 + 194}
\]

\[
= .032
\]

The ‘v1’ in Equation 1 stands for the number of compared groups minus one, ‘v2’ stands for \(N - v1 - 1\), and ‘\(F\)’ is the F statistics value. The result shows that the \(PV\) obtained is equal to the earlier reported eta squared value (\(\eta^2 = .03\)). By applying the statistical indices of MANOVA into the ‘one stop F table’, we learn that the \(F\) value of 3.28 would allow the traditional null hypothesis to be rejected with \(\alpha = .05\) as the tabled \(F\) value is 3.04, but not with \(\alpha = .01\) as the \(F\) value is 4.71. The table further shows that an equivalent \(F\) value for the power level of .50 is 2.48, respectively, and the \(F\) value for the power level of .80 is 4.86, respectively. If the Equation 1 is used to transform these values into \(PV\) values, the values of .03 and .05, respectively are found. As our original \(F\) value of 3.28 falls between the two boundary values, we need to conduct an interpolation between the tabled values in order to estimate where in this range the power of the study falls (id., p. 47):
The ‘$F_{\text{hypothesized}}$’ in Equation 2 stands for the hypothesized size of the effect ($F$ value of 3.28 when $PV=.03$), ‘$F_{.50}$’ stands for the $F$ equivalent of the $PV$ needed to obtain power of .50 ($F$ value of 2.48), and ‘$F_{.80}$’ stands for the $F$ equivalent of the $PV$ needed to obtain power of .80 ($F$ value of 4.86). The power to reject $H_0$ when it is false is estimated to be satisfactory (.60). All the statistically significant MANOVA results presented in the second original publication exceed this power level, as this example represented the weakest $PV$ in the study.

The second ‘magic’ property, articulation, stands for “the degree of comprehensible detail in which conclusions are phrased” (Abelson, 1995, p. 12). I have tried to improve the articulation in the original research publications by writing as explicit “Conclusions”, “Discussion” and “Practical Implications” sections as possible. I wish that the development of my scientific articulation skills is evident when comparing the oldest (Nokelainen & Ruohotie, 2002) publication with the other three more recent original publications (Nokelainen & Ruohotie, 2008; Nokelainen, Silander, Ruohotie & Tirri, 2007; Nokelainen, Tirri & Merenti-Välimäki, 2007).

The third property of the ‘magic’ criteria is named generality, denoting the applicability of the conclusions of the study and is quite synonymous to external validity. However, Abelson (1995) stresses that before publication of the results, quantitative researcher should carry out one or more exact replications of the original study to reproduce the initial results. I have approached this problem in my research work over the years by collecting more than one sample (applied in all original studies) and/or replicating the analyses with randomized sub samples when the original sample is large enough (original studies III and IV). The second important issue within the generality is the art of explaining treatment-by-context interactions: “You never understand a phenomenon unless you can
make it go away or unless you can reverse its direction” (Abelson, 1995, p. 143, a modification of Paul Lazarsfeld’s original oral quotation). This aspect reveals, once again, an irritating weak point in all my original publications: they are based on cross-sectional correlational designs. However, I try to control treatment-by-context interactions in design level by using quasi-experimental design and in data analysis level by applying genuine multivariate approach (e.g., use of MANOVA instead of ANOVA in the second study) whenever it is possible, although I am aware that they only partially compensate for the use of experimental designs.

The fourth ‘magic’ property, interestingness of scientific work contains two key points: Firstly, change of belief by entailing surprising results, and secondly, importance of the issue. It is clear that both of these features are an essential requirements for any academic dissertation. Next, I will describe in a nutshell what were the key ‘excitement factors’ (i.e., “surprise” and “importance”) of my four selected studies. The first original publication surprised by showing that polytechnic institute of higher education employees with established work contracts were less growth motivated and committed to work and organization than their part time or temporary contracted colleagues. Importance of this study builds on two major facts: First, an empirically tested freely available self-rating instrument is provided, and second, it helps explaining employees’ working motivation and satisfaction towards work and, thus, improvement of working conditions.

The second original publication surprised other scholars by showing that “highly gifted individuals, both males and females valued ability as explanations for both success and failures” (Olszewski-Kubilius, 2007, p. 4). I agree with her and carry on with a notion that results revealed some interesting attributional patters by showing that mildly mathematically gifted individuals were more likely to attribute both success and failure to effort. Olszewski-Kubilius (2007, p. 4) evaluates the importance of this study as follows: “The results of this study add to the complexity of the picture regarding attributional style and suggest that there will be differences within a gifted population on this variable.”

The third original publication surprised with three features: First, by defending the use of self-rating instruments in educational studies with
multifaceted discussion; Second, by showing explicitly the effects of item rejection, based on the assumptions of the traditional frequentistic techniques, to the substantive interpretations of the results (i.e., the questionnaire would had been totally different without investigation of the item’s statistical properties); Third, by introducing application of the Bayesian modeling techniques in an educational science context. The major goal of the whole “Theoretical Understandings for Learning in the Virtual University” book (edited by Niemi and Ruohotie, 2002) was to provide educators and administrators some inside views to an important issue of how to support virtual university students to become active and self-directed learners. The specific importance of my article was to present an empirically tested freely available self-rating instrument for the analysis of virtual university student’s learning motivation. If we do some cross linking between the research results so far, we notice that the second original publication showed mathematically oriented polytechnic institution of higher education students to prefer effort over ability as an explanation for success. However, results of this third study, involving more heterogeneous group of polytechnic institution of higher education students, showed that the highly motivated students, as opposite to low motivated students, preferred ability over effort as an explanation for success. However, this discrepancy in the findings is most obviously due to the fact that the within group dispersion of the attributional values of polytechnic institution of higher education student group was not analyzed and taken into account in the second study.

The fourth original publication surprised by two ways: First, results showed that only 22 per cent of the statistical dependencies in an empirical sample \((n = 726)\) between the GOA items were purely linear by nature (i.e., linear mode and mean, unimodal), and second, the results showed that there is a difference between substantive interpretations of the results of linear and non-linear analysis techniques with linear and non-linear dependency samples. Importance of this study is quite obvious, as there is no statistical sense to look for linear inter-item statistical relationships from a data that contains only 22 per cent of such dependencies. However, this is exactly what the researcher is doing if applying a traditional linear frequentistic statistical procedure, like Pearson product moment correlation analysis, on a data that contains mostly non-linear
dependencies. The outcome of such procedure is the most important feature of the Study IV: Researcher is reporting at least partially false results (two out of four research questions were different depending on the statistical approach) and conclusions in his/her report, if not using the right statistical tools for the data. At this point one question remains: How do I know if my data is linear or non-linear by nature? By comparing all the multivariate dependencies, for example, with scatterplots or with the procedure explained in the Study IV. Unfortunately, visual comparison is extremely tedious and we have not published the implementation of the automatic dependency analyzator. The problem is quite simply answered by using non-linear techniques on all the data as those techniques are capable of analyzing both linear and non-linear dependencies.

The last ‘magic’ property, credibility, refers to the believability of the research claim and requires both methodological soundness and theoretical coherence. Researcher must be able to explain how the data fall in a particular pattern and “to match wits with supercilious competitive colleagues who stubbornly cling to presumably false alternative accounts, based on somewhat different clues” (Abelson, 1995, p. 14). It is clear that the results based on sloppy procedures or mistaken statistical analyses will be an easy target of criticism by other scholars. Next, I will concentrate on the possible weak points of the statistical analyses applied in the four original studies as I have already evaluated the methodological and theoretical aspects of the studies in the earlier sections of this discussion.

Table 1 summarized the Bayesian statistical techniques applied in the four original publications and chapter four discussed about their properties and application. However, a wide range of other, more traditional techniques was also used in the four original publications: inter-item correlations (all four studies), MANOVA (Study II), PCA (Study III), EFA (all studies), internal consistency index (Studies I, II and III), MDS (Study I) and CFA (Studies III and IV). Due to discrete indicators, a categorical variable technique was always applied when available: Spearman rank order $r_S$ instead of Pearson product moment $r_P$; categorical indicator EFA instead of continuous indicator EFA; categorical indicator CFA instead of continuous indicator CFA. However, there is an exception to the general rule of favoring non-parametric techniques:
Parametric MANOVA was applied instead of non-parametric Kruskal-Wallis H test in Study II as the point was to control increasing Type I error by using a genuine multivariate technique (both K-W H and ANOVA allow only one-way analysis).

The aforementioned techniques were applied in all four original studies in the same manner: Correlations were calculated to study if the theoretical and empirical dependencies were directed the same way and to test if the inter-item dependencies were strong enough to allow multivariate analysis; MANOVA to test the difference between several respondent group’s perceptions of a correlative multivariate model; PCA to test the dimensionality of model; EFA to test the presence of theoretically plausible latent traits in the empirical evidence; CFA to test the external validity of the selected model with several data sets (study IV) or to compare competing models by their fit to the data (Study III).

Direct oblimin rotation technique in EFA was applied in most of the studies, as the theoretical dimensions of all three inspected models (organizational level growth-orientations, individual level learning motivation and self-attributions) are allowed to correlate. Further, no model modifications were done during the CFA as it is well known that “… when procedures are used that empirically modify a model to make it look as good as possible in a particular sample, all of the model fit indexes will appear unduly optimistic about the quality of the model” (Hu & Bentler, 1995, p. 99).

What comes to the Bayesian modeling techniques presented in this study (BDM, BCM and BUMV), possibilities for misuse by the researcher are considerably more limited when compared to the traditional statistical techniques. Naturally, both approaches share mostly the same study design and data collection concerns (this topic is discussed in detail in section 4.5), but the Bayesian applications allow very limited amount of data and analysis manipulation possibilities. The most severe limitation of the practical research use of BDM is connected to its publicly available implementation, B-Course. The current version of the software (2.0.0) is capable of finding the most probable model out of quite limited data. Empirical testing has shown that with large data sets (e.g., 18 variables and over 500 observations), mixing of the original variable order produces somewhat different models. However, the basic
structure is present in all model variations. In BCM, researcher is able to force classification analysis without variable selection procedure (resembling the “enter technique” in LDA). In all aforementioned analysis techniques, researcher is also able to decide the discretization points for continuous variables manually. This may lead to biased results if it is conducted arbitrary without considering theoretical assumptions for cut points and/or frequency histograms.

My last point to answer to the credibility part of the magic criteria concerns the missing data issue that is evidently present in all four original studies. The traditional social science treatment for the missing data, strongly influenced by Sir R. A. Fisher, is to have none (McKnight, P., McKnight, K., Sidani & Figueredo, 2007). It is true that large, normally distributed samples allow listwise deletion (to have no missing data at all), but even if the sample size is still adequate, we have no idea if the missing data was patterned or not. However, with ‘normal’ social science sample sizes well below 200 observations (or with longitudinal studies), researcher is certainly doing his/her best to keep all the participants in the analysis to the end.

The missing data is often a serious problem with longitudinal studies that are prone to participant dropouts and with so called ‘mono-operations’ where a single measure represents a concept or construct (McKnight et al., 2007). However, in the present cross-sectional studies I have followed the classical test theory’s rule that multiple measures increase reliability by using more than one indicator for each factor or component. It is clear that this rule provides no help with ‘sensitive’ questions (e.g., illegal substance use), but fortunately this is not case with the present research topics. The GOAQ addresses the most sensitive issues of all three self-report instruments applied in this study (by asking questions about superior’s competence as a leader and respondent’s personal work load), but still produces close to none missing data. The three main reasons for a low missing data frequencies in the original four publications are as follows: 1) Participants are invited, motivated and instructed personally to answer to the questionnaire at hand; 2) Questionnaires are not putting too much burden on participants despite of their use of ‘multiple items’ strategy (GOAQ 67 items, SaaS 18 items, APLQ 28 items); 3) The most sensitive data (i.e.,
organizational growth-oriented atmosphere measured by GOAQ) is collected and reported anonymously.

Scholars in all research fields do their best to prevent the missing data problem in the first place by study design, but when we do have missing data, frequentists and Bayesians see the problem quite differently. Frequentists use a statistical imputation technique (e.g., constant replacement techniques, random value imputation or nonrandom value imputation with single or multiple conditions) to allow application of various sample size sensitive statistical analyses (e.g., EFA and CFA). Bayesians, too, operate with imputation techniques, but usually only with extremely small samples or unequal group sizes.

In the original four publications, ‘median substitution’ imputation technique is applied with frequentistic techniques as it performs well also with other than normal distributions (see McKnight et al., 2007). With Bayesian analysis, a relevant missing data class is established and analyzed. Latter is not a possible procedure with traditional frequency-based statistical techniques if the two common conditions are met: 1) Fixed missing data class value is theoretically unjustifiable (e.g., using “0” or “9” with the response scale from “1” to “5”); 2) Missing data class size (e.g., “NA”) is not comparable to the natural class sizes of the data (producing unequal “NA”, “Male” and “Female” group sizes).

6.3 Recommendations

The results of Study I indicated that a psychologically secure work setting (clarity of the job, encouraging and rewarding leadership) together with willingness to work hard (valuation and incentive value of the job) create a work-place atmosphere that is positively growth-oriented and, thus, productive. This same finding was present also in the fourth study: Encouraging leadership was strongly connected to the job valuation and personal know-how developing. Results from the first study also showed that those workers who had part-time contracts, reported to have higher levels of growth motivation and commitment to work and organization than those who had an established contracts. Before concluding that too secure work ruins organizational growth, we must bear in
mind that the study design was cross-sectional, not longitudinal. In the future, it would be interesting to study more closely the dependency between the nature of work contract and the GOA factors over time. According to the results, work content affects employees’ growth orientation: Managers and teachers were more motivated and committed workers than caretakers, cleaners, librarians, office secretaries and other supportive personnel who are working in the lower salary scale. The salary itself might not be the only explanatory variable, also other factors must be taken into account as viable motivating tools: Improvement of work contents based on professional consulting and discussion between employer and employee, conversion training and job rotation when applicable and staff training days. The whole organization’s atmosphere would benefit if all the employees would not only be aware of their own expertise areas, but would also recognize the strengths and responsibilities of their colleagues as well. Finally, the answer to the first research question of this dissertation, asking about the optimal number of dimensions and items in the GOAQ, is that the thirteen-factor and 67-item model was found to be a valid operationalization of the GOA model in the domain of Finnish polytechnic institute of higher education. The GOAQ is also used by several Finnish business sector consults with useful (unpublished) results, demonstrating the good generalizability of the model into different domains. My recommendation is that as the GOAQ is empirically tested instrument, its use in other relevant domains is justifiable.

The results of Study II showed that it is essential that educators and parents understand the influence of different attribution styles on the development of mathematical talent. It is important to know if the attributions for success or failure are stable or unstable, external or internal, as it allows educators and parents predict their expectancies and plan intervention strategies when needed. The theoretical idea of Ellström (2001) essential to this study is that attributions for success or failure affect potential competence, which is a human resource each individual brings to the mathematical problem-solving situation. So, what are the ‘good’ attributions that give rise to individual’s potential competence? The previous research body shows three trends: The first trend sees ability as a more important explanation for success than effort, ‘ability is everything’, the second trend claims that ‘ability without effort goes nowhere’, and the third,
most current trend (e.g., Schunk & Ertmer, 2000; Zimmermann, 2000) states that ‘ability and effort without self-regulation goes nowhere’. The major difference between the research trends is that the first research trend is speaking only about the real ability level, but according to the concept of self-regulation, the real ability level is undistinguishable from the self-perceived ability. This makes sense if we think that effort, part of “Performance or Volitional Control” phase of the cyclical self-regulation process (Figure 4), is sending adjusting signals via self-observation to individual’s self-perceived ability. Attributing success to effort has a self-enhancing and motivating effect as one feels in control of one’s own development. I agree with Alderman (2004) when she suggests that it is up to the teacher or trainer to convince a learner or trainee that mathematical thinking ability as a skill or knowledge is a learnable, unstable quality. The answer to the second research question of this dissertation asking, if the SaaS dimensions are identified in the domain of three mathematically gifted participants, is that all four attributional dimensions were present in all three samples. I recommend that SaaS, although a valid and cultural reference free implementation of the attribution theory, should be applied with a great caution in other target groups than mathematically gifted persons as its development is not explicitly described in the original manual (Campbell, 1996a) and the psychometrical validation process of the instrument is not conducted in other domains.

Study III presented the learning experiences and motivation section of the APLQ, which is an adaptation of the MSLQ (Pintrich, Smith, Garcia & McKeachie, 1991; 1993). Difference with MSLQ and another popular ARS inventory, Learning and Study Strategies Inventory (LASSI, Weinstein, Palmer & Schulte, 1987), is that the latter measures domain-general academic self-regulation (ASR) while the former is developed for measuring college students’ domain-specific ASR (Pintrich et al., 1991). According to Dugan (2005), this “state-trait” issue is under continuous debate. He found in a comparative study that although contextual and individual variables appear important in the domain-specific (i.e., state) ASR, also a domain-general (i.e., trait) level of ASR aids achievement regardless of the context. APLQ aims to capture both domain-specific and domain-general learning experiences and motivation (reported in
Study III), study habits (reported in Nevgi, 2002), quality of teaching and effects, and outcomes of education of adult learners. Results of the third study showed that all six theoretical level motivational dimensions were also present in the empirical evidence. Further, numerous subsequent studies with both polytechnic and university students (e.g., Nokelainen & Ruohotie, 2004, 2005; Kaartinen, 2005) have shown that both APLQ and its online learning version (ACALQ) work quite robustly even with heterogeneous and small sized samples. As an answer to the third research question in this dissertation, I recommend using APLQ when investigating higher education students’ domain-specific learning experiences and motivation.

The results of Study IV showed that some of the highest bivariate correlations in the samples were explained via a third variable in the non-linear Bayesian dependency modeling (BDM). Further, CFA and BDM led to different substantive interpretations in two out of four research questions concerning organizational growth. My recommendation is that the various univariate and multivariate properties of the variables in the data should be carefully examined before selecting the analytical approach. Traditional frequentistic parametric statistical techniques work most efficiently with samples of certain size and continuous measurement level, and consisting of pure linear dependencies. Naturally, traditional non-parametric techniques work best with linear samples with discrete indicators. However, the unpleasant truth is that the non-parametric techniques are more prone to Type II error than parametric techniques (Johnson, 1995; Stephen, 1995). When the nature of dependencies is unknown, I recommend using Bayesian approach as it is able to analyze all inter-item dependency types and it is not prone to the null hypothesis testing errors due to different statistical paradigm. To make this recommendation more concrete, I will mention here a result of an unpublished study of Nokelainen and Silander where they analyzed the nature of inter-item dependencies with a combined sample of 16,500 participants and found that only 49.5 per cent of the dependencies were purely linear by nature. The ‘problem’ with Bayesian analysis techniques presented here is that they only allow investigation of observed variables. Fortunately, Muthén’s CVM (1983) provides non-linear options for EFA and CFA (implemented in Mplus), which I warmly recommend.
I referred in the very beginning of this dissertation to the old saying that knowledge increases pain. I guess that this statement sounds true to many scholars, as we are painfully aware of the two things: Firstly, more we learn to know, more we learn we should also know (limitations of brain capacity and time), and secondly, more we know, more complex the truth that every scientist pursues, becomes (serious problems of giving a simple ‘yes’ or ‘no’ answer when asked).

From a methodological point of view, an example of such complexity that growing awareness of research design and statistical issues may create, I will present here a simple question that has been presented to me many times: “If I collect my non-probability sample with an ordinal indicator self-rating instrument, aren’t I still allowed to use traditional linear parametric analysis techniques such as ANOVA or EFA?” First answer would be: “Of course you are, what other options would you have?” Second answer soon follows: “Of course you are not, read your books about non-parametric tests!” Third answer is: “You should re-design your study to meet the assumptions of those techniques you mentioned, or if experimental design is out of question, perhaps consider qualitative approach.” Fourth answer is getting a bit longer: “Go ahead and use them: Firstly, they are robust techniques enough to handle your data and still produce viable results, and secondly, journal editors and reviewers expect to see results of parametric tests - and besides, they look suspiciously at the use of marginal techniques, such as non-linear techniques, and probably tell you that you are only using sophisticated techniques to cover problems in the theoretical background and research design.” Was that not quite the same as the first answer? However, it is the fifth answer that finally starts to show, painfully, some maturity in this issue: ”Why would you like to use the particular instrument, have you thought about that? It is OK to use it, if it is a valid instrument and produces data that is useful for your research purposes (by operationalizing a relevant theoretical model). Let me see you research design. I see, you plan to analyze mathematical ability with a self-report instrument on a scale from 1 to 5. How about comparing the results of a self-report questionnaire to a real-life math test score? How about administrating a self-report test before and after a math test? I think you should consider more carefully which are the
best data collection techniques and points for your design and how large sample is needed to obtain a specific effect size and power.” The last answer is the one that is most parallel with the text in section 4.6 and reflects the Figure 8 where I present one possible methodological model of professional growth and learning research design.

During the past years I have conducted (e.g., Antikainen, 2005; Luoma, 2001; Ryhänen, 2006; Susimetsä, 2006) or consulted (e.g., Kivinen, 2003; Korpelainen, 2005) Bayesian statistical analyses for quite a large number of professional growth and learning researchers and both national\(^1\) and international\(^2\) research projects. During all these collaborative efforts, and way too many to mention here, I have learned that all the scholars I have had a pleasure to work with have been extremely satisfied with the real-life interpretability and, thus, usability of the results that the Bayesian techniques produce. One reason for such success is that the results have been easy to transform into meaningful interpretations. This is for me the most promising indicator for the future success of the Bayesian statistical applications in educational research. I believe that we all, who are tackling with an endless number of quantitative human science research problems, do have a Bayesian choice. I do hope that this work opens the door for an educational researcher a bit wider into the world of Bayesian statistical applications. My best reward for doing this work would be to see other researchers giving Bayesian approach a try.

---

\(^1\) For example, “Computational Intelligence Techniques for Non-linear Modeling in Social Sciences” lead by Professor Henry Tirri, University of Helsinki, Finland, http://cosco.hiit.fi/Projects/NONE

\(^2\) For example, “International Study of Academic Olympians”, lead by Professor J. R. Campbell, St. John's University, U.S., http://olympiadprojects.com
REFERENCES


Hofer, B. K., Yu, S. L., & Pintrich, P. R. (1998). Teaching college students to be self-regulated learners. In D. Schunk & B. Zimmerman (Eds.), *Self-


Pearl, J. (2000a). *A dialogue with a hostile examiner, Dr. Ex*. Retrieved November 19, 2000, from SEMNET@BAMA.UA.EDU.


Pintrich, P. R., & Sinatra, G. M. (2003). Future directions for theory and research on intentional conceptual change. In G. M. Sinatra & P. R. Pintrich (Eds.),
Intentional conceptual change (pp. 429-441). Mahwah, NJ: Lawrence Erlbaum Associates.


---

*International Self-Concept Research Conference*. University of Western Sydney, Self Research Center.


Investigating Growth Prerequisites in Organizations: A Case Study in Finnish Polytechnic Institution of Higher Education

Petri Nokelainen and Pekka Ruohotie

University of Tampere

Abstract
This study examines the factors of growth-oriented atmosphere in a Finnish polytechnic institution of higher education (n = 447) with categorical exploratory factor analysis, multidimensional scaling and Bayesian unsupervised model-based visualization. The following four research questions were considered: 1) Is the thirteen-factor model of the growth-oriented atmosphere relevant to describe organisational growth prerequisites?; 2) Does the empirical sample support the four domain model of growth-oriented atmosphere?; 3) To what extent employees’ position is connected to growth motivation and commitment to the organization?; 4) Is employee's nature of contract connected to growth motivation and commitment to the organization? Results showed that a thirteen-factor model of the growth-oriented atmosphere was relevant to describe growth prerequisites of Finnish polytechnic institution of higher education employees. The theoretical four group classification of the growth-oriented atmosphere factors was supported by the results: 1) Support and rewards from the management; 2) Incentive value of the job; 3) Operational capacity of the team; 4) Work related stress. Results further showed that managers and teachers had higher growth motivation and level of commitment to work than other personnel including job titles such as cleaner, caretaker, accountant and computer support. Employees across all job titles in the organization, who have temporary or part-time contracts, had higher self-reported growth motivation and commitment to work and organization than their established colleagues.

Keywords
Organizational growth, growth-oriented atmosphere, polytechnic institution of higher education, survey, categorical variable modeling, Finland
1. Introduction

Each one of us experiences his or her own working conditions in a unique way that directly affects on one’s motivation and satisfaction toward work. Attributes of work are, thus, to be studied in real working context because substance of work and organizational atmosphere are not independent factors. The research paradigm associating studies explaining the growth prerequisites of organizations emphasize individual’s own perceptions and interpretations of ones work and working conditions. In addition to the organizational growth prerequisites, professional updating is dependent on personal growth requisites like the professional self-concept and intensity of career motivation.

The Finnish education system consists of comprehensive school, post-comprehensive general and vocational education, higher education and adult education. Comprehensive school provides a nine-year compulsory educational program for all school-age children, beginning at the age of seven. Post-comprehensive education is given by upper secondary schools and vocational schools or institutes. The Finnish higher education system consists of 20 universities and 31 polytechnic institution of higher educations. The higher education system as a whole offers openings for 66 per cent of the relevant age group (universities 29%, polytechnic institution of higher educations 37%).

Polytechnic institution of higher educations have been part of the Finnish school system now for only ten years. The polytechnic institution of higher education evaluation committee decided between 1992 - 1996, on the basis of the fourteen evaluation criteria, which of two-hundred vocational education institutions were promoted to polytechnic institution of higher educations (Liljander, 2002a, p. 10).

This paper has two main goals. First, to present a theoretical model of growth-oriented atmosphere and second, to demonstrate its practical use as a measurement instrument of growth prerequisites with different employee groups of a Finnish polytechnic institution of higher education.

Growth prerequisites are examined on the basis of a fourteen dimensional theoretical model of growth-oriented atmosphere developed by Ruohotie and Nokelainen (2000). The organization investigated in this study received its polytechnic institution of higher education status among the first in 1996. We conducted the first survey investigating growth prerequisites in the organization in 1998 (Ruohotie & Nokelainen, 2000). This research paper reports the findings of the second survey that was conducted in 2002.
Target population is the group of units about which information is wanted, and a survey population is the group of units that we are able to survey. Although the study reported here is based on a non-probability sampling, we believe that the polytechnic institution of higher education investigated in this study represents all the other 30 organizations at least to some extent. We base our claim on the fact that ten years ago all the Finnish vocational institutions that aimed to become high schools had to meet the same criteria (e.g., planning, function and goals of education, curriculum, the development of evaluation and feedback systems) evaluated by the same committee before they were promoted.

The following four research questions are to be considered. First, is the thirteen-factor model of the growth-oriented atmosphere relevant to describe organisational growth prerequisites? Second, does the empirical sample support the four domain model of growth-oriented atmosphere? Third, to what extent employees’ position is connected to growth motivation and commitment to the organization? Fourth, is employee's nature of contract connected to growth motivation and commitment to the organization?

The data analyses applied in this study are exploratory factor analysis for categorical indicators (research question 1), multidimensional scaling (research question 2) and Bayesian unsupervised model-based visualization (research questions 3 and 4).

The paper is organized as follows: First we give condensed up-to-date introduction to the theoretical model of growth-oriented atmosphere. Ruohotie (1996, 2000a, 2000b) has discussed the topic in more detailed level. Second, we represent empirical results considering the four research questions.

2. Growth Prerequisites in Organizations

Professional development includes all developmental functions, which are directed at the maintenance and enhancement of professional competency. In the modern world, updating is, ideally, a continual, lifelong process that addresses such goals as the acquisition of new and up-to-date information, the development of skills and techniques and the elevation of one’s personal esteem (Ruohotie, 1996). The maintenance and enhancement of competency is subject to the combined effect of many factors, ranging from personal traits to salient features of the work environment (Fishbein & Stasson, 1990).

Maurer and Tarulli (1994) have identified the following factors affecting the voluntary involvement of workers in development activities: 1) Perceptions related to the working
environment; 2) perceptions and beliefs regarding the benefits of development; 3) values and judgments; 4) personality factors including: 4.1) identification with work; 4.2) the personal concept of career; 4.3) the need for self-development; 4.4) self-efficacy.

2.1 Organizational Triggers

Changes in organizational structure, areas of responsibility and tasks often require the development of new skills. Individuals respond to such changes both effectively and behaviourally according to their perception of their circumstances, interpreting environmental events or situational change on the basis of personal values and perspectives. Research conducted as part of the Growth Needs Project in Finland show that the following factors are among the keys to the creation and maintenance of growth and high innovative capacity in an organization (Ruohotie, 1996).

Creation of a supportive culture: in a supportive environment innovation becomes a natural part of everyday work. Tasks may be intentionally defined in broad terms, encouraging change and emphasizing the possibility of choice.

Reward of development: in innovative organizations learning, initiative and experiment are prized as inherently valuable.

Supportive and participative management: in innovative organizations it is seen as the duty of management to create a workplace where each individual can reach his or her full potential.

Intensive communication: the more intensive the communication, the more effectively new ideas and alternative points of view can be shared and developed.

Security: in an era of intensifying competition, the organizations that will survive and succeed are those where there is a secure and confident atmosphere for employees. The fear of failure, of blame or of criticism is an effective damper to creative innovation.

Generating continuous enlivening innovation requires at least two things of an organization: First, it must learn to fully develop and utilize the capacity of its personnel, and second, it must show imagination at all times, suspending judgment temporarily when necessary in order to promote the development of new ideas.

2.2 Work Role Triggers

Research results of the Growth Needs Project indicate that motivational aspects of the work environment and the individual’s opportunities to influence it correlate positively with

2.3 Personal Triggers

Events or stages connected to everything from personal factors to life changes — for example, changes in family relationships, health, age and so forth — can cause an individual to reconsider his or her career priorities and goals. In addition, according to Hall (1986, 1990) and Ruohotie (1996), certain personal characteristics predispose an individual to make changes in order to avoid the negative consequences of work pressure or deal with personal frustration at the status quo (i.e., basic personality disposition, motivation for advancement, initiative, stress on performance, hardiness, flexibility, tolerance of ambiguity, independence).

2.4 Factors Contributing to Growth-oriented Atmosphere

Important factors in the development of growth orientation are support and rewards from the management, the incentive value of the job itself, the operational capacity of the team and work related stress. Each of these can further be divided into smaller individual factors.

Management and leaders face such challenges as how to empower people, support the development of their professional identity and how to create careers based on interaction. They should also aim to develop, reward, set goals and evaluate learning in the organization. Successful leadership creates commitment to the job and the organization.

The incentive value of the job depends on the opportunities it offers for learning, for example, the developing nature of the job. Therefore, essential factors for professional growth are the developmental challenges, the employees’ chances to influence, opportunities for the collaborative learning and valuation of the job.

The operational capacity of a team or a group can be defined by its members’ capability to operate and learn together, by the work group co-operation and by the reputation for effectiveness.

Work related stress might become an obstacle to professional growth. Ambiguity, vagueness and role conflicts, a too heavy mental load and demand for continual alterations may stress people and damage the organizational atmosphere. Negative stress quickly suppresses growth and development.
2.5 Theoretical Dimensions of Growth-oriented Atmosphere

In the earlier study dating back to 1998, Ruohotie and Nokelainen (2000) examined the theoretical dimensions of a growth-oriented atmosphere in the same organization as in the current study. The organization consisted of ten geographically separate units. The sample size was 318 employees, 66 per cent out of the survey population of 479 employees. The target population was Finnish polytechnic institution of higher education personnel in 1998 ($N = 7,958$).

Both male ($n = 145$) and female ($n = 147$) respondents’ group sizes were almost identical (46%) with eight per cent ($n = 27$) missing data. Respondents’ age was reported with four classes: 20 to 29 year (5%, $n = 17$), 30 to 39 year (25%, $n = 78$), 40 to 49 year (37%, $n = 120$), and over 50 year (24%, $n = 75$) with nine per cent ($n = 29$) missing data. The job profile contained three groups (7% of missing data): Managers (8%, $n = 25$), teachers (44%, $n = 139$) and other personnel, for example, cleaner, caretaker, librarian, (41%, $n = 131$).

Although the non-response rate was quite high in this study, the job title distribution of the sample (teachers: 44%, managers: 8%, other personnel: 41%, missing: 7%) was parallel both to the survey population (teachers: 47%, managers: 5%, other personnel: 48%) and target population (teachers: 63%, managers: 5%, other personnel: 32%) distributions derived from the public records.

The instrument utilized in the study contained 80 statements. The response options in a five-point summative rating scale (a.k.a. ‘Likert scale’, see DeVellis, 2003, 78-80) varied from 1 (strongly disagree) to 5 (strongly agree).

Ruohotie and Nokelainen (2000) constructed fourteen summated scales (Hair, Anderson, Tatham & Black, 1995, p. 9) to represent the theoretical dimensions of growth-oriented atmosphere. The scales were formed on the basis of both theoretical aspects and the results of exploratory factor analysis (Maximum likelihood with Varimax rotation). The thirteen-factor solution was the most parsimonious representing 67 per cent of the variance within the 80 items. Eigenvalues were between 1.05 and 23.98. Respondents indicated only moderate differences in preferences for various dimensions as mean ratings ranged between
3.2 and 3.8. Internal consistency for each factor was estimated with Cronbach’s alpha coefficient (1970, pp. 160-161). The alpha values ranged from .77 to .93 ($M_\alpha = .84$).

Although the authors report continuous parameters such as mean and alpha on items measured with the non-metric ordinal scale, we consider the results plausible as the underlying phenomenon, a growth-oriented atmosphere is continuous by nature (Marini, Li & Fan, 1996). Johnson & Creech (1983) have studied with simulation studies the categorization error that occurs when continuous variables are measured by indicators with only a few categories. The results indicated that while categorization error does produce distortions in multiple indicator models, under most conditions explored the bias was not sufficient to alter substantive interpretations. However, authors warranted caution in the use of two-, three- or four-category ordinal indicators, particularly when the sample size is small. In the Ruohotie and Nokelainen (2000) study, as well as in the present study, the ordinal scale has five categories and the sample size to the number of the observed variables ratio is acceptable according to empirical and simulation studies (Cattell, 1978; Gorusch, 1983; MacCallum, Widaman, Zhang & Hong, 1999).

Ruohotie and Nokelainen (2000) found that growth-oriented atmosphere generates togetherness and reflects on developing leadership. Multidimensional scaling provided evidence to conclude that factors representing the incentive value of the job, commitment to work and organization, the clarity of the job and growth motivation are the strongest indicators of growth-oriented atmosphere. Ruohotie and Nokelainen (2000) made the following conclusions based on their research findings: 1) Teacher’s professional growth-motivation reflects directly with task value on teacher-pupil relationships and on achievement motivation; 2) Task value has an effect on growth-oriented atmosphere; 3) Growth-oriented atmosphere is the highest in work assignments that offer challenging professional tasks (manager, teacher) and lowest among other workers.
3. Method

3.1 Sample

A non-probability sample included employees that worked in a Finnish polytechnic institution of higher education during the year 2002. The organization is the same as in the 1998 study (Ruohotie & Nokelainen, 2000), but as the organization structure was rearranged in 2000, the number of units has dropped from ten to eight. Four hundred forty-seven participants completed the questionnaire. The sample size is 87 per cent of the survey population of 512 workers, indicating 13 per cent non-response rate. The target population of Finnish polytechnic institution of higher education personnel in 2002 was 9,661. Non-response error was analyzed in the study by comparing job title distributions (manager/teacher/other) between the sample and public employee records. We conclude that the results of this study are not fully but cautiously generalizable to the target population of Finnish polytechnic institution of higher educations, as the target organization's job distribution in the survey population resembles the job distribution of target population.

The average age of respondents' in the sample was 39 years (SD = 9.1, range 22 - 62). Respondents’ job profiles were as follows (with 6%, n = 27 missing data): Teachers (48%, n = 215), managers (7%, n = 30) and other personnel (39%, n = 175).

A majority of the respondents were established employees (64%, n = 287), but the sample included also temporary (25%, n = 109), and part-time (6%, n = 28) workers. Eighty per cent of managers (n = 24) had established contracts and twenty per cent (n =6) had a temporary contract. Over the half of the teachers (67%, n = 143) had established contracts, twenty-one (10%) had part-time, and forty-eight (22%) had temporary contracts. Other personnel had the following contracts: 66 per cent (n = 115) established, three per cent (n = 6) part-time and 29 per cent (n = 51) temporary, respectively.

3.2 Instrument

The Growth-oriented Atmosphere Questionnaire (GOAQ) used in this study was a modified version of the one developed during the Growth Needs project (Ruohotie, 1996). The theoretical basis for the structure of the instrument elicited from the works of Argyris (1972, 1992), Dubin (1977, 1990), Hall (1986, 1990), and Kaufman (1974, 1990). The
latest version of the GOAQ is based on the research findings of the Growth Needs Project’s previous research phase (Ruohotie & Nokelainen, 2000). The original instrument contained 92 items operationalizing fourteen latent dimensions. Each item was measured in a five-point summative rating scale from 1 (strongly disagree) to 5 (strongly agree). According to the results of exploratory factor analysis for categorical variables (CEFA), the 67 strongest loading items were chosen to describe the thirteen dimensions of the updated growth-oriented atmosphere model (see Appendix). The dimension measuring students’ attitudes toward teacher in the 1998 study was dropped out in the current study as it is relevant only for the teachers who represent only 48 per cent of the sample. A demographics sheet was attached to the instrument enquiring respondents’ position in the organization and nature of the contract.

3.3 Procedure
The sample was obtained with non-probability sampling. Each employee of the organization was personally invited via email to complete an online version of the GOAQ. The online questionnaire (Miettinen, Nokelainen, Kurhila, Silander & Tirri, 2005) presented one to five questions at the same page allowing respondents to attach an open comment to each question.

Non-response error was analyzed by comparing the job profiles of the sample with survey population (teachers: 48%, managers: 6%, other personnel: 46%) and target population distributions (teachers: 57%, managers: 9%, other personnel: 34%) derived from public records. Comparison of the job profile distributions shows that the ‘other personnel’ group is seven per cent underrepresented in the sample when compared with the survey population. Teachers are nine per cent underrepresented in the sample when compared with the target population. The sample distribution of job profiles is similar enough to survey and target population distributions to represent Finnish polytechnic institution of higher education personnel in this study.

3.4 Data Analyses
Research questions in this study are addressed with unsupervised multivariate data analysis methods that allow ordinal indicators. Unsupervised methods (e.g., exploratory factor analysis) discover variable structure from the evidence of the data matrix as opposite to supervised methods (e.g., discriminant analysis) that assume a given structure (Venables &
Unsupervised methods are further divided into four sub categories: 1) Visualization methods, 2) cluster analysis, 3) factor analysis and 4) discrete multivariate analysis.

The first research question is investigated with exploratory factor analysis for categorical indicators (CEFA), that is implemented in Mplus (Muthén & Muthén, 2001), and Spearman non-parametric rank-order correlations. The use of CEFA has two major advantages over traditional exploratory factor analysis. First, it allows the use of ordinal indicators as it is based on the categorical variable model developed by Bengt Muthén (1993). Second, it does not require multivariate normality as it applies the general asymptotic distribution free function instead of the usual maximum likelihood estimator (Muthén & Muthén, 2001).

The other three research questions are investigated in this paper with non-linear visualization methods. According to Venables and Ripley (2002), visualization methods are often more effective than clustering methods discovering interesting groupings in the data, and they avoid the danger of over-interpretation of the results as researcher is not allowed to input number of expected latent dimensions. In cluster analysis the centroids that represent the clusters are still high dimensional, and some additional illustration methods are needed for visualization (Kaski, 1997), for example multidimensional scaling (Kim, Kwon & Cook, 2000). We apply in this study non-linear visualization methods as they are capable of analyzing both linear and non-linear dependencies between variables under investigation.

4. Results

Fred Kerlinger (1986) classified weaknesses of rating scales into extrinsic and intrinsic. Extrinsic defect is that scales are way too easy to construct and use. Sometimes a scale is used to measure things that it is not designed to measure. This point was addressed with a pilot study of 12 respondents and an interview of the organizations development manager. The online questionnaire that was used for the pilot study was the near-final version allowing respondents to attach an open comment to each question. This procedure is quite close to what Fowler (1995, pp. 130-131) calls ‘field pre-test with observation’ as with an online questionnaire we are able to ask item-specific comments and even track answering times for each item. The comments from the pilot study and interview were analyzed and wordings improved where necessary. The item structure from the pilot study was analyzed...
with Bayesian dependency modeling that is computationally robust also with small sample sizes. Results of the pilot study showed, for example, that the term ‘manager’ was not clear to all the respondents. Further, some of them did not understand the difference between ‘management’ and ‘manager’. We solved this problem by adding clear definitions of the terms in the opening page of the questionnaire.

According to Kerlinger (1986, p. 495), intrinsic defect of rating scales is their proneness to constant error. He lists four main sources: Halo effect, error of severity (to rate all items too low), error of leniency (to rate all items too high) and error of central tendency (to avoid all extreme judgments). To examine intrinsic defect we analyzed the overall response tendency. Results show that the respondents used the whole scale from 1 (totally disagree) to 5 (totally agree) for all the items but one. The scale for item v82 (“I find self-improvement useful”) range from 2 to 5. Mode frequencies that sum up to the number of items in the questionnaire were as follows: 1) strongly disagree, n = 0; 2) n = 9; 3) n = 27; 4) n = 54; 5) strongly agree, n = 2. This result is as overall distribution of the modes on a five-point summative rating scale is unimodal and only slightly biased towards positive values.

4.1 Research Question 1: Relevance of the thirteen-factor GOA model

Exploratory factor analysis for categorical indicators was conducted to solve the first research question: Is the thirteen-factor model of the growth-oriented atmosphere relevant to describe organisational growth prerequisites? In technical terms, our goal is to find the most relevant factorial structure for observed variables measuring growth-oriented atmosphere.

The GOAQ items were subject to categorical exploratory unweighted least squares factor analysis with Varimax rotation. An initial estimation yielded 14 factors with eigenvalues exceeding unity, accounting for 73 per cent of the total variance. Thirteen-factor Varimax-rotated solution, accounting for 71 per cent of the total variance was found to be most interpretable in terms of meaningful clusters and correspondence to both theoretical and empirical findings of our previous research work. The root mean square residuals (RMSR) help the investigator to examine how well the aspects of the data are captured by the model (Loehlin, 2004, p. 70). RMSR value of .03 was well below a cut-off value of .08 (Hu & Bentler, 1999). Figure 1 presents the thirteen-factor model of the
growth-oriented atmosphere. The individual items related to the dimensions are presented in the Appendix.

-- Insert Figure 1 about here --

Dimensions derived from the factor analysis are strongly related to each other as the correlation coefficients presented in Table 1 are significant at the .01 level (two-tailed). Spearman bivariate coefficients range between .81 and -.52. The average of all coefficients is .26 and the average of total variance explained is seven per cent. Closer examination of the coefficients reveals, as expected, that Growth motivation (GRM) is not affected by Strategic leadership (STR), Know-how rewarding (REW) or Build-up of work requirements (BUI). It is also noteworthy to mention that Psychic stress of the job (PSY) has the only positive correlation with Build-up of work requirements.

-- Insert Table 1 about here --

Growth prerequisites of a polytechnic institution of higher education can be described with the help of the thirteen dimensions that are presented in Table 2. Students’ attitude to teacher dimension that was present in the earlier solution of Nokelainen and Ruohotie (2000) was omitted from this model due to theoretical and technical reasons. Theoretical reason was that the factor is too tightly related to teaching, making it an irrelevant dimension for those employees who do not teach, for example, managers and other personnel. The second, more technical point favoring rejection of the factor was that the items operationalizing the dimension were not selective enough and the full scale was not in use.

Internal consistency measures estimate how consistently individuals respond to the items within a scale. Reliability is, thus, a characteristic of the data in hand, and not of the test (Thompson, 1998). Table 2 shows both lower (Cronbach’s alpha) and upper bound (Tarkkonen’s reliability, see Vehkalahti, 2000) of such measures. The scores in our study range from .75 to .97 (Cronbach’s alpha) and from .79 to .97 (Tarkkonen’s reliability). The most reliable factor was Encouraging leadership (ENC). This finding is partly due to fact
that alpha values tend to get larger as the number of items grows (ENC was measured with 15 items as the other dimensions had three to seven items).

-- Insert Table 2 about here --

4.2 Research Question 2: Validity of the four group classification of GOA factors

Non-metric multidimensional scaling was conducted in order to answer the second research question: Does the empirical sample support the four domain model of growth-oriented atmosphere? In technical terms, we examine what is the geometric two-dimensional structure of the components operationalizing growth-oriented atmosphere.

Figure 2 represents the structure of two dimensional distance measures between cases in our growth-oriented atmosphere data set. Euclidean distance as dissimilarity measure and distance scaling model was applied for ordinal data. First dimension classifies components into two groups. First group contains factors representing operational capacity of the team: Growth motivation (GRM), Incentive value of the job (INV) and Team spirit (TES). Factors in the second group are connected to supporting and rewarding management: Rewarding for know-how (REW), Strategic leadership (STR), and Clarity of the job (CLA). Second dimension is related to the supportive value of the job. Third group of factors in Figure 2 represent the lack of it (Psychic stress of the job, PSY, and increase in the workload, BUI). Fourth group of factors (Encouraging leadership, ENC, Valuation of the job, VAL, Know-how developing, DEV, Commitment to work and organization, COM and Community spirit, COS) represent the positive end of the scale. (Figure 2.)

-- Insert Figure 2 about here --

Examination of the coordinates for scaling Euclidean dimensions in two-dimensional space shows that Growth motivation (1.3945) and Incentive value of the job (1.1274) are the strongest components on the positive end of the first dimension and Psychical stress of the job (-2.4731) is strongest on the negative end together with Rewarding of know-how (-1.5332), and Strategic leadership (-1.4983). STRESS value (.049) indicates that the model fits to the data reasonably well. This result together with visual examination of the Figure 2
verifies the earlier research finding suggesting that Encouraging leadership (ENC) and Commitment to work and organization (COM) are closely situated in the visual space, but in different dimensions.

4.3 Research Question 3: Position and the nature of contract as predictors of growth motivation

Bayesian model-based visualization is applied in this study to investigate the third research question: To what extent employees’ position and the nature of contract are connected to growth motivation? With Bayesian unsupervised model-based visualization we may concentrate on singular summary factors and study each one’s distribution dynamically. Figure 3 is a visualization of the Bayesian network model. The window has following elements: Main window, attribute selection window (upper left corner), low profile window (lower left corner), initial profile window (lower right corner), and high profile window (upper right corner). Main window contains the model in which each dot stands for one respondent \( (n = 447) \). Attribute selection window shows the current component of interest and its discretization (i.e., the classes of data). In Figure 3, the component of interest is growth motivation (KM_GRM). Low profile window shows the distribution of examined variables when sub sample represents the lowest values of the component, high profile window has the same functionality for the high end sub sample. Initial profile window shows the initial distribution of the examined variables. Thin bars in profile window represent initial values, thick bars values of the current sub sample.

Attribute selection window in the upper left part of Figure 3 shows that growth motivation scale is quite biased and thus the upper bound for the lowest category is 3.55. However, inspection of the values in high-scale profile window and high scale sub sample frame gives evidence that managers and teachers has distinguished representation in the highest category of growth motivation as the thick bar is taller than the thin bar that indicates the average value. It is interesting to observe that those respondents with the most insecure contracts, namely temporary and part-time, have higher growth motivation than their established colleagues. (Figure 3.)

-- Insert Figure 3 about here --
4.4 Research Question 4: Position and the nature of contract as predictors of commitment to the organization

The fourth research question is to study to what extent employees’ position and the nature of contract is connected to his or her commitment to the organization. Attribute selection window in the upper left part of Figure 4 shows that scale for commitment to work and organization is balanced: Upper bound for the lowest category is 2.92 and lower bound for the highest category is 4.76. Values in high-scale profile window and high scale sub sample frame suggest that managers and teachers have the highest level of commitment to work as the thick bar is taller than the thin bar indicating the average value. This result is parallel with our earlier research findings (Ruohotie & Nokelainen, 2000). Commitment to work and organization is highest among those respondents with the most insecure contracts. (Figure 4.)

-- Insert Figure 4 about here --

5. Conclusion

We have examined in this paper dimensions of growth-oriented atmosphere in a Finnish polytechnic institution of higher education with categorical exploratory factor analysis, classical multidimensional scaling and Bayesian unsupervised model-based visualization. Thirteen-dimension Varimax-rotated solution in the categorical factor analysis was found to be interpretable in terms of meaningful clusters and correspondence to both theoretical and empirical findings of previous research (Ruohotie & Nokelainen, 2000).

Results of two-dimensional scaling showed that the components on the negative end of the first dimension represent operational capacity of the team. Components on the positive end of the first dimension are related to supporting and rewarding management. Second dimension visualized work-related stress; the most components with the most negative coordinates were psychical stress of the job and increase in the demands of the work. Rewarding for know-how, clarity of the job assignments, and encouraging leadership represented the positive polarity of the second dimension. Research evidence suggests that
the psychic stress caused by the work affects increasingly to the build-up of work requirements.

6. Discussion

The findings of a previous study (Ruohotie & Nokelainen, 2000) conducted in the same domain suggested that growth-oriented atmosphere generates togetherness and reflects on developing leadership. Multidimensional scaling and Bayesian unsupervised model-based visualization both provided evidence to conclude that factors representing encouraging leadership and commitment to work and organization are closely situated in, but in different dimensions. Results further showed that managers and teachers had the highest growth motivation and level of commitment to work. Employees across all job titles in the organization with temporary or part-time contracts, had higher self-reported growth motivation and commitment to work and organization than their established colleagues.

A recent study among 131 employees of a U.S. restaurant chain showed that conscientiousness was the best predictor of job performance against work experience, psychological climate and work effort (Byrne, Stoner, Thompson & Hochwarter, 2005). Results indicated that being conscientious might not be enough to secure the highest levels of performance unless the individual is concurrently willing to work hard, and is a member of a psychologically secure work setting. Majority of the sample worked part-time (82% versus 18% full-time). Work effort is related to following GOA dimensions: Valuation of the Job (VAL), Commitment to Work and Organization (COM) and Clarity of the Job (CLA).

According to Nokelainen, Ruohotie, Tirri & Silander (2002), encouraging leadership could be used as an indicator for empowerment (Walsh, Bartunek & Lacey, 1998; Spreitzer, DeJanasz & Quinn, 1999). We will focus our future studies on the relationship between Commitment to work and Strategic leadership, Incentive value of the job, Valuation of the job, and Psychical stress of the job.
References


Appendix

The Growth-oriented Atmosphere Questionnaire

<table>
<thead>
<tr>
<th>Item</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factor 1: Encouraging leadership (ENC)</strong></td>
<td></td>
</tr>
<tr>
<td>v5. My manager is friendly and easily approachable.</td>
<td>4.0</td>
</tr>
<tr>
<td>v6. My manager pays attention to my suggestions and wishes.</td>
<td>4.0</td>
</tr>
<tr>
<td>v7. My manager works with a team to find solutions.</td>
<td>4.0</td>
</tr>
<tr>
<td>v8. My manager is fair.</td>
<td>4.0</td>
</tr>
<tr>
<td>v9. The employees in my organization are encouraged to develop new working methods and to think creatively.</td>
<td>4.0</td>
</tr>
<tr>
<td>v10. My manager trusts his or her staff and allows them to work independently.</td>
<td>4.0</td>
</tr>
<tr>
<td>v11. The organization promotes self-reliance and employees are encouraged to find new and improved working methods.</td>
<td>4.0</td>
</tr>
<tr>
<td>v13. The managers are interested in the wellbeing of staff.</td>
<td>3.0</td>
</tr>
<tr>
<td>v14. The management strives to improve the working conditions of staff.</td>
<td>4.0</td>
</tr>
<tr>
<td>v15. My goals were agreed in co-operation with my manager.</td>
<td>4.0</td>
</tr>
<tr>
<td>v23. Failures are dealt with in a constructive manner and employees are encouraged to learn from their mistakes.</td>
<td>3.0</td>
</tr>
<tr>
<td>v25. My manager has supported me in the past.</td>
<td>4.0</td>
</tr>
<tr>
<td>v26. My manager knows how to tap into the differing characteristics within the workforce.</td>
<td>3.0</td>
</tr>
<tr>
<td>v27. My manager has succeeded in strengthening the sense of unity in the workplace.</td>
<td>3.0</td>
</tr>
<tr>
<td>v90. This organisation values me as an individual.</td>
<td>4.0</td>
</tr>
<tr>
<td><strong>Factor 2: Strategic leadership (STR)</strong></td>
<td></td>
</tr>
<tr>
<td>v1. The management of my organization provides a clear direction and highlights the key points in education.</td>
<td>3.0</td>
</tr>
<tr>
<td>v2. The management of my organization expresses and enforces accepted values both in spoken form and through its example.</td>
<td>3.0</td>
</tr>
<tr>
<td>v3. The management of my organization embodies distinct values and a clearly defined style of leadership.</td>
<td>3.0</td>
</tr>
<tr>
<td>v4. The management of my organization observes the latest educational developments and uses this information when planning the organization’s activities.</td>
<td>3.0</td>
</tr>
<tr>
<td><strong>Factor 3: Know-how rewarding (REW)</strong></td>
<td></td>
</tr>
<tr>
<td>v20. It is rewarding to achieve my goals.</td>
<td>2.0</td>
</tr>
<tr>
<td>v21. The organization rewards its employees’ professional knowledge and skills.</td>
<td>2.0</td>
</tr>
<tr>
<td>v22. Employees with increased knowledge are given extra responsibility.</td>
<td>3.0</td>
</tr>
<tr>
<td>v24. The organization rewards employees for tackling demanding tasks.</td>
<td>3.0</td>
</tr>
<tr>
<td><strong>Factor 4: Know-how developing (DEV)</strong></td>
<td></td>
</tr>
<tr>
<td>v37. The organization endeavours to always use the latest knowledge in the field.</td>
<td>4.0</td>
</tr>
<tr>
<td>v38. The organization’s employees are given training to increase their professional skills.</td>
<td>3.0</td>
</tr>
<tr>
<td>v39. The organization takes an active interest in its employees’ professional growth.</td>
<td>3.0</td>
</tr>
<tr>
<td>v40. The staff is given the latest information and professional literature.</td>
<td>4.0</td>
</tr>
<tr>
<td>v41. I am given the chance to learn new things and improve myself.</td>
<td>4.0</td>
</tr>
</tbody>
</table>

Note. Data = Employees’ of a Finnish polytechnic institution of higher education. A five-point summative rating scale ranging from 1 (strongly disagree) to 5 (strongly agree) was used.
The Growth-oriented Atmosphere Questionnaire (continued)

<table>
<thead>
<tr>
<th>Item</th>
<th>Median</th>
<th>Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor 5: Incentive value of the job (INV)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>v28. I can work independently and without restrictions.</td>
<td>4.0</td>
<td>4</td>
</tr>
<tr>
<td>v29. I can use my skills at work in a variety of ways.</td>
<td>4.0</td>
<td>4</td>
</tr>
<tr>
<td>v30. My work consists of various differing tasks.</td>
<td>4.0</td>
<td>5</td>
</tr>
<tr>
<td>v31. My work gives me a sense of success and achievement.</td>
<td>4.0</td>
<td>4</td>
</tr>
<tr>
<td>v32. My work gives me personal satisfaction.</td>
<td>4.0</td>
<td>4</td>
</tr>
<tr>
<td>Factor 6: Clarity of the job (CLA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>v46. A clear division of tasks exists between members of teaching staff.</td>
<td>3.0</td>
<td>4</td>
</tr>
<tr>
<td>v47. The organization’s decision making structure is transparent.</td>
<td>3.0</td>
<td>3</td>
</tr>
<tr>
<td>v48. The organization’s goals are transparent.</td>
<td>3.0</td>
<td>4</td>
</tr>
<tr>
<td>v49. The teachers know exactly what their colleagues expect of them.</td>
<td>3.0</td>
<td>3</td>
</tr>
<tr>
<td>Factor 7: Valuation of the job (VAL)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>v42. My manager appreciates my work.</td>
<td>4.0</td>
<td>4</td>
</tr>
<tr>
<td>v43. I am given encouraging feedback on my work.</td>
<td>3.0</td>
<td>4</td>
</tr>
<tr>
<td>v45. I feel that my work is valued.</td>
<td>4.0</td>
<td>4</td>
</tr>
<tr>
<td>Factor 8: Community spirit (COS)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>v54. The organization’s staff feels personally responsible for achieving their goals.</td>
<td>4.0</td>
<td>4</td>
</tr>
<tr>
<td>v55. The staff maintains a demand for high performance.</td>
<td>4.0</td>
<td>4</td>
</tr>
<tr>
<td>v56. The staff possesses a sense of unity and a willingness to strive towards a common goal.</td>
<td>4.0</td>
<td>4</td>
</tr>
<tr>
<td>v57. My colleagues help me when necessary.</td>
<td>4.0</td>
<td>4</td>
</tr>
<tr>
<td>v58. The staff discusses improvements to work and/or their working environment.</td>
<td>4.0</td>
<td>4</td>
</tr>
<tr>
<td>v59. The staff presents new ideas about solving work-related problems.</td>
<td>4.0</td>
<td>4</td>
</tr>
<tr>
<td>v60. The staff wants to improve the quality of teaching.</td>
<td>4.0</td>
<td>4</td>
</tr>
<tr>
<td>Factor 9: Team spirit (TES)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>v50. I have ample opportunities to exchange work-related ideas and experiences with my colleagues.</td>
<td>4.0</td>
<td>4</td>
</tr>
<tr>
<td>v51. We tend to evaluate and analyze our work together to learn from it.</td>
<td>3.0</td>
<td>3</td>
</tr>
<tr>
<td>v52. We solve work-related problems together.</td>
<td>4.0</td>
<td>4</td>
</tr>
<tr>
<td>v53. We advise and guide each other on executing work-related tasks.</td>
<td>4.0</td>
<td>4</td>
</tr>
<tr>
<td>Factor 10: Psychic stress of the job (PSY)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>v78. I feel that I am beginning to dislike my work.</td>
<td>2.0</td>
<td>2</td>
</tr>
<tr>
<td>v79. I feel that it is getting more difficult for me to take the initiative.</td>
<td>2.0</td>
<td>1</td>
</tr>
<tr>
<td>v80. I find it difficult to concentrate.</td>
<td>2.0</td>
<td>1</td>
</tr>
</tbody>
</table>

Note. Data = Employees’ of a Finnish polytechnic institution of higher education. A five-point summative rating scale ranging from 1 (strongly disagree) to 5 (strongly agree) was used.
The Growth-oriented Atmosphere Questionnaire (continued)

<table>
<thead>
<tr>
<th>Item</th>
<th>Median</th>
<th>Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factor 11: Build-up of work requirements (BUI)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>v70. My workplace has too few employees to cope with the workload.</td>
<td>4.0</td>
<td>4</td>
</tr>
<tr>
<td>v72. My workload has increased during the past years.</td>
<td>4.0</td>
<td>5</td>
</tr>
<tr>
<td>v76. My working pace has increased in recent years.</td>
<td>4.0</td>
<td>4</td>
</tr>
<tr>
<td>v77. I feel that I am experiencing fatigue.</td>
<td>3.0</td>
<td>4</td>
</tr>
<tr>
<td><strong>Factor 12: Commitment to work and organization (COM)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>v87. I am happy in my present job.</td>
<td>4.0</td>
<td>4</td>
</tr>
<tr>
<td>v88. &quot;I want to continue in my present job; it gives me job satisfaction.”</td>
<td>4.0</td>
<td>4</td>
</tr>
<tr>
<td>v89. I don’t find going to work each morning disagreeable.</td>
<td>4.0</td>
<td>5</td>
</tr>
<tr>
<td>v91. I do not wish to change jobs.</td>
<td>4.0</td>
<td>4</td>
</tr>
<tr>
<td><strong>Factor 13: Growth motivation (GRM)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>v81. I feel encouraged by having added responsibilities.</td>
<td>4.0</td>
<td>4</td>
</tr>
<tr>
<td>v82. I find self-improvement useful.</td>
<td>5.0</td>
<td>5</td>
</tr>
<tr>
<td>v83. I like to participate in all manner of improvement projects within the organization (such as training, team work and projects, exchanging duties, taking on additional tasks etc).</td>
<td>4.0</td>
<td>4</td>
</tr>
<tr>
<td>v84. I am interested in further training, provided it speeds up my transfer to other, more challenging tasks.</td>
<td>4.0</td>
<td>5</td>
</tr>
<tr>
<td>v85. I like to experiment with new ideas.</td>
<td>4.0</td>
<td>4</td>
</tr>
</tbody>
</table>

*Note. Data = Employees’ of a Finnish polytechnic institution of higher education. A five-point summative rating scale ranging from 1 (strongly disagree) to 5 (strongly agree) was used.*
Figure 1. Thirteen-factor Model of the Growth-oriented Atmosphere.

Figure 2. The Growth-oriented Atmosphere Factors in Two-dimensional Space (MDS, Euclidean Distance Model).

Note. ENC = Encouraging leadership, STR = Strategic leadership, REW = Know-how rewarding, DEV = Know-how developing, INV = Incentive value of the job, CLA = Clarity of the job, VAL = Valuation of the job, COS = Community spirit, TES = Team spirit, PSY = Psychic stress of the job, BUI = Build-up of work requirements; COM = Commitment to work and organization, GRM = Growth motivation.
Figure 3. Bayesian Model-based Visualization of Growth Motivation by Employees’ Position and the Nature of Contract.

Figure 4. Bayesian Model-based Visualization of Commitment to Work and Organization by Employees’ Position and the Nature of Contract.
<table>
<thead>
<tr>
<th>Growth-oriented Atmosphere Factors</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Encouraging leadership (ENC)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Strategic leadership (STR)</td>
<td>.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Know-how rewarding (REW)</td>
<td>.65</td>
<td>.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Know-how developing (DEV)</td>
<td>.68</td>
<td>.40</td>
<td>.61</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Incentive value of the job (INV)</td>
<td>.60</td>
<td>.27</td>
<td>.41</td>
<td>.59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Clarity of the job (CLA)</td>
<td>.72</td>
<td>.47</td>
<td>.57</td>
<td>.61</td>
<td>.48</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Valuation of the job (VAL)</td>
<td>.85</td>
<td>.37</td>
<td>.60</td>
<td>.66</td>
<td>.61</td>
<td>.63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Community spirit (COS)</td>
<td>.55</td>
<td>.32</td>
<td>.39</td>
<td>.55</td>
<td>.46</td>
<td>.48</td>
<td>.56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Team spirit (TES)</td>
<td>.48</td>
<td>.30</td>
<td>.36</td>
<td>.49</td>
<td>.37</td>
<td>.44</td>
<td>.49</td>
<td>.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Psychic stress of the job (PSY)</td>
<td>-.30</td>
<td>-.19</td>
<td>-.21</td>
<td>-.31</td>
<td>-.40</td>
<td>-.30</td>
<td>-.35</td>
<td>-.24</td>
<td>-.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Build-up of work requirements (BUI)</td>
<td>-.19</td>
<td>-.19</td>
<td>-.25</td>
<td>-.16</td>
<td>-.12</td>
<td>-.20</td>
<td>-.23</td>
<td>-.06</td>
<td>-.06</td>
<td>.41</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Commitment to work and organization (COM)</td>
<td>.55</td>
<td>.34</td>
<td>.42</td>
<td>.49</td>
<td>.61</td>
<td>.47</td>
<td>.58</td>
<td>.36</td>
<td>.32</td>
<td>-.50</td>
<td>-.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Growth motivation (GRM)</td>
<td>.20</td>
<td>.04</td>
<td>.04</td>
<td>.24</td>
<td>.29</td>
<td>.11</td>
<td>.21</td>
<td>.30</td>
<td>.22</td>
<td>-.23</td>
<td>.03</td>
<td>.19</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Spearman rank order correlations ($r_s$) were calculated due to ordinal measurement scale.
<table>
<thead>
<tr>
<th>Growth-oriented Atmosphere Factor</th>
<th>Description</th>
<th>$\alpha^b$</th>
<th>$TR^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Encouraging leadership (ENC)$^a$</td>
<td>Management of the organization expresses and consolidates values that direct activities, monitors the development processes of units and defines the direction and focus of operations.</td>
<td>.97</td>
<td>.97</td>
</tr>
<tr>
<td>2. Strategic leadership (STR)</td>
<td>Manager supports and motivates personnel to develop know-how, work methods and work community. He takes advantage of work community member's expert knowledge and he tries to solve problems with them. He pays attention to the expectations and wishes of personnel.</td>
<td>.89</td>
<td>.90</td>
</tr>
<tr>
<td>3. Know-how rewarding (REW)</td>
<td>Organization rewards its employees' professional knowledge and skills. Members of work community gain more responsibility as their know-how increases.</td>
<td>.87</td>
<td>.87</td>
</tr>
<tr>
<td>4. Know-how developing (DEV)</td>
<td>Organization takes an active interest in its employee's professional growth. Members of work community are interested in self-developing.</td>
<td>.88</td>
<td>.90</td>
</tr>
<tr>
<td>5. Incentive value of the job (INV)$^a$</td>
<td>Work gives intrinsic fulfillment by being versatile, autonomous and challenging.</td>
<td>.88</td>
<td>.90</td>
</tr>
<tr>
<td>6. Clarity of the job (CLA)</td>
<td>Personnel has a clear picture of goals and responsibilities. They are aware of decision-making processes and personal expectations.</td>
<td>.87</td>
<td>.90</td>
</tr>
<tr>
<td>7. Valuation of the job (VAL)$^a$</td>
<td>Work contribution is respected by the worker itself, colleagues and management.</td>
<td>.88</td>
<td>.90</td>
</tr>
<tr>
<td>8. Community spirit (COS)$^b$</td>
<td>How community members may learn from each other, for example via dialogue, by analyzing mistakes, participating in collaborative planning and quality development.</td>
<td>.92</td>
<td>.93</td>
</tr>
</tbody>
</table>

$^a$Common dimension as in the previous study in the same organization (Ruohotie & Nokelainen, 2000) with an 80-item version of the questionnaire. $^b$Cronbach’s index of internal consistency. $^c$Tarkkonen’s reliability index.
Table 2
*The Thirteen Dimensions of Growth-oriented Atmosphere (continued)*

<table>
<thead>
<tr>
<th>Growth-oriented Atmosphere Factor</th>
<th>Description</th>
<th>$\alpha^b$</th>
<th>$TR^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>9. Team spirit (TES)</td>
<td>Good team spirit promotes helping each other and taking responsibility over common goals. Work group members discuss about developing work and working environment.</td>
<td>.87</td>
<td>.88</td>
</tr>
<tr>
<td>10. Psychic stress of the job (PSY)</td>
<td>To what extent work and changes relating to it induce psychic strain like fatigue flightiness.</td>
<td>.83</td>
<td>.85</td>
</tr>
<tr>
<td>11. Build-up of work requirements (BUI)</td>
<td>How to cope with changes in the personal workload.</td>
<td>.75</td>
<td>.79</td>
</tr>
<tr>
<td>12. Commitment to work and organization (COM)$^a$</td>
<td>To be truly excited about ones work. How important it is to stay in current job.</td>
<td>.87</td>
<td>.89</td>
</tr>
<tr>
<td>13. Growth motivation (GRM)$^a$</td>
<td>To trust ones abilities in difficult situations, take new challenges and develop ones know-how.</td>
<td>.80</td>
<td>.81</td>
</tr>
</tbody>
</table>

$^a$Common dimension as in the previous study in the same organization (Ruohotie & Nokelainen, 2000) with an 80-item version of the questionnaire. $^b$Cronbach’s index of internal consistency. $^c$Tarkkonen’s reliability index.
Investigating the Influence of Attribution Styles on the Development of Mathematical Talent

Petri Nokelainen

University of Tampere, Finland

Kirsi Tirri

University of Helsinki, Finland

Hanna-Leena Merenti-Välimäki

Espoo-Vantaa Institute of Technology, Finland

The final, definitive version is available at http://online.sagepub.com". 
Abstract

In this paper, we examine the influence of attribution styles on the development of mathematical talent. The study employs a Self-confidence attitude attribute Scales (SaaS) questionnaire (Campbell, 1996b) measuring ability and effort attributions (Weiner, 1974). Participants were three groups of mathematically highly, moderately and mildly gifted Finnish adolescents and adults (n = 203). The results of Bayesian classification modeling showed that items attributing success to effort and failure to lack of effort were the best predictors for the level of mild mathematical giftedness and gender (females). The results of multivariate analysis of variance showed that mathematically highly and moderately gifted reported that ability was a more important reason for success than effort, but mathematically mildly gifted tended to see effort leading to success. Mathematically moderately and mildly gifted attributed failure to lack of effort, mathematically highly gifted attributed failure to lack of ability.
Putting the Research to Use

It is essential that educators and parents understand the influence of different attribution styles on the development of mathematical talent. This study provides understanding how mathematically highly, moderately and mildly gifted adolescents and adults differ in their specific reasons for success and failure. Differences in attribution styles between the three groups of mathematically gifted, measured with the SaaS questionnaire, indicate that it is important to know if the attributions for success or failure are stable or unstable, external or internal.

Knowledge of how learners or trainees use attributions to account for success and failure can help educators and parents to gain a deeper awareness of the mathematically gifted and, thus, predict their expectancies and plan intervention strategies when needed. The information is also applicable to courses concerning the needs of the gifted. Furthermore, the information can be presented directly to mathematically gifted in order to help them develop more insight into their own behavior.
Introduction

In 2000, a total of 180,000 students from 28 Organization for Economic Co-Operation and Development (OECD) member countries and four non-OECD countries (Brazil, Latvia, Liechtenstein and the Russian Federation) participated in the first Programme for International Student Assessment (PISA). The results showed that students from Japan, Korea, New Zealand and Finland scored highest in all tests measuring mathematics literacy (OECD, 2001, p. 78). The Finnish students ranking was even higher, the third, when variation within country was taken into account (OECD, 2001, 80). In 2003 the PISA follow up study focusing on mathematics literacy and involving 276,165 15-year old students was conducted in 41 countries (OECD, 2004). The results of overall student performance in different countries on the mathematics scale showed that Hong Kong students had the highest, and Finnish students had the second highest, mean student score (OECD, 2004, p. 89). The finding of small within country variance in the Finnish sample was repeated.

One logical reason for success in international comparison studies is the Finnish government’s “equal opportunities and high quality education for all” principle. The first practical consequence of the principle is that education is free for all students participating in these assessments. The second consequence is government’s strong financial support for public sector educational institutions. This has led to the situation where there are no appreciable differences in teaching quality or premises between public and special schools. Partly for this reason, only a small minority of the schools in Finland are special schools with entrance examinations and financial support from private or corporate sources. There are no private universities or polytechnics in Finland.

The purpose of this study is to explore the attribution styles—that is, personal explanations for success and failure—Finnish adolescents and adults \((n = 203)\) with varying levels of mathematical giftedness to discover what attributions contribute to or impede the development of mathematical talent.

The first group, “Olympians,” consists of mathematically highly gifted adults who have participated in the international Olympics for mathematics. Tirri and Campbell (2002) reported that 80 per cent of the Finnish Olympians apply their mathematical talent by choosing a career in science. The majority of them are researchers in academia or engineers in technical fields. The Olympians have been very successful in their graduate studies, and they have published articles and books related to their fields. Those Olympians who did not continue in academia chose a career as an engineer or as a CEO or a manager in leading Finnish companies like Nokia. (Tirri, 2002; Tirri & Campbell, 2002.)
The second group, “Prefinalists,” consists of secondary school students, who have taken part in national competitions in mathematics. The group represents the top level of Finnish 15-year old students that participated in the international PISA 2000 study.

The third group, “Polytechnics,” consists of adolescent students from a technical vocational high school who study mathematics as their major subject. In Finland most of the vocational high schools are highly specialized regional institutions training professionals for expert and development careers. This particular institute is the top-rated technically oriented vocational high school in Finland.

Giftedness is not a monolithic construct. There are different levels of giftedness and, thus, the three groups representing mathematically gifted adolescents and adults in this study are not homogenous. Further, we are not able to guarantee that the individuals within each group share the same level of mathematical ability. Intelligence quotient is, especially with children, a useful index of the discrepancy between mental and chronological age. As the participants of this study are adolescents and adults, we did not measure their IQ, but instead we looked at their current or past achievements. Olympians are the most homogenous and mathematically gifted group in this study on the basis of their achievements as Academic Olympians and their traceable academic publication record (Nokelainen, Tirri & Campbell, 2004; Nokelainen, Tirri, Campbell & Walberg, 2004). We classify Olympians for the purpose of this study as mathematically highly gifted. Also Prefinalists have undergone a series of increasingly demanding mathematical tests in order to be included in the Academic Olympians training programme. Their trainers are past Olympians; that is, members of the first group in this study. Prefinalists are classified as mathematically moderately gifted, as we do not yet know how many will be selected to participate as Academic Olympians in the future. Technical vocation high school students, who study mathematics as their major, represent mathematically mildly gifted students in this study. Group membership (1 = “Olympians”, 2 = “Prefinalists”, 3 = “Polytechnics”) showed a strong positive correlation with secondary school mathematics grade average (from 1 (highest) to 7), with a correlation coefficient of \( r(203) = .82, p < .001 \).

Earlier studies of mathematical giftedness have mainly focused on within group differences related to, for example, gender or attribution styles. There are very few between group comparisons, except cross-cultural, reported. Socio-economic differences do exist in Finland, but their impact on children’s educational possibilities is minor because education is free at all levels. As the PISA results indicate (OECD, 2001, 2004), in Finland all individuals are provided with uniformly high-level of basic mathematical training, thus, controlling at least to some extent individual level educational differences. This allows us to interpret possible differences between the groups through differences in individuals’ characteristics such as mathematical giftedness and attribution styles.
All the participants completed the Self-confidence attitude attribute Scales (SaaS) questionnaire (Campbell, 1996a). The instrument included 18 items measuring ability and effort attributions, based on Weiner’s properties of attributional thinking (1974, 1980, 1986, 1994, 2000), on four dimensions: 1) Success due to ability; 2) Failure due to lack of ability; 3) Success due to effort; 4) Failure due to lack of effort.

Our research questions are as follows: 1) “Are the four dimensions of the SaaS instrument (success due to ability, failure due to lack of ability, success due to effort, failure due to lack of effort) identified in this domain?”; 2) “What are the best predictors for the level of mathematical giftedness (highly = “Olympians,” moderately = ”Prefinalists” and mildly = “Polytechnics”) and gender among the SaaS variables?”; 3) “Do the attribution styles differ by the level of mathematical giftedness or gender?”

Theoretical Framework

Properties of Attributional Thinking

Reasons people give for an outcome, such as success or failure in a task, are called attributions (Heider, 1958). Factors, such as specific reason for success and failure, involved in attributional thinking have been shown to be related in achievement settings (Weiner, 1974, 1980, 1986, 1994, 2000). Weiner found in his studies that the four most frequent reasons for success and failure are ability, effort, task difficulty and luck. Subsequent research identified learning strategies as a fifth possible reason for success and failure (Alderman, 2004): It is no good thing trying harder if you do not know how to try.

Dai, Moon and Feldhusen (1998) classify attribution constructs into three groups. First, attribution appraisals are online explanations assessed following actual or manipulated success or failure in performing a specific task. Second, attribution beliefs are domain-specific or domain-general beliefs about the causes of success or failure. Third, attribution styles are generalized, stereotypical patterns of attributions and dispositional beliefs. Attribution styles are assessed in a similar way to attribution beliefs, except that a certain typology is imposed on the data using predetermined criteria. In this study, we examined attribution styles using Weiner’s (1992) classification of reasons for success and failure: 1) Internal and external attributions, referring to within or outside person causes; 2) Stable and unstable attributions, referring to consistent or inconsistent causes over time; 3) Controllable and uncontrollable attributions, referring to the extent a person believes he or she has control over the cause of an outcome. In this study, we examined within person factors (ability and effort) as they have typically been found to be the most frequently cited reasons for success and failure in achievement contexts. Those factors are classified as ‘internal’ attributions. ‘External’ attributions (luck, task difficulty) were
omitted from the study design. Thus, our focus is on stable and unstable internal, controllable and uncontrollable, attributions. For example, most effort attributions are unstable and controllable as opposed to ability attributions that are usually stable and uncontrollable. We will later show in Table 2 how the 18 SaaS items are related to these dimensions. We will also discuss in the later stages of the analysis how the four SaaS factors describe above-mentioned dimensions of reasons for success and failure (Table 4).

**Self-regulation and Attribution Styles**

Self-regulation refers to the process through which self-generated thoughts, feelings, and actions are planned and systematically adapted as necessary to affect one’s learning and motivation (Schunk & Ertmer, 2000, p. 631; Zimmerman, 2000, p. 14). According to social-cognitive theory, self-regulation is dependant on the situation. Therefore, self-regulation is not a general characteristic or a developmental level but is contextually dependent.

Zimmerman (2000) describes self-regulation as cyclical because the feedback from prior performance is used to make adjustments during current efforts. Personal, behavioral, and environmental factors are constantly changing, and therefore an individual has to monitor these changes continuously in order to know whether any adjustments are required. Zimmerman (2000) describes the three feedback loops involved in monitoring one’s internal state, one’s behaviors and one’s environment as the triadic forms of self-regulation.

Figure 1 describes self-regulation of learning tasks as a cyclical, three-phase process (Zimmerman, 1998). The phases in this learning cycle are forethought, performance or volitional control, and self-reflection. **Forethought**, which creates the necessary conditions for learning, consists of task analysis and self-motivation beliefs. **Performance or volitional control**, which guides the learning process and regulates concentration and learning performance, consists of self-control and self-observation. **Self-reflection**, which refers to examining and making meaning of the learning experience, consists of self-judgment and self-reaction. Next we examine more closely the last phase, which contains the focus of this paper, attribution styles.

Self-reflection begins with self-judgment, which is the process whereby individual compares information attained through self-monitoring to extrinsic standards or goals. He or she wants to have fast and accurate feedback on his or her performance as compared to others. Self-judgment leads to attribution interpretations where the learner interprets the reasons for success or failure. Attribution interpretations can lead to positive self-reactions. The individual might interpret the failure of a strategy as the result of too little effort and then increase her efforts, but if she interprets the reason for failure as
being a lack of ability, the reaction is liable to be negative. Attribution interpretations reveal the possible reasons for learning mistakes and help the learner to find those learning strategies which best suit the given situation. They also develop or promote the adaptation process. Self-regulated individuals are more adaptive and evaluate their performance appropriately. Positive reactions (e.g., self-satisfaction) reinforce positive interpretations of oneself as an individual and enhance intrinsic interest in the task.

-- Insert Figure 1 about here --

Ellström (2001) defines qualification as the competence that is actually required by a task and/or is implicitly or explicitly determined by individual qualities. In our study setting, the most interesting point is that competence may also be seen as an attribute of the individual, meaning for example a human resource that the person brings to mathematical problem solving situation. Further, attributions may emphasize formal competence as indicated by degree requirements and certificates, or, the focus of this study, potential competence as indicated by the capacity of the individual to successfully complete tasks and face new challenges on the basis of demonstrated personal attributes and abilities (other than those obtained through formal training). Ellström (2001) has noticed that potential competence may vary greatly between individuals with the same formal qualifications, because they may possess very different levels of inherent ability and may have learned different things outside of school or studies through their working life and recreational activities. Thus, ability attributions affect later performance expectations and, in negative cases, the development or continuation of learned helplessness (Ruohotie & Nokelainen, 2000).

In this study, we concentrate on participants’ self-evaluations on the basis of mathematics achievement and academic ability because causal attributions (see phase “Self-Reflection” in Figure 1) play an important part in the self-regulatory process by being central elements of self-judgment and thus influencing, for example, goal setting and self-efficacy. We are interested to see if the attribution styles of mathematically highly gifted individuals differ from those of the mathematically able individuals.
Investigating the Influence of Attribution Styles

Literature Review

Mathematical Giftedness and Attribution Styles
Campbell has conducted several cross-national studies on Mathematics Olympians (e.g., 1994, 1996b; Nokelainen, Tirri & Campbell, 2004). He made two interesting findings: first, the international data on mathematics self-concept verified the finding that their academic self-concepts fluctuate from grade school to high school, and second, the Olympians attributed effort to be more important in their success than ability (1996b). The latter research finding has been verified by Chan (1996) who reported that adolescent gifted students were more likely to attribute failure to lack of effort than to attribute it to low ability. The American and Taiwanese Olympians have also attributed success and failure more to effort than to ability (Feng, Campbell & Verna, 2001; Wu & Chen, 2001).

Heller and Lengfelder (2000) investigated 100 German Olympians finalists and 135 Prefinalists in mathematics, physics, and chemistry. In contrast to Campbell’s findings, their results showed that participants in both groups valued ability significantly more highly than effort. Effort was estimated to be equally important in the case of failure as in the case of success (Heller & Lengfelder, 2000).

Marsh (1983) found, as he studied relationships among the dimensions of self-attribution, self-concept and academic achievements, that those who attribute academic success to ability and who do not attribute failure to a lack of ability have better academic self-concepts and better academic achievement. Multon, Brown and Lent (1991) have also shown a positive correlation between perceived ability and achievement.

Gender and Attribution Styles
In an American study by Verna and Campbell (2000), a small significant difference between males and females was found with regard to perceptions of ability. The female American Chemistry Olympians considered ability to be a more important factor for success than did the males. However, no difference was found for the effort factor.

Kerr (1994) and Reis (1998) have identified external barriers to gifted women as including the attitudes of parents and school, environmental options and possible discrimination or harassment at school or at work. The possible internal barriers among gifted females included self-doubt, self-criticism, and low expectations. According to Siegle and Reis (1998), gifted girls tend to underestimate their abilities, especially in mathematics, social studies and science.
Instrumentation of Attribution Theory

There is abundant literature and research on attribution theory, especially on attributional properties in achievement settings (Weiner, 1974, 1980, 1986, 1994, 2000) because the role of motivation in academic achievement has proven to be a popular topic. The principle of attribution theory is that students search for understanding, trying to discover why an event has occurred (Weiner, 1974). The interest is apparent as we examine the structure of existing measurement instruments: Biggs’ (1985) 42-item Study Process Questionnaire (SPQ) consists of two scales (motive and strategy) with three approaches: 1) Surface; 2) Deep; 3) Achieving. The questionnaire contains six subscales (surface motive, deep motive, achieving motive, surface strategy, deep strategy, and achieving strategy). Entwistle et. al’s (Ramsden & Entwistle, 1981) Approaches to Studying Inventory (ASI), which is one of the most widely used questionnaire on student learning in higher education, contains subscales including such factors as fear of failure, extrinsic motivation, and achieving orientation. Marsh (e.g., Marsh & O’Neill, 1984) has developed a set of scales (Self-description Questionnaire I-III) for different age groups measuring self-concept with a multifaceted (e.g., mathematics, verbal, academic, physical) view. According to Strein (1995), research results over the past fifteen years have strongly supported the multifaceted view emphasizing domain-specific self-concepts. In this study we apply the Self-confidence attitude attribute Scales (SaaS) questionnaire that was developed by Campbell (1996a) originally for cross-cultural Academic Olympiad studies.

Method

Sample

The Finnish education system includes comprehensive schools, post-comprehensive general and vocational education, higher education and adult education. Comprehensive schools provide a nine-year compulsory educational program for all school-age children, beginning at the age of seven. Post-comprehensive education is provided in upper secondary schools and vocational institutions. The Finnish higher education system includes 20 universities and 30 vocational high schools. The higher education system as a whole offers openings for 66 per cent of the relevant age group (universities 29%, vocational high schools 37%).

Respondents in the first group, “Olympians,” are the Finnish students most gifted in mathematics. The group consists of individuals of different ages who participated in Olympiad Studies in Mathematics during the years 1965-1999. Separate programs exist for the Mathematics, Physics and Chemistry Olympiads. In recent years, programs have been created for Biology and Computer Science
Olympiads as well. Distinct studies have been undertaken in each of these academic areas. In the Mathematics, Physics and Chemistry Olympiad programs, a series of increasingly difficult tests are administered. This testing concludes with the identification of the top national finalists (6-20 Olympians). These individuals are trained to compete in the International Olympiad programs.

The second group, “Prefinalists,” involved in this study consists of secondary school students who have taken part in the national competitions in mathematics during the years 2000 and 2001. Each year schools all over the Finland send their most talented students to this annual competition. The tests of this competition resemble the tests used in academic Olympians.

The third group, “Polytechnics,” are students of Espoo - Vantaa Institute of Technology. They need progressively advanced mathematical skills as they progress in their studies. However, compared to higher-level mathematics studies in universities, technological mathematics studies in vocational high schools are more practically oriented.

In addition, respondents’ parents were asked about their educational level. Over sixty per cent of Olympians (62%) and Prefinalists (65%) parents had an academic degree. Only 24 per cent of Polytechnics parents had the same educational level. Analysis of parental occupation in the three groups showed that Olympians and Prefinalists parents shared similar vocational interests as they were, for example, doctors, teachers and business managers. Polytechnics’ parents were mainly middle class; for example, factory workers. However, in Finland educational level is not a good predictor of socio-economic status. Prefinalists’ parents belong to the highest income class in Finland with the average annual salary of USD 83,597. Both Olympians (USD 49,447) and Polytechnics parents (USD 46,721) earn middle level salaries in the Finnish context.

Procedure
All the participants completed the Self-confidence attitude attribute Scales (SaaS) questionnaire (Campbell, 1996a) based on Weiner’s self-attribution theory (1974). The Mathematics Olympians’ data ($n = 77$) included 68 male and nine female respondents. The sample is quite representative as the total number of Finnish Mathematics Olympians is 84 (14 females and 70 males). The second group ($n = 52$) is a sample from about 200 secondary school national competitors in mathematics. The polytechnic student data ($n = 74$) is a fully representative sample of an advanced mathematics course held at Espoo - Vantaa Institute of Technology in autumn 2001, the total number of participants in which is approximately 3000. Olympiad data was collected between 1997 and 2002, the Polytechnic data in 2001, and the Prefinalists data between 2001 and 2002.
Investigating the Influence of Attribution Styles...

Table 1 shows, except for the Polytechnics, that the gender is biased towards males. This finding is related to the well-documented tendency of females not to pursue careers in Sciences, even though they are equally capable as males (e.g., Enman & Lupart, 2000), unless one or both of their parents are in the same field (Tirri, 2002). The student’s age is a good predictor for group membership.

-- Insert Table 1 about here --

**Measurement Instrument**

The SaaS instrument was mailed to respondents in a traditional paper and pen form (see Table 2). The instrument used a six-point Likert scale ranging from 1 (totally disagree) to 6 (totally agree). The SaaS scale included eighteen items measuring the students’ attributions based on self-attribution theory (Weiner, 1974). Although Weiner’s original conceptualization contained four attributions (ability, effort, difficulty, luck), the statistical analysis based on numerous empirical samples produced only two distinct scales, effort and ability (Campbell, 1996a, 1996b; Feng et al., 2001; Heller & Lengfelder, 2000; Tirri, 2001). In each of these studies a consistent factor structure was found for the ability and effort factors. Statements linking success and effort produced high scores on the effort factor. Statements on the ability factor expressed the view that ability is more important than hard work.

In addition to SaaS, we asked for the following background information from the respondents: gender, age, number of programming languages known, and average of mathematics, physics and chemistry secondary school grades.

-- Insert Table 2 about here --

**Statistical Analyses**

The data analysis began by examining all the items to see if they were technically applicable for linear statistic computations based on multivariate normality assumptions, such as exploratory factor analysis (EFA) and multivariate analysis of variance (MANOVA).

In the second stage we conducted an explorative factor analysis to answer the first research question: 1) “Are the four dimensions of the SaaS instrument (success due to ability, failure due to lack of ability, success due to effort, failure due to lack of effort) identified in this domain?”

Bayesian classification modeling (Silander & Tirri, 1999, 2000) was conducted in the third stage of the analysis to answer the second research question: 2) “What are the best predictors for the level of mathematical giftedness (highly = “Olympians,” moderately = ”Prefinalists” and mildly = “Polytechnics”) and gender among the SaaS variables?” Bayesian classification modeling resembles
linear discriminant analysis (Huberty, 1994), but it is free of most of the assumptions of Gaussian modeling (Nokelainen, Ruohotie & Tirri, 1999, p. 113; Nokelainen & Tirri, 2004).

In the fourth stage we conducted MANOVA (with the Roy-Bargman stepdown analysis and Bonferroni post hoc test) to see if the attribution styles differ by the level of mathematical giftedness or gender. When investigating more than one dependent variable, we applied factorial MANOVA instead of series of ANOVA’s, as it controls for increasing risk of Type I error (falsely rejecting null hypothesis when it is true).

Results

Investigating Variables Statistical Properties

In the first stage, a frequency analysis was carried out for all the variables. Results show that the respondents used the whole scale from 1 (totally disagree) to 6 (totally agree) for all the items. According to Kerlinger (1986, p. 495), before constructing one’s own questionnaire “… one should first ask the question: Is there a better way to measure my variables?” He classifies weaknesses of rating scales into extrinsic and intrinsic. The extrinsic defect is that scales are much too easy to construct and use. Sometimes a scale is used to measure things for which it is not appropriate. Kerlinger defines the intrinsic defect of rating scales as their proneness to constant error. He lists four main sources: Halo effect, the error of severity (to rate all items too low), error of leniency (to rate all items too high) and error of central tendency (to avoid all extreme judgments). To address this issue, we analyzed the overall response tendency. We found that distribution of the modes on a six-point Likert scale was multimodal and slightly biased towards positive values: 1) \(n = 0\); 2) \(n = 5\); 3) \(n = 0\), 4) \(n = 11\); 5) \(n = 0\); 6) \(n = 0\).

Numerous publications declare that certain attributes belong to data that is appropriate for multivariate analysis (e.g., Bradley & Schaefer, 1998, pp. 79-83; Tabachnick & Fidell, 1996, pp. 13-17). The most commonly used criteria for accepting variables for multivariate analysis are: 1) Standard deviation of no more than half the mean; 2) Skewness less than +/- .3 and 3) Correlation between +/- .3 - .7. When we examined the eighteen items using the first two criteria, we noticed that all the items passed the first criteria, but only four items passed the second criteria. As it seemed impossible to take the second criteria literally due to a high rejection rate at the .3 level, we examined the skewness of items in three additional levels (.5, .7, .8). The .7 level proved to be suitable for this data set suggesting rejection of three items (4, 6, and 7). A non-parametric inter-item correlation matrix was produced in order to examine the third criteria. Thirteen items reached the desired level as the values ranged from –
Investigating the Influence of Attribution Styles ... 13

0.48 to 0.71 (\(M = 0.05, SD = 0.16\)). The rejected items were: 4, 7, 10, 14, and 18. We examined multivariate normality with Mahalanobis distances. The maximum values for the two SaaS scales were below critical values obtained from the Chi-square table (alpha = .001), thus not indicating the presence of outliers.

Finally, when we combine the results of the variable selection phase, it seems obvious that at least items “4. I worked harder if I liked the teacher,” and “7. You have to have the ability in order to succeed in most things” should be omitted from further analysis.

**Explorative Factor Analysis**

Our next task, according to the first research question, was to see if the combined sample and three sub samples contained the following four dimensions: 1) Success due to ability; 2) Failure due to lack of ability; 3) Success due to effort; 4) Failure due to lack of effort. We performed the analysis with sixteen items, as the variable rejection based on communalities of two-dimensional PCA structure did not appear to provide a feasible solution. Factor analysis with the Maximum Likelihood extraction method and Direct Oblimin rotation (delta value was set to zero, i.e., letting factors correlate) was conducted for the combined sample (\(n = 203\)) and for each sample separately (Olympians, \(n = 77\); Prefinalists, \(n = 52\); Polytechnics, \(n = 74\)).

A four-factor solution with eight items grouped the variables in all three sub samples and the combined sample as expected. Next we present the eight items operationalizing four SaaS factors. Factor 1, success due to ability, included only one variable, “5. Being smart is more important than working hard”. The logic behind this solution was that all the other two related items (“4. I worked harder if I liked the teacher” and “7. You have to have the ability in order to succeed in most things”) were omitted from further analysis because they did not meet the assumptions of multivariate analysis.

Factor 2, failure due to a lack of ability, included items “3. There are some things you can not do no matter how hard you try” and “13. When I did poorly in school it was because I did not have the needed ability” (alpha = .62). Factor 3, success due to effort, included items “9. Self-discipline is the key to school success,” “12. I had to work hard to get good grades” and “17. Hard work is the key to get good grades” (alpha = .63). Factor 4, failure due to a lack of effort, included items “8. My achievement would have been better if I tried harder” and “16. I could have done better in mathematics if I have tried harder” (alpha = .82). The Cronbach’s alpha values for the four factors within the three groups varied as follows: Olympist data (factor 1 = not calculated, factor 2 = .56, factor 3 = .75, factor 4 = .76), Prefinalists data (factor 1 = not calculated, factor 2 = .60, factor 3 = .67, factor 4 = .84) and Polytechnics data (factor 1 = not calculated, factor 2 = .59, factor 3 = .42, factor 4 = .80).
Although we found only one item measuring the first SaaS dimension, success due to ability, correlations between factors behaved as expected (Table 3). Ability and effort factors correlated negatively with each other and both effort factors, as well as both ability factors, correlated positively.

Bayesian Classification Modeling

We conducted the Bayesian classification modeling with the B-Course program (Myllymäki, Silander, Tirri & Uronen, 2002) to find out which variables measuring attribution styles are the best predictors for the level of mathematical giftedness (highly = “Olympians”, moderately = ”Prefinalists” and mildly = “Polytechnics”) and gender (research question 2). In the classification process, the automatic search
tried to find the best set of variables to predict the class variable for each data item. This procedure resembles the traditional linear discriminant analysis (LDA, Huberty, 1994, pp. 118-126), but the implementation is totally different. For example, a variable selection problem that is addressed with forward, backward or stepwise selection procedure in LDA is replaced with a genetic algorithm approach (e.g., Hilario, Kalousisa, Pradosa & Binzb, 2004; Hsu, 2004) in the Bayesian classification modeling. The genetic algorithm approach means that variable selection is not limited to one (or two or three) specific approach; instead many approaches and their combinations are exploited. One possible approach is to begin with the presumption that the models (i.e., possible predictor variable combinations) that resemble each other a lot (i.e., have almost same variables and discretizations) are likely to be almost equally good. This leads to a search strategy in which models that resemble the current best model are selected for comparison, instead of picking models randomly. Another approach is to abandon the habit of always rejecting the weakest model and instead collect a set of relatively good models. The next step is to combine the best parts of these models so that the resulting combined model is better than any of the original models. B-Course is capable of mobilizing many more viable approaches, for example, rejecting the better model (algorithms like hill climbing, simulated annealing) or trying to avoid picking similar model twice (tabu search).

First, we derived the model for classifying data items according to the class variable level of mathematical giftedness (Olympians, Prefinalists and Polytechnics) with 18 variables of the SaaS scale as predictors (items are listed in Table 2). The estimated classification accuracy for the model was 65 per cent. Second, we derived the model for classifying data items according to the class variable gender. The estimated classification accuracy for the model was 80 per cent.

Table 5 lists the variables ordered by their estimated classification performance in the model. The strongest variables—that is, those that discriminate the independent variables best—are listed first. The percentage value attached to each variable indicates the predicted decrease in the classification performance if the variable were to be dropped from the model. The table shows that the variables in the first two models, level of giftedness and gender, have a clear order of importance. The most important variable for both models is ”12. I had to work hard to get good grades.” If we remove that variable from the first model, it would weaken the performance from 65 per cent to 58 per cent. Removal of the variable from the second model would weaken the performance from 80 per cent to 73 per cent.

Differences in the group means in Table 5 show that the first classification between three groups of mathematically gifted is based on effort attributions as five items out seven measure success or failure due to effort. The first item, “10. The smart kids tried the hardest”, is the best overall predictor variable.
However, the other items are more interesting as they show that both mathematically mildly (Polytechnics) and moderately (Prefinalists) gifted individuals attribute failure to lack of effort, but only mathematically mildly gifted individuals attribute success to effort. Mathematically highly (Olympians) and moderately (Prefinalists) gifted individuals prefer ability as an explanation for their success.

Females in this sample tend to attribute success to effort more than males. Further, they are also more likely to attribute failure to lack of ability than males. Both findings are consistent with existing research (e.g., Alderman, 2004, p. 42; Vermeer, Boekaerts & Seegers, 2000). However, we note that ‘female voice’ in this study belongs to mostly those who are mathematically mildly gifted as they are members of the Polytechnics group. This explains at least to some extent why items measuring effort have such important role in the first two classification models. (Table 5.)

-- Insert Table 5 about here --

**Multivariate Analysis of Variance**

We investigated the third research question “Do the attribution styles differ by the level of mathematical giftedness (highly = “Olympians,” moderately = ”Prefinalists” and mildly = “Polytechnics”) and gender?” with a 3 × 2 factorial multivariate analysis of variance. The four dependent variables were SaaS factors based on both theoretical assumptions (Weiner, 1974) and the results of preceding EFA: Success due to ability, failure due to lack of ability, success due to effort, and failure due to lack of effort. The independent variables were the level of mathematical giftedness and gender. Preliminary assumption testing was conducted to check for normality, linearity, univariate and multivariate outliers, homogeneity of variance-covariance matrices, and multicollinearity. No violations were discovered except for fourth factor, failure due to lack of effort for which the test of homogeneity of variance was not met (Levene’s $p < .001$, Cochran’s $p = .014$, Bartlett-Box’s $p = .002$). Larger variances indicate that a .05 alpha level is overstated and the differences should be assessed using a lower value, for example, .03 (Hair, Anderson, Tatham & Black, 1998, p. 349). For such dependent variables, Tabachnick and Fidell (1996, 382) suggest using Pillai’s criterion instead of Wilks’ lambda.

With the use of Pillai’s criterion, the level of mathematical giftedness multivariate main effect on the SaaS factors was found to be significant, $F(8, 384) = 4.33, p < .001$. The gender multivariate main effect on the SaaS factors was not found to be significant, $F(4, 191) = 0.23, p = .992$. The level of mathematical giftedness and gender multivariate interaction was not found to be significant, $F(8, 384)$
Investigating the Influence of Attribution Styles... 17

= 0.45, \( p = .893 \). The results reflected a modest association between three groups of mathematically gifted and the SaaS factors, partial \( \eta^2 = .08 \) (Pillai’s trace) - .16 (Roy’s Largest Root). This finding suggests that group membership explains attribution styles from eight to 16 per cent. The achieved statistical power for this main effect was 1.0.

To further investigate the impact of the level of mathematical giftedness main effect on the SaaS, a Roy-Bargman stepdown analysis was performed. Stepdown analysis resolves the problem of correlated univariate \( F \) tests with correlated DV’s (Tabachnick & Fidell, 1996, p. 403). The following priority order of SaaS factors from most to least important was developed on the basis of theoretical assumptions, instead of assigning priority on the basis of univariate \( F \), to avoid problems inherent in stepwise regression: (1) success due to ability, (2) failure due to lack of ability, (3) success due to effort, and (4) failure due to lack of effort. In stepdown analysis each SaaS factor was analyzed, in turn, with higher-priority factors treated as covariates and with the highest-priority SaaS factor (success due to ability) tested in a univariate ANOVA. In addition, univariate \( F \) values were calculated to allow correct interpretation of the stepdown analysis. Results of the analysis are summarized in Table 6. An experimentwise error rate of five per cent was achieved by the apportionment of alpha as shown in the last column of Table 6 for each of the SaaS factors.

A unique contribution to predicting differences between the three groups of mathematically gifted was made by success due to ability factor, stepdown \( F(2, 194) = 3.28, p = .040, \eta^2 = .03 \). According to Cohen (1988), the effect size is small, indicating that only three per cent of variance in the dependent variable is attributable to differences in mathematical giftedness. Mathematically highly (mean success due to ability = 3.45, \( SE = .07 \)) and moderately (mean success due to ability = 3.56, \( SE = .09 \)) gifted individuals evaluated the role of ability to be higher than mathematically mildly gifted did (mean success due to ability = 3.28, \( SE = .07 \)). The mean difference between the moderately and mildly gifted reached statistical significance using Bonferroni adjusted alpha level of .013, \( p = .040 \). This result is consistent with both Weiner’s (1986) findings showing high-achieving students’ tendency to use ‘internal-stable-uncontrollable’ causal attributions for success, and Heller and Lengfelder’s (2000) findings for German Olympians and Prefinalists. However, studies of American and Taiwanese Olympians showed opposite results as participants referred more to effort than ability attributions (Campbell, 1996b; Feng, Campbell & Verna, 2001; Wu & Chen, 2001).

After the pattern of differences measured by the first SaaS factor was entered, a difference was also found on failure due to lack of ability, stepdown \( F(2, 193) = 11.20, p < .001, \eta^2 = .10 \). The effect size of this finding is moderate, indicating that ten per cent of variance in the dependent variable is
Investigating the Influence of Attribution Styles... attributable to differences in mathematical giftedness. Mathematically highly gifted (adjusted mean failure due to lack of ability = 3.00, $SE = .09$) students attributed failure to lack of ability more than moderately (adjusted mean failure due to lack of ability = 2.60, $SE = .11$) and mildly (adjusted mean failure due to lack of ability = 2.38, $SE = .09$) gifted. The mean difference between mathematically highly and moderately gifted students reached statistical significance using Bonferroni adjusted alpha level of .013, $p = .028$. Also, the mean difference between the mathematically highly and mildly gifted reached statistical significance using Bonferroni adjusted alpha level of .013, $p < .001$. This ‘internal-stable-uncontrollable’ finding, which according to Marsh (1983) may lead to low academic self-concept and achievement, is against Weiner’s ‘internal-unstable-controllable’ expectation for high-achievers who have failed.

The third step in the analysis was to enter the success due to effort factor. This stepdown reached statistical significance, however, with a small effect size, $F(2, 192) = 3.05$, $p < .05$, $\eta^2 = .03$. Mathematically mildly gifted individuals valued success due to effort higher (adjusted mean success due to effort = 3.04, $SE = .09$) than those who were highly (adjusted mean success due to effort = 2.74, $SE = .09$) and moderately gifted (adjusted mean success due to effort = 2.79, $SE = .10$). The mean difference between mildly and highly gifted students reached statistical significance using Bonferroni adjusted alpha level of .013, $p = .029$. This result showing mathematically mildly gifted Polytechnics preferring ‘internal-unstable-controllable’ attributions was expected.

After the pattern of differences measured by success due to ability, failure due to lack of ability and success due to effort was entered, a difference was also found on the attitude toward failure due to lack of effort, stepdown $F(2, 191) = 11.43$, $p < .001$, $\eta^2 = .11$. According to Cohen (1988), the effect size for this finding is classified as large. Mathematically mildly (adjusted mean failure due to lack of effort = 3.99, $SE = .11$) and moderately (adjusted mean failure due to lack of effort = 3.77, $SE = .12$) gifted students attributed failure to lack of effort more than highly gifted did (adjusted mean failure due to lack of effort = 3.27, $SE = .11$). The mean difference between mildly gifted Polytechnics and highly gifted Olympians reached statistical significance using Bonferroni adjusted alpha level of .013, $p < .001$. In addition, the mean difference between moderately gifted Prefinalists and highly gifted Olympians reached statistical significance using Bonferroni adjusted alpha level of .013, $p < .001$. The group level results of this second SaaS failure factor were not congruent with theoretical expectations. We expected to see mildly gifted individuals prefer ‘internal-stable-uncontrollable’ attributions; that is, failure due to lack of ability, instead of ‘internal-unstable-controllable’.
Summary of Results

In this paper we have examined the influence of attribution styles on the development of mathematical talent in three groups of mathematically gifted Finnish adolescents and adults (n = 203).

All the participants completed the Self-confidence attitude attribute Scales (SaaS) questionnaire (Campbell, 1996a). The instrument included 18 items measuring the students’ ability and effort attributions, based on Weiner’s attribution theory (1974), on four dimensions: (1) success due to ability, (2) failure due to lack of ability, (3) success due to effort, and (4) failure due to lack of effort.

The research questions in this study were as follows: “(1) Are the four dimensions of the SaaS instrument (success due to ability, failure due to lack of ability, success due to effort, failure due to lack of effort) identified in this domain?”, “(2) What are the best predictors for the level of mathematical giftedness and gender among the SaaS scale variables?” and “(3) Do the attribution styles differ by the level of mathematical giftedness (Olympians, Prefinalists, Polytechnics) or gender?”

The first research question was addressed with EFA. The results showed that the four dimensions of the SaaS were present in all samples. The overall alpha values ranged from .62 to .82. Three alpha values were below .60 level in the group level (Olympians, factor 2, $\alpha = .56$; Polytechnics, factor 2, $\alpha = .59$ and factor 3, $\alpha = .42$).

The second research question was analyzed using Bayesian classification modeling. The classification variables were the level of mathematical giftedness (highly = “Olympians”, moderately = “Prefinalists” and mildly = “Polytechnics”) and gender. Eighteen SaaS items were predictors in all the analyses. The results showed that both Polytechnic students and females think that they “had to work hard to get good grades.” When we further examined the female’s preference for effort as a cause for success, we learned that the result was true only for the Polytechnics and Prefinalists samples as there was no difference between female and male Olympians’ responses. Thus Verna and Campbell’s (2000) earlier finding that female Chemistry Olympians considered ability to be a more important factor for success was not repeated in this study. Failure due to lack of ability was the only self-attribute scale that was able to predict respondent’s age. The youngest students (15-28 years) believed more in their abilities than the older ones (29-41 and 42-55 years). We explain this finding with the fact that the younger individuals have not yet reached as high a level in their mathematical studies as the older ones, and thus realized that ‘the more you know, the more you know you ought to know.’
The third research question was analyzed with $3 \times 2$ factorial design MANOVA. Dependent variables were four SaaS factors (success due to ability, failure due to lack of ability, success due to effort, failure due to lack of effort). Independent variables were the level of mathematical giftedness (highly = “Olympians,” moderately =”Prefinalists” and mildly = “Polytechnics”) and gender. Results showed that the level of mathematical giftedness multivariate main effect on the SaaS factors was found to be significant. The gender multivariate main effect on the SaaS factors and the level of mathematical giftedness and gender multivariate interaction were not found to be significant. Highly and moderately mathematically gifted individuals felt that ability is more important to success than effort. According to Dai et al. (1998), such attributions represent self-awareness of high potentialities that constitute a necessary but not sufficient condition for high levels of performance. Mildly mathematically gifted individuals tended to see effort leading to success. Mildly and moderately mathematically gifted students attributed failure to lack of effort. The highly gifted attributed failure to lack of ability. This finding is related to self-concept, as mildly and moderately mathematically gifted individuals tend to judge their efficacy favorably, whereas the highly gifted are likely to base their appraisals of self-efficacy on the actual difficulty levels of the tasks in question (see Dai et al., 1998, for discussion).

Discussion

The theoretical idea of Ellström (2001) essential to this study is that attributions for success or failure affect potential competence, which is a human resource each individual brings to the mathematical problem-solving situation. So, what are the ‘good’ attributions that give rise to individual’s potential competence? The previous research body shows two trends: The first is seeing ability as a more important explanation for success than effort, ‘ability is everything’ (e.g., Heller & Lengfelder, 2000), and another is claiming that ‘ability without effort goes nowhere’ (e.g., Campbell, 1996b; Chan, 1996; Feng, Campbell & Verna, 2001; Wu & Chen, 2001).

Why do several international Academic Olympian research teams end up with contradictory results? The first natural explanation is cultural differences (Campbell, Tirri, Ruohotie & Walberg, 2004). We suggest here that the second natural explanation is statistical. In all of the aforementioned studies, effort and ability attributions were calculated as mean scores of 18 SaaS items: the effort scale was measured with 12 items and the ability scale with six items (see Table 2). The results of variable selection and EFA indicated with the Finnish sample that keeping all 18 SaaS items on the model does not lead to a psychometrically justified solution ($\alpha_{\text{effort}} = .63$ and $\alpha_{\text{ability}} = .27$). We forced the two factor
solution and calculated internationally comparable mean scores for the Finnish Olympian sample \( n = 77 \), but found no difference between ability \( (M = 3.16, SD = 0.32) \) and effort attributions \( (M = 3.18, SD = 0.45) \). It should also be noted that the reported mean difference between ability \( (M = 2.91, SD = 0.58) \) and effort \( (M = 3.21, SD = 0.62) \) scales in the American study (Campbell, 1996b) is quite small.

The third trend in research (e.g., Schunk & Ertmer, 2000; Zimmermann, 2000) says that ability and effort without self-regulation goes nowhere. Figure 1 shows that ‘ability’ is present in the “Self-Reflection” phase that precedes the “Forethought” phase where individuals set goals and plan actions. At this point it is important to make a distinction between an individuals’ real and self-perceived ability level. The real ability level is what the first research trend is speaking about, but according to the concept of self-regulation, it is undistinguishable from the self-perceived ability. This notion comes from the fact that effort is represented in the preceding phase “Performance or Volitional Control” in the cyclical self-regulation process described in Figure 1 via self-control. It is, therefore, sending adjusting signals via self-observation to individual’s self-perceived ability. Our result that shows highly mathematically gifted preferring ability over effort as an explanation for success or failure might indicate that their perceived ability level is so high that they are able to meet the most demanding challenges by adjusting their efforts. According to “attribution asymmetry” phenomenon (Dai et al., 1998), high-ability students tend to attribute their success to both ability and effort. According to them, attributing success to ability represent self-awareness of high potentialities that constitute a necessary but not sufficient condition for success. Attributing success to effort has a self-enhancing and motivating effect as one feels in control of one’s own development.

It is not our primary purpose to compare ability and effort only with questions that ask which is a better explanation for success, but rather to continue with further questions such as “What are the best developmental practices for those individuals who mostly prefer effort over ability as an explanation for success in a given task?” As Campbell (1995, p. 186) says: “Achievers need four qualities: ability, discipline, confidence and good working habits.” Thus, high effort level as a cause for success may indicate that tasks are too demanding and thus individuals feel that too much effort is needed to accomplish the task. Further, this might indicate that an individual needs more support to be convinced that he or she has the ability to succeed. Research has shown that ability is often viewed as a stable and uncontrollable attribution (e.g., SaaS item “7. You have to have the ability in order to succeed in most things“). The worst scenario for a mathematician according to some of the self-attribution theorists (e.g., Dweck, 1999) is to blame personal ability for failure, as it is believed to be something that you either do have or not, an internal and stable attribution. Interestingly the most highly mathematically gifted in this study scored highest in this respect when compared to those who were moderately and
mildly gifted. This finding makes sense if we agree that Olympians represent the highest mathematical giftedness level in this study and, thus, have probably faced a lot more demanding mathematical tasks during their lifetime than the other two group members.

Kay Alderman suggests that it is up to the teacher or trainer to convince a learner or trainee that mathematical thinking ability as a skill or knowledge is learnable, unstable quality (2004, 31). Thus, knowledge of how learners or trainees use attributions to account for success and failure can help teachers or trainers predict their expectancies and plan intervention strategies when needed.

**Limitations of the Study**

In this study, we measured attribution styles with a questionnaire. Such self-reporting allows us to study, as opposed to attribution appraisals or causal beliefs, hypothetical success and failure situations without clear reference to who is the performer. The first possible source of error is the SaaS instrument translation from English to Finnish. To control the error variance due to translation, all the items were re-translated back to English and compared with the original items. However, no pilot study with correlational analyses was conducted. A second possible source of error is cross-cultural differences between the U.S. and Finnish mathematicians since the original instrument was developed for studies among the U.S. mathematics Olympians. Fortunately, the language of the SaaS items is free of cultural references. A third possible source of error is the psychometric properties of the SaaS instrument itself as no alpha values were reported in the original study (Campbell, 1996a).

**References**


Investigating the Influence of Attribution Styles...


Figure 1. Cyclical Self-regulatory Phases (adapted from Zimmermann, 2000).
Table 1

Description of Mathematics Olympians, Prefinalists and Polytechnics Data

<table>
<thead>
<tr>
<th></th>
<th>Olympians (n = 77)</th>
<th>Prefinalists (n = 52)</th>
<th>Polytechnics (n = 74)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>Gender</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Male</td>
<td>68</td>
<td>88%</td>
<td>43</td>
</tr>
<tr>
<td>Female</td>
<td>9</td>
<td>12%</td>
<td>9</td>
</tr>
<tr>
<td>Age</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Median</td>
<td>37 years</td>
<td></td>
<td>17 years</td>
</tr>
<tr>
<td>Range</td>
<td>20 – 55 years</td>
<td></td>
<td>15 – 20 years</td>
</tr>
</tbody>
</table>

Note. Total N of the data is 203.
Table 2
Descriptive Statistics and Attribution Dimensions of the Self-confidence Attribute Attitude Scale

<table>
<thead>
<tr>
<th>Item</th>
<th>Olympians (n = 77)</th>
<th>Prefinalists (n = 52)</th>
<th>Polytechnics (n = 74)</th>
<th>Stable – Unstable a</th>
<th>Controllable – Uncontrollable b</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Effort (12 items)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. I did poorly only when I did not work hard enough.</td>
<td>3.79 0.97</td>
<td>3.56 1.15</td>
<td>3.46 1.15</td>
<td>U</td>
<td>C</td>
</tr>
<tr>
<td>2. You can be successful in anything if you work hard enough at it.</td>
<td>3.27 1.23</td>
<td>3.78 1.03</td>
<td>3.99 0.85</td>
<td>U</td>
<td>C</td>
</tr>
<tr>
<td>6. When I scored low on a test, it was because I didn’t study hard enough.</td>
<td>3.92 0.79</td>
<td>3.63 0.97</td>
<td>3.53 0.95</td>
<td>U</td>
<td>C</td>
</tr>
<tr>
<td>8. My achievement would have been better if I tried harder.</td>
<td>3.29 1.25</td>
<td>3.83 0.96</td>
<td>4.09 0.69</td>
<td>U</td>
<td>C</td>
</tr>
<tr>
<td>9. Self-discipline is the key to school success.</td>
<td>3.27 0.91</td>
<td>3.47 0.92</td>
<td>3.41 1.06</td>
<td>S</td>
<td>C</td>
</tr>
<tr>
<td>10. The smart kids tried the hardest.</td>
<td>2.43 0.93</td>
<td>2.81 0.95</td>
<td>2.22 0.93</td>
<td>S</td>
<td>C</td>
</tr>
<tr>
<td>11. Poor study habits are the main cause of low grades.</td>
<td>3.32 0.91</td>
<td>3.44 0.98</td>
<td>3.54 1.04</td>
<td>U</td>
<td>C</td>
</tr>
<tr>
<td>12. I had to work hard to get good grades.</td>
<td>2.11 0.95</td>
<td>2.08 0.95</td>
<td>2.74 1.03</td>
<td>S</td>
<td>C</td>
</tr>
<tr>
<td>15. When I didn’t understand something, it meant I didn’t put in enough time.</td>
<td>3.69 0.87</td>
<td>3.58 1.07</td>
<td>3.51 0.94</td>
<td>U</td>
<td>C</td>
</tr>
<tr>
<td>16. I could have done better in mathematics if I had worked harder.</td>
<td>3.11 1.17</td>
<td>3.81 1.05</td>
<td>4.01 0.85</td>
<td>U</td>
<td>C</td>
</tr>
<tr>
<td>17. Hard work is the key to get good grades.</td>
<td>2.83 1.05</td>
<td>2.77 1.13</td>
<td>3.00 0.89</td>
<td>S</td>
<td>C</td>
</tr>
<tr>
<td>18. I let people down when I don’t work hard enough.</td>
<td>2.64 1.05</td>
<td>2.62 1.14</td>
<td>2.45 1.09</td>
<td>U</td>
<td>C</td>
</tr>
<tr>
<td><strong>Ability (6 items)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. There are some things you can not do no matter how hard you try.</td>
<td>3.68 1.19</td>
<td>3.29 1.17</td>
<td>2.97 1.19</td>
<td>S</td>
<td>U</td>
</tr>
<tr>
<td>4. I worked harder if I liked the teacher.</td>
<td>3.16 1.22</td>
<td>3.57 1.24</td>
<td>3.77 0.99</td>
<td>U</td>
<td>C</td>
</tr>
<tr>
<td>5. Being smart is more important than working hard.</td>
<td>2.99 1.03</td>
<td>3.23 0.92</td>
<td>2.64 1.03</td>
<td>S</td>
<td>U</td>
</tr>
<tr>
<td>7. You have to have the ability in order to succeed in most things.</td>
<td>3.92 0.74</td>
<td>3.90 0.69</td>
<td>3.86 0.75</td>
<td>S</td>
<td>U</td>
</tr>
<tr>
<td>13. When I did poorly in school it was because I did not have the needed ability.</td>
<td>2.46 1.04</td>
<td>2.37 0.98</td>
<td>2.09 0.80</td>
<td>S</td>
<td>U</td>
</tr>
<tr>
<td>14. Why work in an area where your ability is low?</td>
<td>2.81 1.03</td>
<td>2.75 1.23</td>
<td>2.58 1.16</td>
<td>S</td>
<td>U</td>
</tr>
</tbody>
</table>

a Stable and unstable attributions refer to consistent or inconsistent cause over time. b Controllable and uncontrollable attributions refer to extent person believes he or she has control over the cause of an outcome.
Table 3
*Correlations Between the Self-confidence Attribute Attitude Scale Factors*

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Success due to ability (item 5)</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Failure due to lack of ability (items 3, 13)</td>
<td>.135</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Success due to effort (items 9, 12, 17)</td>
<td>-.194**</td>
<td>-.052</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>4. Failure due to lack of effort (items 8, 16)</td>
<td>.151*</td>
<td>-.286**</td>
<td>.171*</td>
<td>1.000</td>
</tr>
</tbody>
</table>

*Note. Item descriptions are in the Table 2.*

* p < .05. ** p < .01. (Two-tailed.)
Table 4  
*Self-confidence Attribute Attitude Scale Factors by Attribution Dimensions*

<table>
<thead>
<tr>
<th></th>
<th>Controllable</th>
<th>Internal</th>
<th>Uncontrollable</th>
<th>Internal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Items Factors</td>
<td></td>
<td>Items Factors</td>
<td></td>
</tr>
<tr>
<td>Stable</td>
<td>9, 12, 17</td>
<td>Success due to effort</td>
<td>5</td>
<td>Success due to ability</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3, 13</td>
<td>Failure due to lack of ability</td>
</tr>
<tr>
<td>Unstable</td>
<td>8, 16</td>
<td>Failure due to lack of effort</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 5
Importance Ranking of the Self-confidence Attribute Attitude Scale Items by the Level of Giftedness and Gender

<table>
<thead>
<tr>
<th>Class and Predictor variables</th>
<th>Drop c</th>
<th>The Level of Giftedness</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td>Olympians (n = 77)</td>
<td>Prefinalists (n = 52)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M  SD</td>
<td>M  SD</td>
</tr>
<tr>
<td>The Level of Giftedness a</td>
<td>10. The smart kids tried the hardest.</td>
<td>14.36 2.43 0.93</td>
<td>2.81 0.95</td>
</tr>
<tr>
<td></td>
<td>16. I could have done better in mathematics if I had worked harder.</td>
<td>7.92 3.11 1.17</td>
<td>3.81 1.05</td>
</tr>
<tr>
<td></td>
<td>12. I had to work hard to get good grades.</td>
<td>6.93 2.11 0.95</td>
<td>2.08 0.95</td>
</tr>
<tr>
<td></td>
<td>5. Being smart is more important than working hard.</td>
<td>4.95 2.99 1.03</td>
<td>3.23 0.92</td>
</tr>
<tr>
<td></td>
<td>3. There are some things you can not do no matter how hard you try.</td>
<td>3.96 3.68 1.19</td>
<td>3.29 1.17</td>
</tr>
<tr>
<td></td>
<td>4. I worked harder if I liked the teacher.</td>
<td>2.48 3.16 1.22</td>
<td>3.57 1.24</td>
</tr>
<tr>
<td></td>
<td>8. My achievement would have been better if I tried harder.</td>
<td>1.98 3.29 1.25</td>
<td>3.83 0.96</td>
</tr>
<tr>
<td>Gender b</td>
<td>12. I had to work hard to get good grades.</td>
<td>6.93 2.11 0.95</td>
<td>2.08 0.95</td>
</tr>
<tr>
<td></td>
<td>8. My achievement would have been better if I tried harder.</td>
<td>6.93 2.11 0.95</td>
<td>2.08 0.95</td>
</tr>
<tr>
<td></td>
<td>2. You can be successful in anything if you work hard enough at it.</td>
<td>3.96 3.68 1.19</td>
<td>3.29 1.17</td>
</tr>
<tr>
<td></td>
<td>1. I did poorly only when I did not work hard enough.</td>
<td>3.47 3.68 1.19</td>
<td>3.29 1.17</td>
</tr>
<tr>
<td></td>
<td>14. Why work in an area where your ability is low?</td>
<td>2.97 3.68 1.19</td>
<td>3.29 1.17</td>
</tr>
<tr>
<td></td>
<td>5. Being smart is more important than working hard.</td>
<td>1.98 3.68 1.19</td>
<td>3.29 1.17</td>
</tr>
</tbody>
</table>

aClassification accuracy is 65 per cent. bClassification accuracy is 80 per cent. cDecrease in predictive classification if item is dropped from the classification model.
Table 6
Tests of the Level of Giftedness, Gender, and Their Interaction

<table>
<thead>
<tr>
<th>IV</th>
<th>DV</th>
<th>Univariate F</th>
<th>df</th>
<th>Stepdown F&lt;sup&gt;a&lt;/sup&gt;</th>
<th>df</th>
<th>α</th>
<th>η&lt;sup&gt;2&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Level of Giftedness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUC_ABI</td>
<td>3.28*</td>
<td>2,194</td>
<td>3.28*</td>
<td>2,194</td>
<td>.01</td>
<td>.03</td>
<td></td>
</tr>
<tr>
<td>FAI_ABI</td>
<td>11.99***</td>
<td>2,194</td>
<td>11.20***</td>
<td>2,194</td>
<td>.01</td>
<td>.10</td>
<td></td>
</tr>
<tr>
<td>SUC_EFF</td>
<td>4.12*</td>
<td>2,194</td>
<td>3.05*</td>
<td>2,192</td>
<td>.01</td>
<td>.03</td>
<td></td>
</tr>
<tr>
<td>FAI_EFF</td>
<td>17.80***</td>
<td>2,194</td>
<td>11.43***</td>
<td>2,191</td>
<td>.01</td>
<td>.11</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUC_ABI</td>
<td>0.00</td>
<td>1,194</td>
<td>0.00</td>
<td>1,194</td>
<td>.01</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td>FAI_ABI</td>
<td>0.91</td>
<td>1,194</td>
<td>0.92</td>
<td>1,193</td>
<td>.01</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td>SUC_EFF</td>
<td>0.13</td>
<td>1,194</td>
<td>0.11</td>
<td>1,192</td>
<td>.01</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td>FAI_EFF</td>
<td>0.18</td>
<td>1,194</td>
<td>0.09</td>
<td>1,191</td>
<td>.01</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td>Group by gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUC_ABI</td>
<td>1.00</td>
<td>2,194</td>
<td>1.00</td>
<td>2,194</td>
<td>.01</td>
<td>.01</td>
<td></td>
</tr>
<tr>
<td>FAI_ABI</td>
<td>0.09</td>
<td>2,194</td>
<td>0.16</td>
<td>2,193</td>
<td>.01</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td>SUC_EFF</td>
<td>0.23</td>
<td>2,194</td>
<td>0.38</td>
<td>2,192</td>
<td>.01</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td>FAI_EFF</td>
<td>0.73</td>
<td>2,194</td>
<td>0.26</td>
<td>2,191</td>
<td>.01</td>
<td>.00</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Roy-Bargman stepdown F.<br><br>Note. α = Adjusted alpha level. η<sup>2</sup> = Effect size for the stepdown F. SUC_ABI = Success due to ability. FAI_ABI = Failure due to lack of ability. SUC_EFF = Success due to effort. FAI_EFF = Failure due to lack of effort.* p < .05. ** p < .01. *** p < .001.
CHAPTER 7

Modeling Students’ Motivational Profile for Learning in Vocational Higher Education

Petri Nokelainen and Pekka Ruohotie
University of Tampere

Introduction

The Abilities for Professional Learning (APL) questionnaire (Ruohotie, 2000b) was designed to measure 15 factors of motivation and learning strategies. This chapter examines the validity of six factors of the motivational part of the APL self-report instrument. The next chapter will focus on the development and validation of instruments to assess strategic skills in learning.

We agree with Boekaerts (2001, p. 18) that the problem with extensively used self-report questionnaires is that they do not provide insight into students’ personal motivation traits in actual classroom learning situations as they are in most cases developed for measuring traits of domain-specific motivation. However, there is a shift in the way the research methodology is applied in the context of motivational research, as there is growing number of moments when the students are not present simultaneously in the classroom where it is easiest to conduct the observation of general motivation traits. We refer here to computer-supported collaborative learning (CSCL) and other forms of distance learning exercised in the virtual university studies. Järvelä and Niemivirta (2001, 108) approach this issue by challenging the researchers within the framework of computer-supported collaboration and learning to focus on both learners’ intrinsic and extrinsic motivational processes. In a series of empirical case studies they approach this challenge with multi-faceted research methodology using questionnaires, videotaping of students’ learning activity, stimulated recall-interviews, and repeated process-oriented interviews (Järvelä & Niemivirta, 2001, pp. 115-122).

It is clear that in order to study complex phenomenon such as student motivation, both multi-faceted theoretical and methodological approach is required. Unfortunately, severe limitations for data collection exist in the context of virtual university learning.
Teacher or tutor has few face-to-face discussions with the students. In fact, as an expert in certain area, she has to deal each time with a new group of attendees’. Furthermore, study processes including collaborative actions, such as group works are conducted through the distance learning platforms.

We believe that as the volume in virtual studies increases, the personal face to face tutoring time for each student decreases and leads to situation where the guidance is conducted via synchronous (chat) or asynchronous (email) tools. In this hectic situation the teacher or tutor has to make good use of all the available information channels in order to understand the personal needs of each student. Nokelainen and Tirri (2002) discuss in their article how to utilize computerized self-rating instruments in the context of virtual university studies.

Theoretical Framework

Learner-centered Learning and Motivation

Learning in the virtual university is intended to be learner-centered, thus the various learning environments are expected to meet the needs of diverse student populations, while aiming to provide both high quality and cost-effective learning opportunities (Nokelainen, Silander, Tirri, H., Nevgi & Tirri, K., 2001).

Throughout the 1990's learner-centered learning environments and computer-mediated communication (CMC) systems such as problem-based, project-based, cognitive apprenticeships, constructivist learning environments, and goal-based scenarios, have rather focused on the affordances they provide learners for effecting their way of learning and thinking, than transmitting information from teachers to learners (Land & Hannafin, 2001).

Learner-centered learning is supported theoretically by various overlapping pedagogical concepts such as self-directed learning (Candy, 1991), student-centered instruction or learning (Felder & Brent, 1996), active learning (Ramsden, 1992), vicarious learning (Lee & McKendree, 1999) and cooperative learning (Felder & Brent, 2001). For example, self-directed learning involves dimensions of process and product referring to four related phenomena: personal autonomy, self-management, learner-
control and autodidaxy (Candy, 1991). All these dimensions are present in the process of learner-centered learning where the locus of control is shifted from teacher to the learner who has now a greater responsibility for her own learning.

Learning tasks in learner-centered learning environments include such techniques as substituting active learning experiences for lectures, holding students responsible for material that has not been explicitly discussed in class, assigning open-ended problems and problems requiring critical or creative thinking, and using self-paced and/or cooperative learning. The research findings of educational literature prove convincingly that properly implemented student-centered learning fosters motivation and elicits deeper understanding toward the subject being taught (Felder & Brent, 1996, 2001; Dillinger, 2001).

The learner-centered principles (APA, 1997) present a shift from the traditional teacher-centered approach to learner-centered approach of instruction providing learner with valuable real-life skills (Dillinger, 2001). Motivational and affective factors play a significant role on the list consisting of motivational and emotional influences on learning, intrinsic motivation to learn, and effects of motivation on effort (APA, 1997). However, practical implications that benefit network learners are difficult to conduct as these factors are presented in broad and abstract level.

Specific form of constructivist theory is under continuous debate, but researchers agree (see Bonk & Cunningham, 1998; Dalgarno, 2001; Leflore, 2000; de Vries, 2001) that following characteristics are included: learner construction of meaning (von Glasersfeld, 1995); social interaction to help students learn (Vygotsky, 1978); and student problem-solving in “real world” contexts (Duffy & Jonassen, 1991). Bonk and Cunningham (1998, 32) suggest that we are able to examine motivational issues, such as meaningfulness of studies and self-regulation of learning, in more detailed level from the constructivistic point of view. This thought is supported also by Leflore (2000) who argues, with implemented list of guidelines (p. 113-115), that constructivist approach offers a suitable theoretical base for Web-based learning.
Self-rated Questionnaires in Distance Education

Numerous instruments specialized to collect quantitative data from students in distance education have been developed over time. The interest in motivational issues is apparent as we examine the structure of the instruments. For example, Bigg’s (1985) 42-item Study Process Questionnaire (SPQ) consists of two scales (motive and strategy) with three approaches: (1) surface, (2) deep, and (3) achieving. The questionnaire contains six subscales (surface motive, deep motive, achieving motive, surface strategy, deep strategy, and achieving strategy). Also Entwistle’s and his colleagues (Ramsden & Entwistle, 1981) Approaches to Studying Inventory (ASI), as one of the most widely used questionnaire on student learning in higher education, contains subscales including, for example, fear of failure, extrinsic motivation, and achieving orientation.

Richardson made an extensive review of quantitative research instruments applied to studying campus-based and distance education (2001). He found that broadly the same factors underlied the responses given by distance learning and campus-based students. Richardson concluded that a productive rapprochement can be achieved between the two previously separate research communities, and the findings in the mainstream research literature concerned with approaches to studying in campus-based higher education will be broadly valid for understanding approaches studying in distance education (Richardson, 2001, viii).

Development of Professional Learning Instruments


As the MSLQ was targeted for measuring college student learning, Finnish researchers developed the first adaptation, Motivation and Learning Strategies for Adult Education, for the field of vocational education in the “Growth Needs” project during 1992-1994 (Ruohotie, 1994). The first version of the instrument, with 40 items measuring motivational scale, was tested with both adult learners of several companies acting in
Finnish business sector and adolescent learners of vocational education institutes. Total sample size of this initial study was 156 respondents (Ruohotie, 1994; Kautto-Koivula, 1996).

After this period, ongoing project titled “Motivation and Self-Regulated Learning in Vocational Education” (Ruohotie 1999), continued the development process producing the second version, Motivated Strategies for Professional Learning (Ruohotie, 1998). This version had the same number of items (40) on the motivational scale, but the phrasing of items was improved and the relationship of items to theoretical structure was validated. This version was tested with adolescent learners of various Finnish vocational education institutes; total sample size was 693 students.

**Abilities for Professional Learning Questionnaire**

The latest version, Abilities for Professional Learning (APLQ, Ruohotie, 2000b), still retains the same basic structure as the MSLQ measuring both motivational factors and learning strategies based on Pintrich’s motivational expectancy model (1988). (Table 1.)

<table>
<thead>
<tr>
<th>Table 1. Elaboration of the Motivational Scale for Professional Learning</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Motivation and Learning Strategies for Adult Education</strong> (Ruohotie, 1994)</td>
</tr>
<tr>
<td><strong>Value Components</strong></td>
</tr>
<tr>
<td>Intrinsic Goal Orientation</td>
</tr>
<tr>
<td>Extrinsic Goal Orientation</td>
</tr>
<tr>
<td>Task Value of Learning</td>
</tr>
<tr>
<td><strong>Expectancy Components</strong></td>
</tr>
<tr>
<td>Intrinsic Control Beliefs</td>
</tr>
<tr>
<td>Extrinsic Control Beliefs</td>
</tr>
<tr>
<td>Self-Efficacy</td>
</tr>
<tr>
<td>Expectancy for Success</td>
</tr>
<tr>
<td><strong>Affective Components</strong></td>
</tr>
<tr>
<td>Test Anxiety</td>
</tr>
<tr>
<td>Self-Worth</td>
</tr>
</tbody>
</table>
Method

Sample and Procedure

Our target organization in this study, a Finnish polytechnic institution of higher education, has over 7000 students and a staff of over 700, approximately half of which are teachers. Most of the students are adolescents (18 – 27 years), but training programmes for adults exist, too. The institution is situated in eight geographical units in the Province of Southern Finland within an area of seven municipalities. This is the most populated area in Finland and about half of all Finns live within the Polytechnic's sphere of influence. The polytechnic institution of higher has five degree programmes: 1) Technology and communications; 2) Business and administration; 3) Natural resources; 4) Culture; 5) Health care and social services.

The sample consists of 459 students who had started their studies in 1998 or 1999. They were thus second (n = 281, 61%) or third (n = 164, 35%) year students during March 2000. The sample represents 32% of the survey population (n = 1436) and consists of 323 (70%) female and 134 (29%) male participants. The respondents were divided into three groups based on their age: (1) below 21 years (n = 195, 43%), (2) from 21 to 27 years (n = 225, 49%), and (3) above 27 years (n = 35, 8%). Most of the students (n = 443, 96%) are adolescents, only ten respondents (2%) study in the adult degree program. All of the eight geographical units and five degree programmes of the polytechnic institution are represented in this study.

The heads of each degree program were asked for their opinion about the propositions in the APLQ, and the principle teachers were trained by the researchers to give support in problematic situations to the students as they fill out the questionnaire. The APLQ was delivered in two forms: as a traditional paper and pen test (n = 459), and as a computerized adaptive test (n = 205) based on Bayesian probabilistic modeling. In this article, we report the results and structural validation of the paper and pen questionnaire.

Measurement Instrument

Students responded to 28 items on a 5-point scale. The total number of items in the APLQ is 116 as the questionnaire measures four dimensions of professional learning: 1)
Learning experiences and motivation, 28 items; 2) Study habits, 40 items; 3) Quality of teaching, 23 items; 4) Effects and outcomes of education, 25 items. The response options varied from 1 (completely disagree) to 5 (completely agree). Students also completed a short background information questionnaire containing following sections: age, gender, starting year of studies, degree program, and degree mode (adolescent or adult). As the detailed theoretical background of the APLQ is reported elsewhere (Ruohotie, 2000b, 2002), we describe here only the features that are relevant to the psychometric validation process of the first, motivational part of the questionnaire.

Motivation category of the APLQ has three sections: 1) Value section; 2) Expectancy section; 3) Affective section. The value section has three subscales: 1.1) Intrinsic goal orientation; 1.2) Extrinsic goal orientation; 1.3) Meaningfulness of study. The expectancy section consists of two subscales: 2.1) Control beliefs and 2.2) Self-efficacy. The affective section includes one component: 3.1) Test anxiety. (Table 2.) Reader is encouraged to see Appendix for full item list of APLQ.

**Table 2.** Theoretical Structure and Related Items of the Motivational Scale of Abilities for Professional Learning Questionnaire

<table>
<thead>
<tr>
<th>Motivational Scale (A)</th>
<th>Factors</th>
<th>Items (28) in APLQ</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Value Components</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.1 Intrinsic Goal Orientation</td>
<td>MF1</td>
<td>A1, A18, A23, A25</td>
</tr>
<tr>
<td>1.2 Extrinsic Goal Orientation</td>
<td>MF2</td>
<td>A8, A13, A27</td>
</tr>
<tr>
<td>1.3 Meaningfulness of Study</td>
<td>MF3</td>
<td>A4 (rev.), A5, A11, A15, A19, A24, A28</td>
</tr>
<tr>
<td><strong>2. Expectancy Components</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1 Control Beliefs</td>
<td>MF4</td>
<td>A2, A10, A20, A26</td>
</tr>
<tr>
<td>2.2 Self-Efficacy</td>
<td>MF5</td>
<td>A6, A7, A12, A17, A22</td>
</tr>
<tr>
<td><strong>3. Affective Components</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.1 Test Anxiety</td>
<td>MF6</td>
<td>A3, A9, A14, A16, A21</td>
</tr>
</tbody>
</table>
Statistical Analyses

The statistical procedures were conducted in five stages: 1) Variable selection based on descriptive statistics; 2) Analysis of communalities; 3) Bayesian dependency modeling; 4) Explorative factor analysis; 5) Bayesian unsupervised model-based visualization.

First, we analyzed all the twenty-eight items to see if they were technically applicable for linear statistic computations, such as explorative and confirmatory factor analysis, based on multivariate normality assumption. Bayesian modeling techniques are in most cases free from such presuppositions (Congdon, 2001, p. 1; Nokelainen, Ruohotie & Tirri, 1999, p. 113). Second, we investigated the data exploratively with factor analysis to see if the six dimensions (motivational factors) are identified in this domain. In the third stage of the analysis we examined probabilistic dependences between all variables with Bayesian search algorithm (Myllymäki, Silander, Tirri & Uronen, 2001) in order to find a model with the highest probability. The fourth stage was to conduct the traditional construct validity test with Cronbach’s alpha (Cronbach, 1951). The task for the last analysis method, Bayesian unsupervised model-based visualization (Kontkanen, Lahtinen, Myllymäki & Tirri, 2000), was to provide information how the derived factors are interrelated in individual level.

Results

Variable Selection Based on Descriptive Statistics

The aim of the first stage was to examine learning motivation variables and choose the acceptable ones for explorative factor analysis. Frequency analyses were carried out for all the variables. The results are presented in Table 3.

Table 3 shows that respondents used the whole scale from 1 to 5, except for five propositions the scale was from 2 to 5. This leads to a notion that distribution of the mode is strongly biased towards positive values: 1) \( n = 2 \); 2) \( n = 3 \); 3) \( n = 8 \), 4) \( n = 15 \); 5) \( n = 0 \). (Table 3.)
Table 3. Descriptive Statistics of the Motivational Scale of Abilities for Professional Learning Questionnaire

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>3.57</td>
<td>0.90</td>
<td>-0.20</td>
<td>0.11</td>
<td>-0.54</td>
<td>0.23</td>
</tr>
<tr>
<td>A2</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>4.08</td>
<td>0.80</td>
<td>-0.69</td>
<td>0.11</td>
<td>0.18</td>
<td>0.23</td>
</tr>
<tr>
<td>A3</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>1.77</td>
<td>0.91</td>
<td>1.07</td>
<td>0.11</td>
<td>0.38</td>
<td>0.23</td>
</tr>
<tr>
<td>A4</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>2.59</td>
<td>1.13</td>
<td>0.30</td>
<td>0.11</td>
<td>-0.83</td>
<td>0.23</td>
</tr>
<tr>
<td>A5</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>4.18</td>
<td>0.78</td>
<td>-0.79</td>
<td>0.11</td>
<td>0.46</td>
<td>0.23</td>
</tr>
<tr>
<td>A6</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>3.34</td>
<td>0.83</td>
<td>-0.13</td>
<td>0.11</td>
<td>0.06</td>
<td>0.23</td>
</tr>
<tr>
<td>A7</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>3.07</td>
<td>1.06</td>
<td>0.00</td>
<td>0.11</td>
<td>-0.75</td>
<td>0.23</td>
</tr>
<tr>
<td>A8</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>3.68</td>
<td>1.00</td>
<td>-0.56</td>
<td>0.11</td>
<td>-0.08</td>
<td>0.23</td>
</tr>
<tr>
<td>A9</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>2.85</td>
<td>1.16</td>
<td>0.00</td>
<td>0.11</td>
<td>-1.04</td>
<td>0.23</td>
</tr>
<tr>
<td>A10</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>3.10</td>
<td>1.05</td>
<td>0.06</td>
<td>0.11</td>
<td>-0.81</td>
<td>0.23</td>
</tr>
<tr>
<td>A11</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>4.08</td>
<td>0.76</td>
<td>-0.46</td>
<td>0.11</td>
<td>-0.26</td>
<td>0.23</td>
</tr>
<tr>
<td>A12</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>3.57</td>
<td>0.85</td>
<td>-0.32</td>
<td>0.11</td>
<td>-0.21</td>
<td>0.23</td>
</tr>
<tr>
<td>A13</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>2.88</td>
<td>0.98</td>
<td>0.16</td>
<td>0.11</td>
<td>-0.46</td>
<td>0.23</td>
</tr>
<tr>
<td>A14</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>2.22</td>
<td>1.17</td>
<td>0.75</td>
<td>0.11</td>
<td>-0.41</td>
<td>0.23</td>
</tr>
<tr>
<td>A15</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>3.23</td>
<td>1.02</td>
<td>-0.04</td>
<td>0.11</td>
<td>-0.59</td>
<td>0.23</td>
</tr>
<tr>
<td>A16</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>2.52</td>
<td>1.09</td>
<td>0.37</td>
<td>0.11</td>
<td>-0.66</td>
<td>0.23</td>
</tr>
<tr>
<td>A17</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>3.72</td>
<td>0.81</td>
<td>-0.34</td>
<td>0.11</td>
<td>-0.05</td>
<td>0.23</td>
</tr>
<tr>
<td>A18</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>3.65</td>
<td>0.92</td>
<td>-0.26</td>
<td>0.11</td>
<td>-0.42</td>
<td>0.23</td>
</tr>
<tr>
<td>A19</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>4.05</td>
<td>0.90</td>
<td>-0.82</td>
<td>0.11</td>
<td>0.25</td>
<td>0.23</td>
</tr>
<tr>
<td>A20</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>4.29</td>
<td>0.71</td>
<td>-0.77</td>
<td>0.11</td>
<td>0.34</td>
<td>0.23</td>
</tr>
<tr>
<td>A21</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>1.83</td>
<td>1.01</td>
<td>1.22</td>
<td>0.11</td>
<td>0.94</td>
<td>0.23</td>
</tr>
<tr>
<td>A22</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>3.80</td>
<td>0.82</td>
<td>-0.42</td>
<td>0.11</td>
<td>-0.01</td>
<td>0.23</td>
</tr>
<tr>
<td>A23</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>3.37</td>
<td>0.93</td>
<td>-0.10</td>
<td>0.11</td>
<td>-0.34</td>
<td>0.23</td>
</tr>
<tr>
<td>A24</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>4.28</td>
<td>0.73</td>
<td>-0.72</td>
<td>0.11</td>
<td>0.04</td>
<td>0.23</td>
</tr>
<tr>
<td>A25</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>3.47</td>
<td>0.88</td>
<td>-0.45</td>
<td>0.11</td>
<td>0.04</td>
<td>0.23</td>
</tr>
<tr>
<td>A26</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>3.18</td>
<td>0.99</td>
<td>-0.04</td>
<td>0.11</td>
<td>-0.57</td>
<td>0.23</td>
</tr>
<tr>
<td>A27</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>3.14</td>
<td>1.08</td>
<td>-0.10</td>
<td>0.11</td>
<td>-0.63</td>
<td>0.23</td>
</tr>
<tr>
<td>A28</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>3.95</td>
<td>0.80</td>
<td>-0.31</td>
<td>0.11</td>
<td>-0.54</td>
<td>0.23</td>
</tr>
</tbody>
</table>

Numerous publications declare that certain attributes belong to data that is appropriate for multivariate analysis (e.g., Tabachnick & Fidell, 1996, pp. 13-17; Thompson, 1999; Bradley & Schaefer, 1998, pp. 79-83). The most common criteria’s for accepting variables for further multivariate analysis are: 1) Standard deviation maximum half the mean; 2) Skewness less than +/- .3; 3) Correlation between +/- .3-.7. Data examination results for the items in the learning motivation part of the instrument are presented in Table 4. The table indicates lowest rejection rate (7%) for the third criteria and highest (93%) for the second criteria.
Table 4. Criteria for Motivational Scale Item Rejection

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Items (n = 28)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Accepted n (%)</td>
</tr>
<tr>
<td>1. SD</td>
<td>25 (89)</td>
</tr>
<tr>
<td>2. Skewness</td>
<td>+/- .3</td>
</tr>
<tr>
<td></td>
<td>+/- .5</td>
</tr>
<tr>
<td></td>
<td>+/- .7</td>
</tr>
<tr>
<td></td>
<td>+/- .8</td>
</tr>
<tr>
<td></td>
<td>+/- .9</td>
</tr>
<tr>
<td>3. Correlation</td>
<td>26 (93)</td>
</tr>
</tbody>
</table>

The first and second criteria for item rejection were inspected with the help of descriptive statistics presented in Table 3. Examination suggested rejection of three variables (A3, A14, A21) based on the first criteria, and rejection of 26 variables based on the second criteria. As it seemed impossible to take the second criteria literally due to high rejection rate at .3 level, we examined the skewness of variables at four additional levels of which .8 proved to be suitable for this data set (Table 4). In order to strengthen our opinion about the right acceptance level based on skewness of the variables, we compared the results to our previous study aimed to develop the first version of the APLQ instrument (Ruohotie & Nokelainen, 2000a). The data consisted of 156 adults working in a Finnish telemedia company. We also used for selected evaluation tasks the comparison data consisting of 512 adolescents studying at Finnish polytechnic institution. The results were almost identical with this comparison data set suggesting rejection level between .8 and .9.

Spearman nonparametric correlation was calculated due to use of discrete scale in the questionnaire. The variables correlated satisfactory and their scale ranged from -.34 to .74 ($M = .10, SD = .20$). We suggest rejection of two variables (A13, A15) from the analysis on the basis of the third criteria. Examination of the correlation matrix proved that 255 out of 392 correlations (65%) were significant at the .01 level and 30 correlations (8%) were significant at the .05 level. The correlation matrix contained 107 uncorrelated pairs of variables (27%). The significant correlations on both levels (.01 and .05) and non-
significant correlations were uniformly distributed. (Table 4.) Inter-item correlations (.41 - .78) did not indicate multicollinearity or singularity.

After comparing the results of the rejection criteria’s we decided to reject four variables, namely A3, A13, A15, and A21. Reader should notice that we intend to use the variables in following phases three and four that are based on Bayesian modeling.

Table 5 shows the results of reliability analysis supporting our rejection decisions. In the original solution Cronbach’s Alpha level is good for only two factors (Self-Efficacy, Test Anxiety) and satisfactory for three other (Intrinsic Goal Orientation, Meaningfulness of Study, and Control Beliefs), but poor for Extrinsic Goal Orientation. The optimized 24-item solution shows essential improvement on Extrinsic Goal Orientation and Meaningfulness of Study. The overall reliability for the motivational scale of the APLQ questionnaire is good for both original (Alpha = .73) and optimized (Alpha = .79) solutions. We found similar reliability estimates (original = .79, optimized = .83) as testing these two solutions with our comparison data (n = 512). (Table 5.)
Table 5. Reliability Comparison of the Original 28 Item and Optimized 24 Item Solutions

<table>
<thead>
<tr>
<th>Motivational Scale</th>
<th>Original solution 2</th>
<th>Optimized solution 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Value Components</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MF1 Intrinsic Goal Orientation</td>
<td>A1, A18, A23, A25</td>
<td>A1, A18, A23, A25</td>
</tr>
<tr>
<td>MF2 Extrinsic Goal Orientation</td>
<td>.64 (.66) 1</td>
<td>A1, A18, A23, A25</td>
</tr>
<tr>
<td>MF4 Control Beliefs</td>
<td>A2, A10, A20, A26</td>
<td>A2, A10, A20, A26</td>
</tr>
<tr>
<td>MF5 Self-Efficacy</td>
<td>A6, A7, A12, A17, A22</td>
<td>A6, A7, A12, A17, A22</td>
</tr>
<tr>
<td><strong>2. Expectancy Components</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MF6 Test Anxiety</td>
<td>A3, A9, A14, A16, A21</td>
<td>A9, A14, A16, A21</td>
</tr>
</tbody>
</table>

1) Values presented inside parenthesis are derived from comparison data (n = 519)
2) Cronbach’s Alpha for 28 item solution is .73 (.79)
3) Cronbach’s Alpha for 24 item solution is .79 (.83)

Communalities

In the second phase of the analysis we examined communalities in the original polytechnic institution of higher education data (n = 459) and in three sub samples (n = 200, n = 150, n = 100) derived from it. We also inspected communalities of the comparison data (n = 512) collected from a Finnish polytechnic institution. It was interesting to see that the lower bound of initial estimates and total variance explained had notable better values as the sample size decreased, but the upper bound of initial estimates remained at the same level. The results presented in Table 6 suggest the use of 21-item solution as the basis for further analysis.
Table 6. Item Rejection Based on Communalities ($h^2$) in Five Samples

<table>
<thead>
<tr>
<th>Items</th>
<th>Rejected Items</th>
<th>Data$^1$</th>
<th>Sample Size</th>
<th>Initial Estimates</th>
<th>Factors</th>
<th>Total Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td>A3,A13,A15,A21</td>
<td>1</td>
<td>459</td>
<td>.16 -.58</td>
<td>6</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>200</td>
<td>.21 -.61</td>
<td>6</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>150</td>
<td>.23 -.63</td>
<td>7</td>
<td>62</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>100</td>
<td>.31 -.62</td>
<td>8</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>512</td>
<td>.18 -.53</td>
<td>7</td>
<td>62</td>
</tr>
<tr>
<td>21</td>
<td>A1,A3,A4,A9,A13,A15,A21</td>
<td>1</td>
<td>459</td>
<td>.24 -.58</td>
<td>6</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>200</td>
<td>.31 -.60</td>
<td>6</td>
<td>61</td>
</tr>
<tr>
<td>18</td>
<td>A1,A3,A4,A9,A13,A15,A10,A18,A21,A25</td>
<td>1</td>
<td>459</td>
<td>.18 -.57</td>
<td>5</td>
<td>60</td>
</tr>
<tr>
<td>17</td>
<td>A1,A3,A4,A9,A13,A15,A10,A18,A21,A25,A26</td>
<td>1</td>
<td>459</td>
<td>.28 -.56</td>
<td>4</td>
<td>56</td>
</tr>
</tbody>
</table>

1 Polytechnic Institution of Higher Education; 2 Polytechnic Institution

Bayesian Dependence Modeling

The third stage of the analysis was to build a Bayesian network out of 28-item solution measuring self-evaluated motivation. The rationale for this procedure was to examine dependencies between variables by both their visual representation and probability ratio of each dependency.

A Bayesian network is a representation of a probability distribution over a set of random variables, consisting of an directed a cyclic graph (DAG), where the nodes correspond to domain variables, and the arcs define a set of independence assumptions which allow the joint probability distribution for a data vector to be factorized as a product of simple conditional probabilities. Graphical visualization of Bayesian network contains two components: 1) Observed variables visualized as ellipses and 2) dependences visualized as lines between nodes. Solid lines indicate direct causal relations and dashed lines indicate dependency where it is not sure if there is a direct causal influence or latent cause. Variable is considered as independent of all other variables if there is no line attached to it. We have proved in our earlier research work that Bayesian networks are useful for explorative analysis of causal structures between observed variables (Ruohotie & Nokelainen, 2000a; Nokelainen et al., 2001).
Probabilistic dependencies between all of the motivational scale variables are presented in Table 7. Bayesian search algorithm evaluated the data set in order to find the model with the highest probability. During the extensive search, 738,861 models were evaluated. Visual examination of the network model shows that “Control Beliefs” scale is two-dimensional: one dimension represents negative attitude towards one’s own capabilities, for example, A10 (“It is my own fault if I fail to learn theory”), while the other dimension is about the learners beliefs of her capability to manage learning tasks, for example, A20 (“I will acquire the required professional skills if I work hard enough”). The analysis of the dependency paths in the network model suggests that variables A4 (“In my opinion craftsmanship can be acquired through practice without theory lessons”), A13 (“I am only interested in mastering learning tasks that are required in real working life”), and A15 (“Studying feels often burdensome and/or frustrating”) should be removed from the final model. The network model indicates that MF1, MF3, and MF5 dimensions form an isolated homogenous cluster together with the “optimistic” part of the MF4 dimension. Closer examination of the probability ratios of dependences reveal that also variables A1 (“I prefer to study theoretical subjects and work tasks which are demanding and from which I can learn something new”), A3 (“During an exam I wonder how I am performing in comparison to other students”), A9 (“When answering essay questions I am also concerned about the other questions on the test that I cannot answer”), and A25 (“When given an opportunity I choose exercises and literature from which I can learn something new even if it means receiving lower grades than I could get by choosing those that I already know something about”) should be omitted from the model. (Table 7.)

The variable rejection based on visual examination of the network model is in general parallel with the results obtained from the first and second phases of the analysis. When interpreting the results we must bear in mind that the relationships between variables in the network model are presented with dashed lines, and thus it is not sure if there is a direct causal influence or latent cause.
Table 7. Bayesian Network Model and the Importance Ranking of the 28 Item Solution Measuring Self-evaluated Motivation

<table>
<thead>
<tr>
<th>Network Model</th>
<th>Dependency</th>
<th>Probability ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A14-&gt;A21</td>
<td>1 : Inf.</td>
</tr>
<tr>
<td></td>
<td>A14-&gt;A7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A17-&gt;A6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A22-&gt;A2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A17-&gt;A22</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A17-&gt;A2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A22-&gt;A12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A11-&gt;A28</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A22-&gt;A23</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A5-&gt;A11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A24-&gt;A5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A20-&gt;A19</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A8-&gt;A27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A17-&gt;A20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A19-&gt;A24</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A26-&gt;A10</td>
<td>1 : 1.000.000</td>
</tr>
<tr>
<td></td>
<td>A21-&gt;A16</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A23-&gt;A18</td>
<td></td>
</tr>
<tr>
<td></td>
<td>A21-&gt;A3</td>
<td>1 : 16343</td>
</tr>
<tr>
<td></td>
<td>A3-&gt;A9</td>
<td>1 : 648</td>
</tr>
<tr>
<td></td>
<td>A1-&gt;A17</td>
<td>1 : 180</td>
</tr>
<tr>
<td></td>
<td>A23-&gt;A25</td>
<td>1 : 65</td>
</tr>
<tr>
<td></td>
<td>A4_REV, A13,A15</td>
<td>No importance</td>
</tr>
</tbody>
</table>

The advantage of applying Bayesian dependency modeling as a tool to identify independent variables was clearly proved by the early stage identification of both A1 and A4 variables. The comparison of the results of the first (Tables 4), second (Table 6) and third phase (Table 7) variable selection are presented in Table 8. Reader should bear in mind that argument for rejection in the first and second phase is the technical quality of the data, while in the third phase it is the strength of dependencies between variables.
**Table 8.** Comparison of the First, Second, and Third Phase Item Selection

<table>
<thead>
<tr>
<th>Selection Method</th>
<th>Rejected Items</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Phase 1</strong></td>
<td></td>
</tr>
<tr>
<td>SD, Skewness Correlation Matrix</td>
<td>A3, A13, A15, A21</td>
</tr>
<tr>
<td><strong>Phase 2</strong></td>
<td></td>
</tr>
<tr>
<td>Communalities</td>
<td>A1, A3, A4, A9, A13, A15, A21</td>
</tr>
<tr>
<td><strong>Phase 3</strong></td>
<td></td>
</tr>
<tr>
<td>Bayesian Dependency Network</td>
<td>A1, A3, A4, A9, A13, A15</td>
</tr>
</tbody>
</table>

**Explorative Factor Analysis**

The fourth phase of the analysis is based on exploratory factor analysis. We chose Maximum Likelihood algorithm due to adequate sample size and the first phase variable selection. The rotation method applied in this study was Direct Oblimin as we let the motivational factors correlate with each other.

Table 9 shows the correlations between the variables and the factors. Each variable’s relationship to the theoretical model of motivation is indicated with gray background. Reader should notice that values below +/- .20 are omitted from the output.
Table 9. Factor Structure Matrix of the 21 Item Solution

<table>
<thead>
<tr>
<th>Items</th>
<th>MF5</th>
<th>MF6</th>
<th>MF4</th>
<th>MF2</th>
<th>MF1</th>
<th>MF3</th>
</tr>
</thead>
<tbody>
<tr>
<td>A17</td>
<td>0.74</td>
<td>-0.45</td>
<td>0.30</td>
<td>-0.39</td>
<td>-0.39</td>
<td></td>
</tr>
<tr>
<td>A7</td>
<td>0.65</td>
<td>-0.41</td>
<td>0.24</td>
<td>-0.37</td>
<td>-0.25</td>
<td></td>
</tr>
<tr>
<td>A23</td>
<td>0.60</td>
<td>-0.26</td>
<td>0.24</td>
<td>-0.37</td>
<td>-0.25</td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td>0.57</td>
<td>-0.35</td>
<td>0.34</td>
<td>-0.22</td>
<td>-0.35</td>
<td></td>
</tr>
<tr>
<td>A25</td>
<td>0.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A28</td>
<td>0.49</td>
<td>-0.32</td>
<td>-0.26</td>
<td>-0.35</td>
<td>-0.35</td>
<td></td>
</tr>
<tr>
<td>A18</td>
<td>0.47</td>
<td>-0.32</td>
<td>-0.26</td>
<td>-0.35</td>
<td>-0.35</td>
<td></td>
</tr>
<tr>
<td>A12</td>
<td>0.44</td>
<td>-0.27</td>
<td>0.30</td>
<td>-0.35</td>
<td>-0.35</td>
<td>-0.38</td>
</tr>
<tr>
<td>A14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A20</td>
<td>0.37</td>
<td>-0.26</td>
<td>0.42</td>
<td>-0.24</td>
<td>-0.22</td>
<td>-0.39</td>
</tr>
<tr>
<td>A8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A22</td>
<td>0.52</td>
<td>-0.43</td>
<td>0.67</td>
<td>-0.22</td>
<td>-0.39</td>
<td>-0.39</td>
</tr>
<tr>
<td>A27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A6</td>
<td>0.38</td>
<td>-0.31</td>
<td>0.59</td>
<td>-0.31</td>
<td>-0.28</td>
<td></td>
</tr>
<tr>
<td>A19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A5</td>
<td>0.32</td>
<td>-0.23</td>
<td>0.49</td>
<td>-0.36</td>
<td>-0.60</td>
<td></td>
</tr>
<tr>
<td>A11</td>
<td>0.43</td>
<td>-0.26</td>
<td>0.49</td>
<td>-0.36</td>
<td>-0.60</td>
<td></td>
</tr>
<tr>
<td>A24</td>
<td>0.30</td>
<td>-0.26</td>
<td>0.49</td>
<td>-0.36</td>
<td>-0.60</td>
<td></td>
</tr>
</tbody>
</table>


The correlations between motivational factors are presented in Table 10. Low correlation indicates that factors are nearly orthogonal. This is the case when we compare “Intrinsic Goal Orientation” to “Control Beliefs”, “Test Anxiety”, and “Extrinsic Goal Orientation”. “Self-Efficacy” correlates relatively strongly with “Test Anxiety” and “Meaningfulness of Study”. Examination of the Table 10 justifies our earlier decision to allow the factors to correlate, as there are only three uncorrelated pairs out of 15.
Table 10. Factor Correlation Matrix

<table>
<thead>
<tr>
<th>Factor</th>
<th>MF5</th>
<th>MF6</th>
<th>MF4</th>
<th>MF2</th>
<th>MF1</th>
<th>MF3</th>
</tr>
</thead>
<tbody>
<tr>
<td>MF5 Self-Efficacy</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MF6 Test Anxiety</td>
<td>0.31</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MF4 Control Beliefs</td>
<td>0.18</td>
<td>0.15</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MF2 Extrinsic Goal Orientation</td>
<td>0.28</td>
<td>0.07</td>
<td>0.17</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MF1 Intrinsic Goal Orientation</td>
<td>0.26</td>
<td>0.06</td>
<td>0.00</td>
<td>0.10</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>MF3 Meaningfulness of Study</td>
<td>0.44</td>
<td>0.15</td>
<td>0.16</td>
<td>0.33</td>
<td>0.23</td>
<td>1.00</td>
</tr>
</tbody>
</table>

The factor structure matrix presented earlier in Table 9 encourages us to test the reliability of two optional solutions: the first 21 item solution is based solely on technical appropriateness of variables for linear multivariate analysis, and the second utilizes the results of the Bayesian network model of motivation in the decision making process. The reliability estimates of the two solutions were verified with a comparison data ($n = 512$). The comparison of the optimized solutions is presented in Table 11.

Table 11. Reliability Comparison of the Two Optimized Solutions

<table>
<thead>
<tr>
<th>Motivational Scale</th>
<th>Optimized solution $a^2$</th>
<th>Optimized solution $b^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Items (21)</td>
<td>Alpha</td>
</tr>
<tr>
<td>1. Value Components</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MF1 Intrinsic Goal Orientation</td>
<td>A19, A23, A25</td>
<td>.57 (.62)$^1$</td>
</tr>
<tr>
<td>MF2 Extrinsic Goal Orientation</td>
<td>A6, A8, A27</td>
<td>.62 (.53)</td>
</tr>
<tr>
<td>2. Expectancy Components</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MF4 Control Beliefs</td>
<td>A2, A10, A20, A26</td>
<td>.60 (.58)</td>
</tr>
<tr>
<td>MF5 Self-Efficacy</td>
<td>A7, A12, A17, A18, A22</td>
<td>.76 (.78)</td>
</tr>
<tr>
<td>3. Affective Components</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MF6 Test Anxiety</td>
<td>A14, A16</td>
<td>.62 (.61)</td>
</tr>
</tbody>
</table>

1) Values presented inside parenthesis are derived from comparison data ($n = 519$)
2) Cronbach’s Alpha for the solution is .79 (.84)
3) Cronbach’s Alpha for the solution is .76 (.82)
We conducted confirmatory factor analysis in order to compare the original 28-item solution to two optimized (24 and 21 item) solutions. The results of model testing using $\chi^2$, information criteria, root mean square error of approximation measures, and residual measures are presented in Table 12. The relative chi-squares for both models indicated an acceptable fit between the hypothetical model and the sample data. The root mean square error of approximation values indicates a reasonable error of approximation for all three models. When compared to baseline model, the optimized 21-item solution proves to be the best by both comparative fit index and Tucker-Lewis coefficient. When the models are compared by standardized root mean square residual indices, the optimized 24-item solution is slightly better than 21-item model. (Table 12.)

<table>
<thead>
<tr>
<th></th>
<th>Original 28 Item Solution</th>
<th>Optimized 24 Item Solution</th>
<th>Optimized 21 Item Solution b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-Square Test of Model Fit</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Value</td>
<td>1093.435</td>
<td>761.370</td>
<td>651.878</td>
</tr>
<tr>
<td>Degrees of Freedom</td>
<td>335</td>
<td>237</td>
<td>174</td>
</tr>
<tr>
<td>$\chi^2 / df$</td>
<td>3.26</td>
<td>3.21</td>
<td>3.75</td>
</tr>
<tr>
<td>P-Value</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>RMSEA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estimate</td>
<td>0.071</td>
<td>0.070</td>
<td>0.078</td>
</tr>
<tr>
<td>90% C.I.</td>
<td>0.066 [0.075, 0.064]</td>
<td>0.064 [0.075, 0.071]</td>
<td>0.071 [0.084, 0.071]</td>
</tr>
<tr>
<td>Probability RMSEA &lt;= .05</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>CFI</td>
<td>0.792 [0.821, 0.765]</td>
<td>0.821 [0.834, 0.800]</td>
<td>0.834</td>
</tr>
<tr>
<td>TLI</td>
<td>0.765 [0.792, 0.765]</td>
<td>0.792 [0.800, 0.792]</td>
<td>0.800</td>
</tr>
<tr>
<td>SRMR</td>
<td>0.073 [0.064, 0.073]</td>
<td>0.064 [0.068, 0.064]</td>
<td>0.068</td>
</tr>
</tbody>
</table>

Bayesian Unsupervised Model-based Visualization

The fifth and last stage of the analysis was to examine the dispersion of single data vectors ($n = 469$) in three-dimensional space. The data was mapped into six dimensions (MF1, MF2, MF3, MF4, MF5, MF6) according to the 21 item optimized solution (see Optimized Solution b in Table 11) from which Bayesian algorithm produced one optimal
Bayesian unsupervised model-based visualization differs from Multidimensional Scaling (MDS) in the following ways: 1) Bayesian approach is parameter-free and the user input is not required, instead, prior distributions of the model offer a theoretically justifiable method for affecting the model construction; 2) Bayesian methods work with probabilities and can hence be expected to produce smooth and robust visualizations with discrete data containing nominal and ordinal attributes; 3) Bayesian approach has no limit for minimum sample size; 4) Bayesian modeling assumes no multivariate normal model (Kontkanen, Lahtinen, Myllymäki & Tirri, 2000). Researchers (Ruohotie & Nokelainen, 2000b; Ruohotie, Nokelainen & Tirri, 2002) have applied this multidimensional data visualization method earlier to research questions concerning growth-oriented atmosphere.

Figure 1 presents two-dimensional visualization of Intrinsic Goal Orientation (MF1) and Test Anxiety (MF6). Distribution of MF1 is asymptotic, but MF6 is skewed towards small values. The figure shows clearly that the polarity changes in the model: respondents with strong goal orientation have little test anxiety. On the other hand, those who have weak goal orientation are more likely to suffer from test anxiety. Examination of Figure 2 shows that distribution of extrinsic goal orientation (MF3) is skewed to left ranging from 2.0 to 5.0. It is interesting to see that the “top” group of seventeen respondents located in the lower right corner in each three figure has consistent values on dimensions MF1, MF3, MF5, and MF6, but not on dimensions MF2 and MF4.

Figure 3 presents profiles of two polar subsamples selected from the visualization of the three-dimensional model. Upper sample consists of twenty students, and lower sample seventeen students. Visual inspection of the figure reveals that the best discriminating dimensions between two selections are MF1, MF3, MF4, and MF5. Members of the upper sample have low intrinsic goal orientation and studying feels burdensome. If they fail on learning task, it is natural to accuse lack of effort (MF4) or ability (MF5). Members of the lower sample are eager to learn new things, as their studies feel meaningful. They believe in success in studies due to high level of ability and
effort. The two groups have most in common with dimension MF2 indicating that both groups’ extrinsic goal orientations are above average.

Figure 1. Bayesian Model-based Visualization of Intrinsic Goal Orientation (MF1) and Test Anxiety (MF6)

Figure 2. Bayesian Model-based Visualization of Extrinsic Goal Orientation (MF2) and Meaningfulness of Study (MF3)
Figure 3. Bayesian Model-based Visualization of Profiles of Two Groups Presenting the Extremes of Polarity on the Motivational Scale
Discussion

In this article we have examined construct validity of motivational scale of the Abilities for Professional Growth measurement instrument. The instrument is based on Pintrich’s motivational expectancy theory. We have discussed the new role of self-assessment instruments on the research field of motivational study. Technical development process of the APLQ measurement instrument is described thoroughly, but description of the theoretical framework is limited to essential parts as it is presented in full length in the second chapter of this book.

The analysis was based on a sample of 459 adolescent students of Finnish polytechnic institution of higher education. The total number of participants in this study was 1127 as we validated some of the models with two additional data sets. The statistical procedures were conducted in five stages. Exploratory factor analysis, Bayesian dependence modeling and Bayesian unsupervised model-based visualization were the statistical techniques applied in this study.

The other goal of this study was to find the core items out of 28 in the original model for the use of limited time tasks and time series analysis. The results of factor analysis and Bayesian dependence modeling showed that the six factor 21-item model has similar descriptive power than the original 28-item version. The optimized 21-item solution (b) proved to be satisfactory when compared to baseline model. The optimized 24-item solution was slightly better than 21-item solution when compared by residual-based fit indices.

We intend to carry on the development process of the 21-item solution as it was found to be interpretable in terms of meaningful clusters and correspondence to theoretical structure. Empirical findings of our previous research work also supports acceptance of this model.

Polarity change in the test anxiety scale was inspected with Bayesian model-based visualization. With this new method it was easy to profile each respondent, or group, by it’s motivational level on each of the six dimensions. In addition, we examined profiles of two polar sub samples, first one representing students with high, and second with low motivational level. We concluded by the visual inspection of the profiles that the students
with low motivation are likely to accuse lack of effort or ability if they fail in learning tasks, as opposite to highly motivated students who believe that they succeed in their studies due to high level of ability and effort. The both groups had extrinsic goal orientations above average.

Our next task is to collect data from virtual university students and test if the 21-item solution is applicable also to that domain. In the future, it is interesting to see if the instrument is as non-discriminating in the virtual university population as it seems to be for the students of polytechnic institutions of higher education.

References


APPENDIX

Abilities for Professional Learning Questionnaire (APLQ)

Part A  Learning Experiences and Motivation

A1  I prefer to study theoretical subjects and work tasks which are demanding and from which I can learn something new.
A2  I am able to learn even the most difficult subjects if I use good study methods.
A3  During an exam I wonder how I am performing in comparison to other students.
A4  In my opinion craftsmanship can be acquired through practice without theory lessons.
A5  I believe that theory lessons will prove useful in practical training and later in working life.
A6  I expect to get excellent grades in my vocational/occupational studies.
A7  I am confident that I understand even the most difficult aspects of my studies.
A8  I want to receive as high grades as possible.
A9  When answering essay questions I am also concerned about the other questions on the test that I cannot answer.
A10  It is my own fault if I fail to learn theory.
A11  It is important for me to learn theory related to my profession.
A12  I am confident that I will learn the skills related to my field.
A13  I am only interested in mastering learning tasks that are required in real working life.
A14  Nervousness during exams affects my performance.
A15  Studying feels often burdensome and/or frustrating.
A16  When taking part in a practical examination I am concerned about failing and what will happen as a result.
A17  I am confident that I will learn even the most difficult theoretical subjects and work tasks.
A18  I prefer to study theoretical subjects that interest me even if I find them difficult.
A19  I am very interested in my field of study as well as in the new information related to it.
A20  I will acquire the required professional skills if I work hard enough.
A21  I am really nervous in all test situations.
A22  I am confident that I will do well in my studies.
A23  I find it most rewarding when I can research a subject as thoroughly as possible.
A24  I believe that my vocational/occupational studies will be of practical benefit to me.
A25  When given an opportunity I choose exercises and literature from which I can learn something new even if it means receiving lower grades than I could get by choosing those that I already know something about.
A26  If I do not understand theory, it is because I am not trying hard enough.
A27  It is important for me to do well in my studies and show others (my family, friends, colleagues) what I am capable of.
A28  It is essential for me to understand the topics contained in my vocational/occupational studies.
Investigating the Number of Non-linear and Multi-modal Relationships Between Observed Variables Measuring A Growth-oriented Atmosphere

P. NOKELAINEN1*, T. SILANDER2, P. RUOHOTIE1 and H. TIRRI3

1University of Tampere; 2Complex Systems Computation Group; 3Nokia Research Center

Abstract. This study investigates the number of non-linear and multi-modal relationships between observed variables measuring a Growth-oriented Atmosphere. The sample \( n = 726 \) represents employees of three vocational high schools in Finland. The first stage of analysis showed that only 22 per cent of all dependencies between variables were purely linear. In the second stage two sub samples of the data were identified as linear and non-linear. Both bivariate correlations and confirmatory factor analysis (CFA) parameter estimates were found to be higher in the linear sub sample. Results showed that some of the highest bivariate correlations in both sub samples were explained via a third variable in the non-linear Bayesian dependence modeling (BDM). Finally, the results of CFA and BDM led to different substantive interpretations in two out of four research questions concerning organizational growth.

Key words: categorical data, survey data, non-linear modeling, structural equation modeling, organizational atmosphere

1. Introduction

When an organizational researcher wants to study dependencies between observed and latent variables, for example, the factors of a growth-oriented atmosphere (Ruohotie, 1996), the assumptions for the data may become quite challenging in traditional frequentistic statistical analysis. A few examples of such assumptions are the assumption of continuous measurement level, multivariate normality and linearity of both the data and phenomena under investigation. For example, the result of violating multivariate normality assumption is that chi-square becomes too large (too many models are rejected) and standard errors become too small (significance tests have too much power). In addition, normal distribution analysis sets minimum requirements for the number of observations.

As a first response to the continuous measurement level and multivariate normality assumptions, Muthén and Kaplan (1985) suggest that treating the ordinal variables as continuous does produce viable results as long as the frequency distributions are unimodal with an internal

---

1 * Author for correspondence: Ed.Lic Petri Nokelainen, Research Centre for Vocational Education, University of Tampere, P.O. Box 229, FIN-13101 Hämeenlinna, Finland. Email: petri.nokelainen@uta.fi.
mode. Johnson & Creech (1983) have simulation studies to examine the categorization error that occurs when continuous variables are measured by indicators with only a few categories. The results indicated that while categorization error does produce distortions in multiple indicator models, under most conditions explored the bias was not sufficient to alter substantive interpretations with a large simulation sample \((n = 5000)\). However, the authors urged caution in the use of two-, three- or four-category ordinal indicators, particularly when the sample size is small (in that study ten sub samples of 500 cases were examined). They were also worried about situations in which it is not certain that a normal distribution accurately reflects the true distribution of many underlying variables.

The second solution that is presented theoretically by Muthén (1983, 1984, 1989) and applied in practice in LISREL by Jöreskog (2003) is to estimate tetrachoric (for binary variables) or polychoric (for categorical variables) correlations among the ordinal variables and use these correlations to estimate the model using asymptotic distribution free function (ADF) by Browne (1984). Amos (Arbuckle, 1999) software package uses Browne’s original name, but EQS by Bentler (1995) describes it as arbitrary generalized least squares (AGLS) and LISREL (Jöreskog & Sörbom, 1998) together with Mplus (Muthén & Muthén, 2001) call it weighted least squares (WLS). The main advantage is that the estimator is not dependent on multivariate normality. However, a limited number of variables (recommendation is below 20) and the very large samples that are needed to produce good estimates are the limitations of this approach. For example, a simulation study by Yung & Bentler (1994) suggests more than 2000 observations. In addition, Olsson, Foss, Troye and Howell (2000) show that ADF estimation performs poorly when the model is incorrectly specified.

The third approach to address modeling problems with ordinal non-normal data is the categorical variable model (CVM) developed by Muthén (1993). The model, which is implemented in the Mplus program, uses the general ADF function but without the aforementioned limitations. We consider this a viable frequentistic approach to the analysis of ordinal variables in organizational research.

However, none of the techniques described earlier address the problem of non-linear dependencies between observed variables. In this paper, we argue that the Bayesian modeling approach (see e.g., Bernardo & Smith, 2000; Myllymäki, Silander, Tirri & Uronen, 2002), named after English reverend Thomas Bayes (1702-1761) for his contributions (Bayes, 1763), is a viable alternative to frequentistic statistical techniques addressing all the abovementioned modeling problems. Although the subject domain of this paper is organizational atmosphere (or climate) research, we believe that this discussion applies equally to all other research areas in the social sciences when the measurement level of the indicators is categorical.

The essential benefits of Bayesian modeling are summarized in (Congdon, 2001). Next we discuss more thoroughly the three most important features for this study.
1. The researcher is capable of inputting a priori information to the model. The source of subjective information could be, for example, an interview with an expert on a certain topic, or previously collected data. For example, an adaptive online questionnaire is able to profile respondents with Bayesian probabilistic modeling and thus personalize the total number of questions asked from each person. In this application field a priori profile information, which is applied in the profiling stage when considering the closest match profile for the current person, is gained from earlier responses from a similar population. In this paper, a priori information is needed only when Bayesian networks are developed from data; that is, in the part of the paper where we perform Bayesian dependence modeling (BDM, results are presented in Tables 2 and 3). Prior beliefs are quantified by calculating the equivalent sample size (ESS) (e.g., Heckerman, Geiger & Chickering, 1995) parameter value with equation 1

$$ESS = \frac{|V_1| + |V_2| + \ldots + |V_n|}{2N}$$

where $|V_i|$ denotes the number of values of each variable $V_i$ in the model and $N$ denotes the number of variables $V_i$ in the model (Myllymäki, Silander, Tirri & Uronen, 2002). ESS is a parameter that regulates our behavior when new data is entered; that is, how we update our beliefs when new evidence is presented. If the ESS value is small, new evidence has greater impact on our beliefs than if the parameter value is large. Equation 1 is applied in this paper for two reasons: First, equal priors are set to all variables in the model, as we have no reason to favor any single variable. Second, the ESS parameter calculation is related to so-called Jeffrey’s prior that is commonly used as a non-informative prior in Bayesian analysis.

2. Bayesian modeling is designed to analyze discrete categorical variables. Organizational researchers collect some of their data with paper and pencil or web-based online surveys. The most typical question types in survey research are dichotomous and multiple-choice questions. In both cases, the categories are discrete (e.g., have no overlap and are mutually exclusive) and exhaust the possible range of responses. One of the major differences between traditional Gaussian and Bayesian models lies in the fact that the latter does not require a multivariate normal distribution of the indicators (e.g., observed variables) or underlying phenomena. This feature is especially useful for a researcher who collects her data with, for example, Likert-scale (DeVellis, 2003, pp. 78-80) questions as the response options from 1 to 7 produce data that is more qualitative than quantitative in nature. The measurement level of such an item is ordinal and it is not advisable to model it with traditional statistical analysis that rely on the concept of normal distribution, and require the calculation of mean and standard deviation.

3. Bayesian modeling is able to analyze both linear and non-linear dependencies between variables. Phenomena under investigation are seldom purely linear or continuous in nature. Unfortunately most commonly applied traditional linear Gaussian models (e.g., regression and
factor analysis) are statistically inadequate for understanding non-linear dependencies between variables. Bayesian dependence models for discrete data allow the description of non-linearities as Bayesian theory gives a simple criterion, the probability of the model, to select among such models.

A major drawback with the Bayesian approach at this moment is that only a few applications are capable of analyzing latent variable models. Examples of such software are BUGS\textsuperscript{1}, which studies Bayesian inference using Gibbs sampling (see Congdon, 2001, 2003), and DSIGoM\textsuperscript{ii}, which is based on the grade of membership analysis. As the major goal of this paper is to investigate the number of non-linear and multi-modal relationships in real-life organizational research data sets, we limit the investigation to dependencies between observed variables and use bivariate correlations (both $r_p$ and $r_s$) and Bayesian dependence modeling (B-Course\textsuperscript{iii}) to reach that goal.

The research questions in this study are: (1) What kind of and how many non-linearities are captured by discrete Bayesian networks? (2) Is there a difference between the results of linear bivariate correlations and Bayesian dependence modeling? (3) Does an empirical sample containing pure linear dependencies have better overall fit indices in CFA than a sample containing less linear dependencies? (4) Does an empirical sample containing pure linear dependencies have higher CFA parameter estimates than a sample containing less linear dependencies? (5) Is there a difference between the substantive interpretations of the results of CFA and BDM with linear and non-linear samples?

2. Theoretical Background

2.1 Growth-oriented Atmosphere

Ongoing learning and self-development by employees are critical to the mission of any modern organization. In order to be successful, educational organizations must provide effective professional development programs for employees over the entire course of their careers (Lawler, 1994). Professional development includes all developmental functions that are directed at the maintenance and enhancement of professional competency. In the modern world, updating is ideally a continual, lifelong process that addresses such goals as the acquisition of new and up-to-date information, the development of skills and techniques and the elevation of one’s personal esteem. The maintenance and enhancement of competency is subject to the combined effect of many factors, ranging from personal traits to the salient features of the work environment (Fishbein & Stasson, 1990).

Research has shown that important factors in the development of a growth orientation are support and rewards from the management, the incentive value of the job itself, the operational capacity of the team and work related stress (Argyris, 1990; Dubin, 1990; Hall, 1990; Kaufman, 1990; Nokelainen & Ruohotie, 2003; Ruohotie, 1996).
Management and leaders face challenges such as how to develop and reward learning, empower people, support the development of professional identity, create careers based on interaction, set goals for learning and plan development, evaluate learning and its development and create commitment to the job and the organization.

The incentive value of the job depends on the opportunities it offers for learning; that is, developmental challenges, the employees’ chance to exert influence, opportunities to learn collaboratively and the dignity of the job.

The operational capacity of a team or a group can be defined by its members’ capability to operate and learn together, by the work group co-operation and by the reputation for effectiveness.

Work related stress might become an obstacle to professional growth if a heavy mental load and demand for continual alterations causes stress for people and thus suppresses organizational growth and development.

Ruohotie and Nokelainen (2000) examined the theoretical dimensions of a growth-oriented atmosphere (GOA) in a Finnish vocational education high school. The organization consisted of ten geographically separate units. The sample size was 318 employees, 66 per cent out of the survey population of 479 employees. The target population was Finnish vocational high school personnel in 1998 ($N = 7,958$).

The instrument utilized in the study contained 80 statements. The response options on a 5-point Likert scale varied from 1 (strongly disagree) to 5 (strongly agree).

Ruohotie and Nokelainen (2000) constructed fourteen summated scales (Hair, Anderson, Tatham & Black, 1995, p. 9) to represent the theoretical dimensions of a growth-oriented atmosphere. The scales were formed on the basis of both theoretical aspects of growth-orientation (Ruohotie, 1996) and the results of exploratory factor analysis (Maximum likelihood with Varimax rotation). The fourteen-factor solution was the most parsimonious, representing 67 per cent of the variance within 80 items. Eigenvalues were between 1.05 and 23.98. Respondents indicated only moderate differences in preferences for various dimensions as mean ratings ranged between 3.2 and 3.8. Internal consistency for each factor was estimated with Cronbach's alpha coefficient (1970, pp. 160-161). The alpha values ranged from .77 to .93.

Although the authors report continuous parameters such as mean and alpha on items measured with the nonmetric ordinal scale, we consider the results plausible as the underlying phenomenon, a growth-oriented atmosphere, is continuous by nature. The ration of the sample size to the number of observed variables in the Ruohotie and Nokelainen study (2000) was acceptable according to empirical and simulation studies (e.g., Cattell, 1978; Gorusch, 1983; MacCallum, Widaman, Zhang & Hong, 1999).

Ruohotie and Nokelainen (2000) found that a growth-oriented atmosphere generates togetherness and influences developing leadership. Multidimensional scaling provided evidence that
factors representing the incentive value of the job, commitment to work and organization, clarity of
the job and growth motivation are the strongest indicators of growth-oriented atmosphere. They
reached the following conclusions on the basis of their research findings: 1) Teacher’s professional
growth-motivation reflects directly with task value on teacher-pupil relationships and on
achievement motivation; (2) Task value has an effect on a growth-oriented atmosphere; 3) Growth-
oriented atmosphere is strongest in work assignments that offer challenging professional tasks
(manager, teacher) and lowest among other workers.

2.2 Bayesian Modeling Approach
In this paper, we study two different kinds of non-linearities among the items measuring the
fourteen factors of the GOA: 1) Non-linear relationships between continuous variables and 2) multi-
modal relationships between continuous variables.

The term ‘non-linear’ is not very informative since it seems to include many different
dependence patterns between random variables. As a mathematical concept, linearity (of a
mapping) is well defined. In statistics, when describing the relationship between two variables as
linear, we usually assume that the mean of the variables is a linear function of the means of some
other variables (possibly in some special context, e.g., when certain variables are fixed). However,
there are many situations in which describing only these linear relationships of variables misses the
important aspects of dependencies. The most self-evident shortcoming is that the dependencies
between variables may be very non-linear. Moreover, if some variables are measured on a nominal
scale, the concept of linearity is not meaningful at all. When using ordinal scales, linearity (based
on means) is also often considered dubious. It may also be a conceptually misleading notion even if
the dependencies are mathematically linear. For example, in the case of multi-modality the
relationship between means may be linear, but the mean of the dependent distribution may lie in a
low probability region (i.e., values close to the mean are rare).

Discrete Bayesian networks operate on a nominal finite scale. Thus it is trivial that these
networks are capable of modeling this type of non-linearity. Any dependence between variables,
one of which is measured on a nominal scale, is non-linear. Consequently, non-linearity due to the
scale is not studied in this paper. However, it is worth noting that when data contains nominal scale
variables that are not totally independent of all the other variables of the data, Bayesian networks
are capable of modeling non-linearities.

Mathematically, linearity is well defined between two sets of continuous variables. However, in
this paper, we only study simple non-linear relationships. In our study, the dependence between
variables X and Y is considered non-linear if the mean of the conditional distribution of Y is not a
monotonic (i.e., continuously increasing or decreasing) function of X. Similarly, the dependence
between variables $X$ and $Y$ is considered multi-modal if the mode of the conditional distribution of $Y$ is not a monotonic function of $X$.

This study resembles to some extent the work by Hofmann and Tresp (1996, 1998) in which they use the method of Parzen windows to allow non-linear dependencies between continuous variables. The purpose of their work was to demonstrate the possibility of building Bayesian networks that can capture non-linear relationships. By using discrete variables this is rendered trivial, but our objective is to find out to what extent this is possible in other situations; that is, how many and what kind of non-linearities are captured by discrete Bayesian networks.

Given the identically and independently distributed multivariate data set $D$ over variables $V$ and the prior probability distribution $\pi$ over Bayesian networks (Pearl, 1998), Bayesian probability theory allows us to calculate the probability $P(G \mid D, \pi)$ of any Bayesian network $G$ (Heckerman, Geiger and Chickering, 1995). Different networks can then be compared by their probability. Finding the most probable Bayesian network for any given data is known to be NP-hard (Non-deterministic Polynomial-time hard) which means that the automatic discovery of the most probable network is a mission impossible (Chickering, 1996). An example of an NP-hard problem is the ‘subset sum problem’ (e.g., Cormen, Leiserson & Rivest, 1996, p. 951): Given a set of integers, does any subset sum exactly to zero? For example, given the set {$-5, -3, 1, 2, 9$}, the answer is YES because the subset {$-3, 1, 2$} sums to zero. Fortunately stochastic search methods have proven to be successful in finding high probability networks (Chickering, Geiger & Heckerman, 1995). Once the network $G$ has been constructed using data $D$, we can use it to calculate predictive joint distributions $P(V \mid G, D)$. Bayesian network structure can be used to effectively calculate the conditional marginals of the predictive joint distribution for single variables, $P(V_i \mid A, G, D)$, where $A$ is any subset of the variables of $V$. In this paper we study only the marginals, where $A$ is a singleton {$V_j$} and there is either an arrow from $V_i$ to $V_j$ or an arrow from $V_j$ to $V_i$ (we say that $V_i$ and $V_j$ are adjacent in $G$).

The Bayesian dependence network (Heckerman, Geiger & Chickering, 1995; Myllymäki, Silander, Tirri & Uronen, 2002; Silander & Tirri, 2000) is a representation of a probability distribution over a set of random variables, consisting of an directed acyclic graph (DAG), where the nodes correspond to domain variables, and the arcs define a set of independence assumptions which allow the joint probability distribution for a data vector to be factorized as a product of simple conditional probabilities. A graphical visualization of the Bayesian network contains two components: (1) observed variables visualized as ellipses and (2) dependences visualized as lines between nodes. A variable is considered to be independent of all other variables if there is no line attached to it. Such networks (see Tables 2 and 3) are calculated in this paper using the aforementioned B-Course software (Myllymäki, Silander, Tirri & Uronen, 2002). We have shown
in our earlier research that Bayesian networks are useful for the exploratory analysis of statistical relationships between observed variables (see e.g., Ruohotie & Nokelainen, 2000).

3. Method

3.1 Sample and Procedure
The sample \( n = 726 \) was collected during the year 2001 with a 69-item web-based self-report questionnaire that is a revised version of the previous one (Nokelainen, Ruohotie & Tirri, 2002). The instrument had a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The data consists of adult employees from three Finnish vocational high schools (D21, \( n = 447 \); D22, \( n = 71 \); D23, \( n = 208 \)). Respondents had three different kinds of job profiles (with 4% missing data, \( n = 31 \)): Managers (6%, \( n = 46 \)), teachers (61%, \( n = 462 \)), and administrative personnel (29%, \( n = 223 \)). The nature of a respondent’s contract was categorized into three classes (with 3%, \( n = 26 \) missing data): Established (70%, \( n = 533 \)), temporary (22%, \( n = 169 \)), and part-time (5%, \( n = 34 \)) employees.

3.2 Measuring Non-linearities
To measure non-linear dependencies captured by Bayesian networks, every variable was tested in each network by conditioning it one by one with its immediate neighbors in the network. It was observed whether the modes and means of the conditional distributions were linear and whether the conditional distributions were unimodal. Linearity of modes and means was tested by recording whether the means and modes were increasing or decreasing functions of the conditioning variable. Even clear departures from line-like behavior were accepted as linear as long as the direction of correlation (positive, negative) did not change. Therefore, in these experiments, a ‘linear’ relationship is one that can be more or less adequately modeled by line describing how the central tendency of the dependent variable varies as a function of the independent variable. In measuring the unimodality of conditional distributions, we judged the dependence to be unimodal if (and only if) none of the conditional distributions \( P(Y|X) \) were clearly multimodal. We acknowledge the possible presence of Simpson’s Paradox (e.g., Moore & McCabe, 1993, p. 190) when detecting non-linear relationships with sign changes, but prefer to label it as a statistical fact. The reversal of the direction of a comparison or an association when data from several groups are combined to form a single group becomes a paradox only when associated with dubious causal interpretation.
4. Results

4.1 Research question 1: What kind of (and how many) non-linearities are captured by discrete Bayesian networks?

Research evidence based on Bayesian network modeling of six independent empirical data describing the dimensions of the GOA showed that only 22 per cent of all dependencies were purely linear; that is, linear mode, linear mean, unimodal. This is the best data for traditional linear analysis as no information is lost due to non-linearity. The total number of non-linear dependencies in the data was 57 per cent. Sixteen per cent of dependencies were purely non-linear (non-linear mode, non-linear mean, multimodal). Multimodality was the most common violation of linearity in all samples. The results show that Bayesian networks capture non-linear dependencies in the data, as 46 per cent of the pairwise (unconditional) dependencies of the models are vaguely, and 23 per cent severely, non-linear.

4.2 Research question 2: Is there difference between the results of linear bivariate correlations and Bayesian dependence modeling?

To answer the second research question, we compared the results of linear correlational analysis and non-linear Bayesian network modeling. Our hypothesis was that the results of both linear and non-linear analysis should be the same if the level of non-linearity in the sample has no practical effect. We used in this stage of the analysis only D21 (n = 447) and D23 (n = 208) samples. Those two samples were chosen as the sample sizes of the D21 and D23 data are more appropriate for the analysis than D22 (n = 71). The D21 data represents in this analysis a linear sample, as 24 per cent of its dependencies are unimodal and have a linear mode and mean. The D23 data is more non-linear in nature as it has five per cent less similar pure linear dependencies.

We begin by analyzing the fourteen GOA factors sum correlations of the linear D21 data with the Spearman rank order method ($r_s$) and mean correlations with Pearson’s product moment method ($r_p$). Because the results of both correlational analyzes proved to be alike, we study only the Pearson product moment correlations here.

A comparison of the Bayesian dependence modeling (BDM) and correlation solutions is presented in Table 1. The left hand side shows the visualization of the network where nodes represent variables and arches represent dependencies between them. The strength of each dependence on the model is indicated with a color; a darker color indicates a stronger statistical relationship between the two variables. Importance ranking corresponding to the color of the arcs in the final model is presented in the middle part of the table. The column on the right hand side contains the results of the correlation matrix. BDM shows nine strong and five weaker relationships between the GOA factors.

-- Insert TABLE 1 about here --
Team spirit (TES) is the most important variable in the model as it has a direct statistical relationship to Growth motivation (GRM), Community spirit (COS), Valuation of the job (VAL) and Developing of know-how (DEV) factors. The role of the Valuation of the job (VAL) factor is also important as it is a connecting node to the Rewarding of know-how (REW) factor, which in turn is connected to the Encouraging leadership (ENC) factor. If leaders encourage their subordinates, they feel more commitment to work and organization, the incentive value of the job increases and they feel less psychic stress.

The Bayesian dependence network model is generally congruent with the correlation matrix as both methods found the same two factors, namely the Build-up of work requirements (BUI) and the Students attitudes towards the teacher (STA), independent of all the other factors. Third factor not belonging to the model is the Growth motivation (GRM) as it has only a weak relationship to the Team spirit (TES) factor. Non-linear modeling found nine strong dependencies between the factors as the correlational analysis found five. However, the results were almost identical, as all but two of the high correlation dependencies on the matrix were also present in the Bayesian model. The missing dependencies were between the Valuation of the job (VAL) and Developing of know-how (DEV) factors, \( r(447) = .718, p < .01 \), and the Rewarding of know-how (REW) and Developing of know-how (DEV) factors, \( r(447) = .650, p < .01 \). However, BDM suggests that the dependence between the two factors and the Developing of know-how (DEV) factor is mediated by the Encouraging leadership (ENC) factor. This finding, as it is repeated later in this paper with the non-linear sample, could be interpreted as superior enabling her subordinates’ motivation and commitment to work. The non-linear model provides new information by revealing the relationship between the Team spirit (TES), Valuation of the job (VAL) and Rewarding of know-how (REW) factors. The Bayesian model indicates that encouraging leadership has stronger effect on the clarity of the job factor than on the strategic leadership factor. The correlation matrix supports this finding as it shows higher positive correlation between the Encouraging leadership and Clarity of the job factors, \( r(447) = .709, p < .01 \).

Next we compare the correlations and BDM of the non-linear data D23. The commitment to work and organization (COM) is the central factor in the model presented in Table 2. The factor has direct connections to the Psychic stress of the job (PSY), the Incentive value of the job (INV), the Clarity of the work role (CLA), Team spirit (TES) and Encouraging leadership (ENC) factors. Closer examination of the frequency distributions (not presented in Table 2) shows, that employees at the vocational high school have high commitment to work and organization, team spirit and valuation of the job. Employee’s responses show that they are disappointed in the rewarding of know-how and their work roles are not explicit. As with the linear data D21, the correlation matrix shows high correlation, \( r(208) = .699, p < .01 \), between the Valuation of the job (VAL) and
Developing of know-how (DEV) and Rewarding of know-how (REW) factors. Again, the BDM shows that the dependency between the two variables is mediated by the Encouraging leadership (ENC) factor.

-- Insert TABLE 2 about here –

4.3 Research question 3: Does an empirical sample containing pure linear dependencies have better overall fit indices in CFA than sample containing less linear dependencies?

We tested the GOA model fit to both samples, linear D21 and non-linear D23, with confirmatory (restricted) factor analysis. Our hypothesis was, that linear data should fit the model better than the data that contains more non-linear dependencies.


Table 3 shows the model fit indices of the confirmatory factor analysis. The first section of the table presents measures of absolute fit that determine the degree to which the model predicts the observed correlation matrix (Hair et al., 1995, p. 683). The root mean square error of approximation (RMSEA) is designed to evaluate the approximate fit of the model in the population (Kaplan, 2000, p. 112). The estimate was in both samples below the fair fit level of .08 (Hair et al., 1995, p. 685), indicating good fit (Browne & Kudeck, 1993). The upper limit of the 90 per cent confidence interval was also above the cutoff value in both samples. The standardized root mean square residuals (SRMR) help the investigator to examine how well the aspects of the data are captured by the model (Loehlin, 2004, p. 70). SRMRs were in both samples below a cut-off value of .08 (Hu & Bentler, 1999).

The second section of Table 3 presents incremental fit measures that compare the proposed model to a baseline model that all other models should be expected to exceed (Hair et al., 1995, 685). The Tucker-Lewis index (TLI), a.k.a. the Non-normed Fit Index (NNFI), and a similar measure, the comparative fit index (CFI), were both slightly below the recommended level of .90 (Tucker & Lewis, 1973) in both samples.

The results indicate that the GOA model was performing slightly better with the linear D21 sample than non-linear D23 sample. We also tested the model with CFA implemented in Mplus that uses ADF function (weighted least squares) instead of ML and allows categorical indicators (Muthén &
Muthén, 2001). The fit indices for the linear sample were also slightly better than for the non-linear sample with this method.

-- Insert TABLE 3 about here --

4.4 Research question 4: Does an empirical sample containing pure linear dependencies have higher CFA parameter estimates than sample containing less linear dependencies?

The factor covariances of the GOA model presented in Table 4 show that the linear sample (D21) has higher overall parameter estimates and smaller error variances than the non-linear sample. Next we will discuss the differences in strength of dependencies between the two samples in more detailed manner as we interpreted the results via four different theoretical aspects.

-- Insert TABLE 4 about here --

4.5 Research question 5: Is there difference between substantive interpretations of the results of CFA and BDM with linear and non-linear samples?

Bollen (1989, 281) states “nonsense results for individual parameters can occur in conjunction with good overall fit measures …”. We examined both Bayesian dependence models (see Tables 2 and 3) and component fit measures (see Table 4) to see if the estimates between factors make sense according to the GOA model. We will focus on the following four aspects of the model: 1) Support and rewards from management have a positive influence on how an employee experiences her know-how as being rewarded and developed and her work as being valued; 2) The incentive value of the job has a positive influence on the development of know-how and valuation of the job; 3) The operational capacity of the team and team spirit correlate positively; 4) Work-related stress hinders the development of all the other factors of the GOA.

The first aspect is that support and rewards from management are essential in the development of growth orientation (Ruohotie, 1996). The correlational analysis shows that both the Valuation of the job (VAL) and Development of know-how (DEV) factors are connected to the Know-how rewarding (REW) factor. The relationship is also present in corresponding Bayesian models (see Tables 2 and 3), but in both samples the models include the Encouraging leadership (ENC) as a mediating component between the two factors. The CFA covariance matrix presented in Table 4 shows strong positive covariances between the Encouraging leadership (ENC), Know-how rewarding (REW), Know-how developing (DEV), Valuation of the job (VAL) and Clarity of the job (CLA) factors. The parameter estimates range between 0.752 and 1.116 (linear sample D21), and 0.634 and 0.798 (non-linear sample D23). Also Bayesian models for both data contain the aforementioned relationships.
The second aspect is that the incentive value of the job depends on the opportunities it offers for learning; that is, challenging work and high valuation of the work (Ruohotie, 1996). Factor covariances presented in Table 3 support the GOA theory in both samples (D21 $\phi_{INV,DEV} = 0.537$; D21 $\phi_{INV,VAL} = 0.585$; D23 $\phi_{INV,DEV} = 0.518$; D23 $\phi_{INV,VAL} = 0.537$). The Bayesian model for the linear sample (D21) shows only partial support for the second theoretical assumption as there is a connection between the Incentive value of the job (INV) and the Know-how developing (DEV) factors, but no direct connection exists between the Valuation of the job (VAL) factor and the two other factors (Table 1). The non-linear sample (D23) does not show any evidence that supports the GOA theory (Table 2). In both Bayesian models the Encouraging leadership (ENC) factor acts as a mediator between the variables under investigation.

The third aspect is the relationship between the Community (COS) and Team spirit (TES) factors. According to Argyris (1992) and Ruohotie (1996), community members should discuss developmental aspects of their work and learn from each other. Factor covariances support the GOA theory in both samples: $\phi_{COS,TES} = 0.691$ (D21), and $\phi_{COS,TES} = 0.534$ (D23) (Table 5). The dependence between the two factors is also present in both Bayesian dependence models (Tables 2 and 3).

The fourth aspect is that work related psychic stress (PSY) hinders people from giving their best performance in the work, and may thus become an obstacle to professional growth (see e.g., Edwards & Rothbard, 1999; Ruohotie, 1996). Table 4 shows that factor covariances between the Psychic stress of the job (PSY) and the other factors are stronger in the linear sample, as the parameter estimates range between –0.219 and –0.547 (linear sample D21), and –0.138 and –0.319 (non-linear sample D23). The highest (negative) covariances in both samples are between PSY and Commitment to work and organization (COM). The dependence between PSY and COM is also present in both Bayesian dependence models, but no other support for the fourth theoretical aspect is present (Tables 2 and 3).

The results are summarized in Table 5. The first observation is that both Bayesian dependency models do not support the second theoretical assumption about the relationship between Incentive value of the job (INV) and Know-how developing (DEV) and Valuation of the job (VAL). However, INV and VAL have a statistical dependence also in BDM derived from the linear sample (Table 1). The second observation is that the fourth theoretical assumption about the negative influence of Psychic stress (PSY) on all the other factors is only partially supported in both Bayesian models. Finally, the linear sample (D21) has in most cases higher CFA parameter estimates than the non-linear sample (Table 4). This is a natural finding because only linear factor covariances between variables are estimated.

-- Insert TABLE 5 about here --
5. Discussion

This study investigated the number of linear and non-linear dependencies between the items measuring fourteen dimensions of Growth-oriented Atmosphere. An empirical sample represented employees of three Finnish vocational high schools \((n = 726)\). Bayesian theory was discussed and Bayesian dependence models for discrete data were introduced as a model family capable of describing non-linearities. Next, empirical data was analyzed to find out if non-linear dependencies weaken the robustness of bivariate linear statistical methods (represented by correlation analysis) when compared with non-linear modeling (represented by Bayesian dependency modeling).

Investigation of empirical data \((n = 726)\) showed that only 22 per cent of all dependencies between variables were purely linear (linear mode, linear mean, unimodal). Sixteen per cent of dependencies were purely non-linear (non-linear mode, non-linear mean, multimodal). Multimodality was the most common reason for the violation of linearity in both data sets.

Investigations were continued with two sub samples of the vocational high school data, namely D21 \((n = 447)\) and D23 \((n = 208)\). The D21 sample represents in this study linear empirical data with 24 per cent of pure linear and 15 per cent of pure non-linear dependencies and the D23 sample represents non-linear data with only 16 per cent of pure linear dependencies and 18 per cent of pure non-linear dependencies.

The subject domain interpretations of linear correlational analysis and non-linear Bayesian dependence modeling (BDM) were compared to learn if the results would lead to different subjective interpretations. The results showed that in general Bayesian network models were congruent with the correlation matrixes as both methods found the same variables independent of all the other variables. However, non-linear modeling found with both linear and non-linear samples a greater number of strong dependencies between the GOA factors. Comparison of the correlations and dependencies in Bayesian networks showed, that in both samples linear correlations indicated a direct connection between know-how rewarding, know-how developing and valuation of the job, as Bayesian models indicated indirect connections between the variables with encouraging leadership acting as a mediator between them.

Further, we focused on the following four aspects of the GOA theory because our motivation was to investigate if there were differences between the results of linear (CFA) and non-linear (BDM) analysis with the linear and non-linear samples: 1) Support and rewards from the management have a positive influence on how an employee experiences her know-how as being rewarded and developed and her job as being valued; 2) The incentive value of the job has a positive influence on the development of know-how developing and valuation of the job; 3) The operational capacity of the team and team spirit correlate positively; and 4) Work-related psychic stress hinders the development of the GOA.
Results showed that the analysis techniques produced similar results for two out of four theoretical aspects, namely the first and third. Different results leading to different substantive interpretations were considered for the second and fourth theoretical aspect as follows. The BDM was able to find only partial support from the linear sample, and no support at all from the non-linear, sample for the assumption that the incentive value of the job would have a positive influence on the development of know-how developing and valuation of the job. Both Bayesian dependency models suggested that the components under investigation are not directly related, but instead indirectly connected to each other via encouraging leadership. The fourth theoretical aspect was supported in both linear analyses, as the Psychic stress of the job factor was negatively related all the other factors. However, in both Bayesian dependency models the PSY factor was related only to commitment to work and organization. Finally, linear methods (e.g., bivariate r and CFA) found stronger statistical relationships between factors measuring the Growth-oriented Atmosphere from a linear than non-linear sample. We fully agree with Grilli and Rampichini (2004) when they state that use of a proper model is always a desirable feature of the analysis because we may expect the resulting inferences to be generally more reliable.

Notes

i BUGS is available at http://www.mrc-bsu.cam.ac.uk/bugs/

ii DSIGoM is available at http://www.dsisoft.com/grade_of_membership.html

iii B-Course is available at http://b-course.cs.helsinki.fi/obc/

iv Growth-oriented Atmosphere Questionnaire (GOAQ) is available at http://www.uta.fi/aktkk/goaq/

References


<table>
<thead>
<tr>
<th>Table 1</th>
<th>Bayesian Network Model of the Dimensions of Growth-oriented Atmosphere (Linear Sample, n = 447)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Network Model</strong></td>
<td><strong>Dependence</strong></td>
</tr>
<tr>
<td></td>
<td>TES-&gt;COS</td>
</tr>
<tr>
<td></td>
<td>ENC-&gt;CLA</td>
</tr>
<tr>
<td></td>
<td>ENC-&gt;COM</td>
</tr>
<tr>
<td></td>
<td>TES-&gt;VAL</td>
</tr>
<tr>
<td></td>
<td>VAL-&gt;REW</td>
</tr>
<tr>
<td></td>
<td>VAL-&gt;ENC</td>
</tr>
<tr>
<td></td>
<td>ENC-&gt;DEV</td>
</tr>
<tr>
<td></td>
<td>CLA-&gt;STR</td>
</tr>
<tr>
<td></td>
<td>COM-&gt;PSY</td>
</tr>
<tr>
<td></td>
<td>REW-&gt;ENC</td>
</tr>
<tr>
<td></td>
<td>DEV-&gt;INV</td>
</tr>
<tr>
<td></td>
<td>COM-&gt;INV</td>
</tr>
<tr>
<td></td>
<td>TES-&gt;DEV</td>
</tr>
<tr>
<td></td>
<td>TES-&gt;GRM</td>
</tr>
</tbody>
</table>

**Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed).**

*Note.* ENC = Encouraging leadership, STR = Strategic leadership, REW = Know-how rewarding, DEV = Know-how developing, INV = Incentive value of the job, CLA = Clarity of the job, VAL = Valuation of the job, COS = Community spirit, TES = Team spirit, PSY = Psychic stress of the job, COM = Commitment to work and organization, GRM = Growth motivation.
Table 2
Bayesian Network of the Dimensions of Growth-oriented Atmosphere (Non-linear Sample, n = 208)

<table>
<thead>
<tr>
<th>Network Model</th>
<th>Dependence</th>
<th>Probability ratio</th>
<th>$r_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENC-&gt;DEV</td>
<td>1:Inf.</td>
<td>.742**</td>
<td></td>
</tr>
<tr>
<td>ENC-&gt;VAL</td>
<td>.723**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COM-&gt;CLA</td>
<td>.611**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ENC-&gt;REW</td>
<td>.665**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COM-&gt;ENC</td>
<td>.636**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLA-&gt;STR</td>
<td>1:1.000.000</td>
<td>.582**</td>
<td></td>
</tr>
<tr>
<td>COM-&gt;INV</td>
<td>.693**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLA-&gt;COS</td>
<td>1:58948</td>
<td>.500**</td>
<td></td>
</tr>
<tr>
<td>COS-&gt;TES</td>
<td>1:7695</td>
<td>.565**</td>
<td></td>
</tr>
<tr>
<td>COM-&gt;TES</td>
<td>1:34</td>
<td>.586**</td>
<td></td>
</tr>
<tr>
<td>COM-&gt;PSY</td>
<td>1:1.7</td>
<td>-.373**</td>
<td></td>
</tr>
</tbody>
</table>

** Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed).

Note. ENC = Encouraging leadership, STR = Strategic leadership, REW = Know-how rewarding, DEV = Know-how developing, INV = Incentive value of the job, CLA = Clarity of the job, VAL = Valuation of the job, COS = Community spirit, TES = Team spirit, PSY = Psychic stress of the job, COM = Commitment to work and organization.
### Table 3

*Model Fit Indices of the Growth-oriented Atmosphere Model*

<table>
<thead>
<tr>
<th>Sample</th>
<th>$\chi^2$</th>
<th>df</th>
<th>$\chi^2/df$</th>
<th>p</th>
<th>RMSEA</th>
<th>C.I. (90)</th>
<th>SRMR</th>
<th>TLI</th>
<th>CFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>D21$^a$</td>
<td>4448.113</td>
<td>1375</td>
<td>3.235</td>
<td>.000</td>
<td>.071</td>
<td>.068</td>
<td>.073</td>
<td>.835</td>
<td>.853</td>
</tr>
<tr>
<td>D23$^b$</td>
<td>2897.437</td>
<td>1375</td>
<td>2.107</td>
<td>.000</td>
<td>.073</td>
<td>.069</td>
<td>.077</td>
<td>.810</td>
<td>.831</td>
</tr>
</tbody>
</table>

$^a$ Linear sample D21 $n = 447$. $^b$ Non-linear sample D23 $n = 208$.

*Note. RMSEA = Root Mean Square Error of Approximation with 90 per cent confidence interval. TLI = Tucker-Lewis coefficient. CFI = Comparative Fit Index.*
### Table 4
**Factor Covariances of the Growth-oriented Atmosphere Model**

*Data D21 (n = 447)*

<table>
<thead>
<tr>
<th></th>
<th>f01_enc</th>
<th>f02_str</th>
<th>f03_rew</th>
<th>f04_dev</th>
<th>f05_inv</th>
<th>f06_cla</th>
<th>f07_val</th>
<th>f08_cos</th>
<th>f09_tes</th>
<th>f11_psy</th>
<th>f13_com</th>
</tr>
</thead>
<tbody>
<tr>
<td>f01_enc</td>
<td>Encouraging leadership</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f02_str</td>
<td>Strategic leadership</td>
<td>0.417</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f03_rew</td>
<td>Know-how rewarding</td>
<td>0.752</td>
<td>0.500</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f04_dev</td>
<td>Know-how developing</td>
<td>0.871</td>
<td>0.365</td>
<td>0.651</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f05_inv</td>
<td>Incentive value of the job</td>
<td>0.543</td>
<td>0.212</td>
<td>0.377</td>
<td>0.537</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f06_cla</td>
<td>Clarity of the job</td>
<td>0.886</td>
<td>0.490</td>
<td>0.647</td>
<td>0.707</td>
<td>0.423</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f07_val</td>
<td>Valuation of the job</td>
<td>1.116</td>
<td>0.391</td>
<td>0.720</td>
<td>0.827</td>
<td>0.585</td>
<td>0.799</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f08_cos</td>
<td>Community spirit</td>
<td>0.510</td>
<td>0.288</td>
<td>0.384</td>
<td>0.506</td>
<td>0.323</td>
<td>0.501</td>
<td>0.552</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f09_tes</td>
<td>Team spirit</td>
<td>0.541</td>
<td>0.274</td>
<td>0.392</td>
<td>0.484</td>
<td>0.333</td>
<td>0.498</td>
<td>0.586</td>
<td>0.691</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>f11_psy</td>
<td>Psychic stress of the job</td>
<td>-0.361</td>
<td>-0.226</td>
<td>-0.277</td>
<td>-0.316</td>
<td>-0.387</td>
<td>-0.363</td>
<td>-0.434</td>
<td>-0.219</td>
<td>-0.240</td>
<td>1.000</td>
</tr>
<tr>
<td>f13_com</td>
<td>Commitment to work and organization</td>
<td>0.686</td>
<td>0.360</td>
<td>0.495</td>
<td>0.570</td>
<td>0.575</td>
<td>0.560</td>
<td>0.725</td>
<td>0.366</td>
<td>0.370</td>
<td>-0.547</td>
</tr>
</tbody>
</table>

*Data D23 (n = 208)*

<table>
<thead>
<tr>
<th></th>
<th>f01_enc</th>
<th>f02_str</th>
<th>f03_rew</th>
<th>f04_dev</th>
<th>f05_inv</th>
<th>f06_cla</th>
<th>f07_val</th>
<th>f08_cos</th>
<th>f09_tes</th>
<th>f11_psy</th>
<th>f13_com</th>
</tr>
</thead>
<tbody>
<tr>
<td>f01_enc</td>
<td>Encouraging leadership</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f02_str</td>
<td>Strategic leadership</td>
<td>0.503</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f03_rew</td>
<td>Know-how rewarding</td>
<td>0.632</td>
<td>0.491</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f04_dev</td>
<td>Know-how developing</td>
<td>0.736</td>
<td>0.509</td>
<td>0.677</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f05_inv</td>
<td>Incentive value of the job</td>
<td>0.462</td>
<td>0.243</td>
<td>0.373</td>
<td>0.518</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f06_cla</td>
<td>Clarity of the job</td>
<td>0.684</td>
<td>0.600</td>
<td>0.587</td>
<td>0.633</td>
<td>0.366</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f07_val</td>
<td>Valuation of the job</td>
<td>0.798</td>
<td>0.388</td>
<td>0.661</td>
<td>0.736</td>
<td>0.537</td>
<td>0.563</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f08_cos</td>
<td>Community spirit</td>
<td>0.367</td>
<td>0.369</td>
<td>0.388</td>
<td>0.460</td>
<td>0.207</td>
<td>0.585</td>
<td>0.419</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f09_tes</td>
<td>Team spirit</td>
<td>0.351</td>
<td>0.269</td>
<td>0.306</td>
<td>0.384</td>
<td>0.268</td>
<td>0.383</td>
<td>0.402</td>
<td>0.534</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>f11_psy</td>
<td>Psychic stress of the job</td>
<td>-0.218</td>
<td>-0.184</td>
<td>-0.186</td>
<td>-0.179</td>
<td>-0.138</td>
<td>-0.341</td>
<td>-0.221</td>
<td>-0.304</td>
<td>-0.250</td>
<td>1.000</td>
</tr>
<tr>
<td>f13_com</td>
<td>Commitment to work and organization</td>
<td>0.569</td>
<td>0.383</td>
<td>0.487</td>
<td>0.613</td>
<td>0.602</td>
<td>0.586</td>
<td>0.604</td>
<td>0.390</td>
<td>0.440</td>
<td>-0.319</td>
</tr>
</tbody>
</table>

*Note.* Data D21 error variances range between 0.040 and 0.091. All p-values were $p < .001$. Data D23 error variances range between 0.055 and 0.106. All p-values range between $p < .001$ -.022.
**Table 5**  
*Summary of the Results based on Theoretical Assumptions*

<table>
<thead>
<tr>
<th>Theoretical assumptions</th>
<th>D21&lt;sup&gt;a&lt;/sup&gt;</th>
<th></th>
<th></th>
<th>D23&lt;sup&gt;b&lt;/sup&gt;</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Support and rewards from the management</td>
<td>ENC (+) → REW, VAL, DEV</td>
<td>●&lt;sup&gt;e&lt;/sup&gt;</td>
<td>☷&lt;sup&gt;b&lt;/sup&gt;</td>
<td>BDM&lt;sup&gt;c&lt;/sup&gt;</td>
<td>CFA&lt;sup&gt;d&lt;/sup&gt;</td>
<td>●&lt;sup&gt;e&lt;/sup&gt;</td>
</tr>
<tr>
<td>2. The incentive value of the job</td>
<td>INV (+) → DEV, VAL</td>
<td>○&lt;sup&gt;f&lt;/sup&gt;</td>
<td>●&lt;sup&gt;e&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Operational capacity of the team</td>
<td>COS (+) ↔ (+) TES</td>
<td>☷&lt;sup&gt;g&lt;/sup&gt;</td>
<td>☷&lt;sup&gt;g&lt;/sup&gt;</td>
<td></td>
<td></td>
<td>●&lt;sup&gt;e&lt;/sup&gt;</td>
</tr>
<tr>
<td>4. Work-related stress</td>
<td>PSY (-) → All the other factors</td>
<td>○&lt;sup&gt;f&lt;/sup&gt;</td>
<td>☷&lt;sup&gt;g&lt;/sup&gt;</td>
<td></td>
<td></td>
<td>○&lt;sup&gt;f&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

---

<sup>a</sup> Linear sample D21 n = 447.  
<sup>b</sup> Non-linear sample D23 n = 208.  
<sup>c</sup> Bayesian Dependence Model.  
<sup>d</sup> Confirmatory Factor Analysis.  
<sup>e</sup> Research evidence supports the theoretical assumption.  
<sup>f</sup> Research evidence supports the theoretical assumption partially.  
<sup>g</sup> Research evidence supports the theoretical assumption strongly.  
<sup>h</sup> Research evidence does not support the theoretical assumption.

*Note.* ENC = Encouraging leadership, REW = Know-how rewarding, DEV = Know-how developing, INV = Incentive value of the job, VAL = Valuation of the job, COS = Community spirit, TES = Team spirit, PSY = Psychic stress of the job.