Chapter 10
Beyond the subject-matter level: The use of the domain-specific knowledge state test and knowledge profiles

1 Introduction

Previous chapters have dealt with investigations into the structure of the PKS at the subject-matter level. In this chapter we investigate the structure of the PKS along a variety of dimensions. Our discussion in relation to the structure of the PKS introduces a new approach towards the 'structure of knowledge' problem. In the theoretical part of this text, we discuss this new and distinct approach which is based on extensive analysis of theories, models and practice-based strategies found in literature. This base is exploited to define a set of 'dimensions' that might be helpful to construct 'knowledge profiles'. Four types of dimensions are designated: cognitive-psychological dimensions, educational-psychological dimensions, psychometrical dimensions and content-based dimensions.

In the empirical part of this chapter, we provide the reader with data that support the relevance and validity of the knowledge profile dimensions. Two approaches will be adopted, based on data gathered during an investigation involving a large sample of university students. First, we analyze the extent to which the parameters along the dimensions give information about different components of the PKS. Second, we analyze the discriminatory power of the knowledge profile dimensions to make apparent PKS differences between a variety of student sub-populations.

2 Capturing the structure of the prior knowledge state

The quality and impact of the PKS has been a major issue in our research about the role of prior knowledge at university level in the previous chapters. An important conclusion from this research body indicates that it looks promising to analyze in more detail the complex of components of the PKS. A first and promising attempt in this direction focused on the structure of the PKS along the content dimension (chapter 9). Note that we use the term 'component' as an operationalization of a parameter. Components refer to these subparts of the prior knowledge state that relate to a specific parameter along a knowledge profile dimension, as our knowledge state tests did refer to at the content dimension.

Three important concepts are presented in the former paragraph: the PKS, components of the PKS and a complex of components. These three concepts refer to the value attached to the specific 'structure' of the PKS. Our earlier research revealed that such a structure could be indicated in prior knowledge, e.g. along the content dimension. We did for instance discriminate between the optimal requisite prior knowledge state and the mathematics prior knowledge state. But it was also suggested that the differentiation of components of the PKS along other dimensions could be helpful to diagnose and support educational practice.

3 The structure of knowledge issue

The structure of knowledge issue has been dealt with by a variety of theoretical disciplines: cognitive psychology, educational psychology, artificial intelligence, etc. From a pragmatic point of view, the issue has also been of prime importance in applied sciences like instructional psychology, curriculum development theories and psychometrics.

Disciplines like cognitive psychology, educational psychology, artificial intelligence, etc. have - from their points of view - highlighted the 'structure of knowledge' resulting in a puzzling variety of approaches, focuses, models, theories, research attempts (Ausubel (1968), De Groot (1946), Mayer (1979), Reigeluth and Stein (1983)). We give a short outline of some of their specific theoretical contributions. Other relevant information can be found in chapter 2.

An early, cognitive-theoretical approach appears in the work of Ausubel (1968) who argues that new knowledge is only acquired to the extent that it is meaningfully related to existing knowledge. Ausubel maintains that knowledge is organized primarily in a hierarchical fashion, which implies that mastery of higher knowledge levels assumes mastery of all lower knowledge levels. Additionally, Ausubel advances that the various pieces of information integrated within a particular knowledge structure are highly interrelated. Thus, the more structured prior knowledge, the more flexible and easy the acquisition of new knowledge becomes.

Ausubel's conceptualization of learning as the assimilation into prior knowledge, is echoed and extended in Mayer's (1979) schema theory. New knowledge is - according to Mayer - assimilated into a hierarchy of progressively more specific content within the learner's knowledge base. Thus, the basic learning process can be described as the assimilation of new knowledge within hierarchically ordered schemata.

Another benchmark is set by the 'elaboration theory'. According to this theory, multiple access avenues become available to the learner by the activation of alternate relational paths . This theory also depended on the assumption that knowledge acquisition is facilitated to the extent that information is organized in a hierarchically integrated mode (Reigeluth and Stein, 1983).

These theories are important since they stress the 'structured' and hierarchical nature of prior knowledge, but they all lack empirical support to ground their validity (Reigeluth and Stein, 1983). Additional support, especially for the idea of the hierarchical nature of the knowledge organization need to be found.

From a rather pragmatic, e.g. instructional-psychological, point of view the structure-of-knowledge paradigm needs to be investigated in detail in order to find more efficient ways for using instructional technology. Our research into ways of utilizing prior knowledge indicate e.g. that the different components of the PKS should be taken into account (e.g. at the subject-matter level) and that components of the PKS along other dimensions could be helpful in educational settings for diagnoses and as a basis for intervention.

If we summarize the variety of theoretical and pragmatic approaches we found in literature, four main types of dimensions to structure knowledge can be conceptualized:

Content related dimensions
Cognitive-psychological dimensions
Educational-psychological dimensions
Item characteristic dimensions

In the next parts of this text, these dimensions will be recapitulated and discussed in more detail. But more important is our attempt to reveal some new insight into the hierarchical nature and structure of knowledge. As indicated earlier this will be the main issue in the empirical part of this text.

It should also be noted that - when putting forward these theories and approaches towards prior knowledge - we adopt an information processing view towards learning (Sternberg, 1985a & 1985b; see chapter 1). This is inspired by our dynamic approach towards knowledge acquisition; a departure which is also advocated by information processing theory.

4 Knowledge profiles: definition

As such, the concept of 'knowledge profiles' is not found in literature. Only 'student profiles' (Wolf, et al., 1991) and 'cognitive profile' (Letteri et al., 1982) have some similarity in meaning. This is certainly the case for the studies by Letteri et al. (1980, 1982), described below.

Our use of the concept 'profile' is derived from the practice, common in educational research, of plotting as a graph (or profile) the raw or standardized scores of a group or individual on certain parameters (Keeves, 1988). In analyzing research data, comparisons are made between persons or groups of persons in terms of a set of measurements about related aspects. For each person or group a profile can be obtained by combining the results of the set of parameters. The comparison between individual profiles is known by the generic term 'profile analysis'.

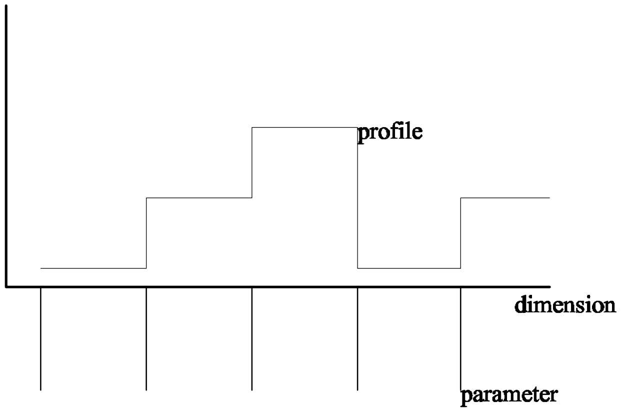

Figure 1: Example of a profile

Figure 1 shows the relationships between some key concepts. A 'dimension' is the basis to construct a knowledge profile. Each dimension represents an approach towards the structure of knowledge since it introduces a related set of PKS components. These components are named 'parameters'.

From an instructional-psychological point of view, knowledge profiles can give practical indications of student achievement and learning, thereby making it possible to direct the learning process. In a recent overview of student assessment, Wolf et. al. (1991) advocated this approach. According to these authors, there is a need for new educational psychometrics in line with the much changed perceptions of educational achievement. The educational world must come to terms with new premises, multiple paths towards the prior knowledge state, more developmentally-oriented assessments and the students entering school with widely varying backgrounds. For our part, we take account of these changes by trying to identify the components of the prior knowledge state, by implementing prior knowledge state tests and by setting in motion plans to use these tests several times a year to monitor students' progress.
In this context it is necessary to come to an agreement on the relevant parameters for describing student performance and it is critical to develop ways of looking at 'student profiles': " unless we develop these kinds of differentiated portraits of student performance within a domain, it is difficult to envision student assessment ever informing, rather than merely measuring, the educational process" (Wolf, et. al., 1991).

As suggested above, our concept 'knowledge profiles' is related - to a certain degree - to Letteri's use of the concept. His work at the Center for Cognitive Studies, University of Vermont, focuses on the development of an individual's cognitive profile. A profile in Letteri's conception is based on a continuum along which a variety of cognitive dimensions are put one next to the other. Letteri combines up to seven dimensions such as scanning (focus), breadth of categorization and cognitive complexity. An individual's cognitive profile is the diagram, the picture that results after positioning individual scores in relation to each of these dimensions along the continuum (Letteri, 1980).

According to Letteri and Kuntz (1982) very high correlations between an individual's cognitive profile and its performance of intellectual tasks, the ability to learn and school performance have been detected. The results of the Letteri studies are quite amazing. His cognitive profiles can for instance help to separate seventh and eight grade students into significantly different achievement levels; can account for as much as 87 per cent of the variance in post-test scores; and predict those scores at a level of .05 or better.

Moreover, the results show that a cognitive profile is a basic determinant of academic achievement and can accurately help to identify specific learning deficits that contribute significantly to low academic achievement. Cognitive profiles seem to be reliable predictors of low/high academic achievement (Letteri, 1980).

Although the Vermont studies provide evidence for the potential of cognitive profiles, some critical remarks are needed.

First, the work at Vermont concentrates rather on 'cognitive styles' and 'characteristics of cognitive functioning'. This is particularly obvious if we look in more detail to the dimensions used to construct the cognitive profiles. So there is a clear distinction between their and our approach. Our knowledge profile dimensions clearly consist of parameters referring to structure-aspects of knowledge. Second, the Letteri profiles seem not to be appropriate for adult learners because they are based on theories and research concerning child development. Third, we perceive profiles as diagrams based on a single dimension along which the parameters are clearly interrelated. The use of a set of such dimensions results in the construction of multiple knowledge profiles (i.e. one for each dimension). Letteri constructs only one profile, based on a variety of dimensions. His major focus is on the correlation between this profile and school performance. As a consequence, the remedial power of his profiles remains restricted. The Letteri profile is therefore to be considered as an instrument to differentiate groups performing below or above average, taking into account the perceived correlations between the relative positioning on the profile and external measures of school performance. His remedial method is thus not based on the specific overall profile structure but on related external measures of school performance. Finally, it is suspicious that there is not more recent work on Letteri's profiles.

In contrast, our multi-profile approach generates 'profile analysis', an application of multivariate analysis of variance (MANOVA) in which several dependent variables are measured on the same scale (Tabachnick and Fidell, 1989). This profile analysis can provide relevant information with diagnostic and remedial value. In this way, the profiles help us to identify learning deficits to be remediated.

5 Theoretical dimensions

In this part, we review a representative sample of dimensions and parameters currently found in literature. Each of these dimensions is based on a specific model or theory of knowledge structuring, which will only be discussed in short. As suggested in part 3 of this chapter, the 'structure of knowledge' issue has been dealt with from a large variety of viewpoints, resulting in a quartet of dimensions.

A first group of dimensions is classified according to common models of the economics domain. Other sets of dimensions are based on theories of knowledge representation, knowledge structure, learning theories, text representation models and psychometric theory.

A first question in relation to each dimension is whether they are applicable as structures to find components of the PKS. Secondly - and this will be discussed in the empirical part of this text - are these dimensions useful to differentiate groups of students. In this way we can scrutinize the descriptive, explanatory and remedial prospect of our knowledge profiles.

5.1 Content dimensions

Economics subdomain dimensions

Content is one of the most important dimensions for categorizing domain-specific knowledge. Classification based on the parameter 'subdomain' involves dividing the economics domain into several widely-accepted subject areas. One possible way of subdividing the domain is based on the curriculum of the University of Maastricht, containing nine parameters.

1. Reporting
2. Financing
3. Organization
4. Marketing
5. Macro-economics
6. Micro-economics
7. Public finances
8. International economic affairs
9. Behavioural and social sciences

Curriculum level dimension

Students are expected to have mastered certain aspects of a science by certain points in their study. These points are here referred to as the curriculum levels (first and second year). A team of economists have helped to define these levels. These curriculum levels are subsequent, but too broad to be supposed hierarchical.

1. First year level
2. Second year level

Curriculum accent dimension

Within economics it is common to differentiate between two main streams, representing a different accent, i.e. general economics and business administration on the one hand and quantitative economics on the other hand.

1. General economics and business administration
2. Quantitative economics

5.2 Cognitive-psychological dimensions

Propositional dimension

Knowledge representation as used in schema theories (see chapter 2) takes certain propositions or nodes as a starting point. A proposition is the smallest unit that can be qualified as true or false. According to most schema theories there are six kinds of nodes: Physical State (PS, statement that refers to an ongoing state in the physical or social world), Physical Event (PE, statement that refers to a state change in the physical or social world), Internal State (IS, statement that refers to an ongoing state of knowledge, attitude, or belief in an individual), Internal Event (IE, refers to a state change in knowledge, attitude or belief in an individual), Goal (G, statement that refers to an achieved or unachieved state that a person wants) and Style (S, statement that refers to details about the style or manner in which an action or event occurred).

Further examples and elaborations of these parameters are given in Dochy and Bouwens (1990c). These nodes are used in the representational theory of Graesser (1981) to represent knowledge as a network of labelled statement nodes that are interrelated by directed arcs (see further). As such, the nodes do not have any hierarchical relationship.

1. PS
2. PE
3. IS
4. IE
5. G
6. S

Node relation dimension

The 'Node relation dimension' is based on characteristics of the interrelations between propositions (see propositional dimension). Relations between propositions can be classified as node relation or arc parameters: Reason (R, a Goal node is a reason for another Goal node), Initiate (I, a State or Event initiates another Goal node), Consequence (C, a State, Event or Goal node that has the consequence of another State or Event node), Manner (M, an Event or Goal node occurs with some

style), Property (P, a person, object or entity has some property that is a State node) (see also Dochy and Bouwens, 1990e).
These arc parameters are not of a hierarchical nature. The overview on the next page shows what specific relations between the propositions have been identified.

1. G - G REASON
2. PS - G INITIATE IS - G PE - G IE - G
3. PS - PE CONSEQUENCE IS - PE PE - PE IE - PE G - PE PS - PS IS - PS PE - PS IE - PS G - PS
4. PE - S/G MANNER IE - S/G GE - S/G
5. PS - PS PROPERTY

Cognitive complexity dimension

McDaniel (1991) proposes five levels of cognitive complexity. These parameters are designed to measure thinking processes by determining the cognitive complexity apparent in written interpretations of complex situations.

Level 1: Unilateral Descriptions
The situation is simplified. It focuses on one idea or argument. Alternatives are not identified. No new information, meaning, or perspectives are brought in. 'Good-bad' and 'either-or' assertions are made. Appeals to authority or simple rules. Information is simply paraphrased, restated or repeated.

Level 2: Simplistic Alternatives
Simple and obvious conflicts are identified, but the conflicts are not pursued or analysed. Develops a position by dismissing or ignoring one alternative and supporting the other with assertions and simple explanations rather than through deeper assessment of the situation.

Level 3: Emergent Complexity
More than one possible explanation or perspective is identified. Complexity is

established and preserved. New elements are introduced. Supports position through comparisons and simple causal statements.

Level 4: Broad Interpretations
Broad ideas help define and interpret the situation. Ideas within the perspective established are manipulated. There is a clearly recognizable explanatory theme. Ideas are integrated into 'subassemblies' each supporting a component of the explanation.

Level 5: Integrated Analysis
The situation is restructured or reconceptualized and the problem is approached from a new point of view. A network of cause-and-effect relationships is constructed. Ideas are integrated and extrapolated. One arrives at new interpretations by analogy, application of principles and generalizations. An organizing framework is constructed, connections are given and consequences are predicted.

1. Unilateral descriptions
2. Simplistic alternatives
3. Emergent complexity
4. Broad interpretations
5. Integrated analysis

5.3 Educational-psychological dimensions

Discussing the first two dimensions of this type, i.e. the behavioural and the content dimensions, it is helpful to mention Component Display Theory (CDT, Merrill, 1983) which makes use of related concepts. CDT can be described as a set of prescriptive relationships that can be used to guide the design and the development of learning activities. One of the basic assumptions of CDT is that there is more than one type of learning and more than one kind of memory structure. Primary aspects of CDT are objectives, learning activities and tests. According to CDT, all objectives or test items can be classified in cells of a matrix, based on a content and a performance dimension. The CDT content dimension distinguishes 'facts, concepts, principles and procedures'. The CDT performance dimension differentiates 'remember, use and find'. This is also conform to Gagné's tripartite: 'information, skills and strategies'. CDT holds that this performance-content matrix can be considered as a taxonomy, thus suggesting a hierarchical base for the two determining dimensions in the matrix.

Behavioural dimension

The much used distinction between 'declarative and procedural knowledge' can be further operationalized as 'to know, to understand, and to apply'. These concepts are considered as equivalent to recognition and reproduction (to know) and production (to apply) (De Corte, 1976). As shown below, 'understanding', also called 'insight' is intermediate category (Keeves, 1988). The concepts can also be related to the classification: Appreciation, recognition and reproduction of information (=declarative) or production or applications (interpretative, convergent, divergent or evaluative production) which can be viewed as procedural (Keeves, 1988). The three parameters do also correspond with the basic taxonomic levels proposed by several educationalists such as Bloom, Guilford, De Corte and De Block (Keeves, 1988). Most researchers agree that these parameters are of a hierarchical nature which has also been supported by empirical evidence (Keeves, 1988). Research suggests also that there is some justification for treating Bloom's lower levels as being taxonomic. This should not be the case for the levels 'synthesis' and 'evaluation' (Madaus, et. al., 1973).

1. Know 2. Understand	1. Recognition, Reproduction	1. Appreciation, recognition, and reproduction of information	1. Declarative
3. Apply	2. Production	2. Interpretative convergent, divergent or evaluative production	2. Procedural

Content dimension

At the content dimension we can differentiate five parameters: facts, concepts, relations, structures and methods. This is in accordance with the work of Guilford where he distinguishes several product parameters, and the work of other authors (Keeves, 1988). These parameters are widely accepted as being hierarchical (Keeves, 1988).

1. Facts
2. Concepts
3. Relations
4. Structures
5. Methods

Epistemological dimension

Following Brachman (1985), we may differentiate between five levels of knowledge representation. These parameters can also be considered as the most current combinations between behaviour - and content parameters, as given between brackets: knowledge identification (identifying facts and concepts), knowledge conceptualization (insight in concepts), epistemological analysis (to know and understand relations and structures), logical analysis (to know and understand methods), implementational analysis (application of methods). These levels are considered as hierarchical since they are a combination of the hierarchical behaviour and content dimension.

1. Knowledge identification
2. Knowledge conceptualization
3. Epistemological analysis
4. Logical analysis
5. Implementational analysis

Layers of knowledge dimension

We may differentiate between: static knowledge (description of concepts and relations), knowledge of different types of inferences, knowledge representing elementary tasks (procedures) and strategic knowledge (Clancey, 1983).
The first layer contains the static knowledge of the domain: domain concepts, relations and complex structures, such as models of processes or devices.
The second layer is the inference layer. In this layer we describe what inferences can be made on the basis of the knowledge in the static layer. Two types of entities are represented at the inference layer: meta-classes and knowledge sources. Meta-classes describe the role domain-concepts can play in reasoning. For example, a domain concept like infection can play the role of a finding in a consultation process, but it may also play the role of hypothesis. Knowledge sources describe what types of inferences can be made on the basis of the relations in the domain layer. Examples are the specification and the generalization of knowledge sources, which both make use of a subsumption relation in the domain layer. The third layer is called the task layer. At this level the basic objects become goals and tasks. Tasks are ways in which knowledge sources can be combined to achieve a particular goal. The fourth layer is the strategic layer in which knowledge resides which allows a system to make plans (i.e. create a task structure), control and monitor the execution of tasks, diagnose when something goes wrong and repairing impasses.

1. Static knowledge layer
2. Inference layer
3. Task layer
4. Strategic layer

5.4 Item characteristics dimensions

Although these dimensions are of a completely different nature, they are in the context of our research purposes of importance. Items are used to measure the mastery of (prior) knowledge. Moreover, test items are clues to the activation of prior knowledge. The way the individual is instigated to show his mastery of knowledge and the way certain knowledge is (re)presented to the learner can also be related to the 'structure of knowledge issue'.

Number of propositions dimension

A proposition is the smallest unit that can be considered as a separate statement that can be judged as true or false. In schema theories (Dochy and Bouwens, 1990c), propositions or nodes have a central function in the structure of schema. It is assumed that the amount of propositions determines the degree of structure needed to answer a question correctly. Three parameters have been identified in relation to this dimension:

1. < 5 propositions
2. > 4 < 10 propositions
3. > 9 propositions

Information level dimension

The stem of an item is the general information which is given and which must not be evaluated as true or false. This correct information precedes the question(s) for which this information should be taken into account. A stem can be connected to one or more subsequent questions. Therefore, the spatial and logical distance between the general information part of an item and the question part of the whole item is larger than for simple items without stem.

1. Items with a stem
2. Items without a stem

Representation level dimension

Adhering the representation structure, used in the research of Boekaerts (1979), i.e. visual, verbal and symbolic representation, we used four parameters along this dimension. These parameters are also closely related to the four content levels in Guilford's structure-of-intellect model: figural, symbolic, semantic (the verbal factor) and behavioural (nonverbal information) and the Twyman (1985) categories: verbal, pictorial and schematic. Since test-items are always - in part - based on a textual representation of information, our dimension distinguishes only between parameters that are combinations of knowledge representation.

1. Textual-graphical
2. Textual
3. Textual-schematic
4. Textual-symbolic

5.5 Hierarchical and non-hierarchical dimensions

Of importance in this theoretical discussion of the dimensions for the construction of knowledge profiles is the hierarchical or non-hierarchical nature of the dimensions. Empirical validation of the dimension and their further application, has to take this particularity into account. Table 1 presents a summary.

Table 1: Hierarchical and non-hierarchical dimensions

Profile dimensions	Hierarchical

Economics subdomains dimension	-
Curriculum level dimension	-
Curriculum accent dimension	-
Propositional dimension	-
Node relation dimension	-
Cognitive complexity dimension	+
Behavioural dimension	+
Content dimension	+
Epistemological dimension	+
Layers of knowledge dimension	+
Number of propositions dimension	-
Information level dimension	-
Representation level dimension	+

6 Looking for empirical evidence: Data collection, data screening and general validation procedure

6.1 Data Collection

When looking for empirical evidence to support the theoretical assumptions and theoretical constructs in the former part of this text, research data were gathered from a sample. The DS KST (see chapter 7) was administered to 627 students of the University of Maastricht and the Dutch Open University, starting studying 'economics'. The 154 items were analyzed **-** separately **-** by the researchers. In reviewing the items, the researchers attempted to classify each item on each one of the 13 dimensions.

Table 2: Applying profile dimensions: applicability in practice

Profile dimensions	applica-bility
Economics subdomains dimension	+
Curriculum level dimension	+
Curriculum accent dimension	+
Propositional parameters dimension	-

Node relation dimension	+
Cognitive complexity dimension	-
Behavioural dimension	+
Content dimension	+
Epistemological dimension	+
Layers of knowledge dimension	-
Number of propositions dimension	+
Information level dimension	+
Representation level dimension	+

The project team categorized aIl items according to the parameters of these dimensions. In relation to three dimensions the researchers encountered too many difficulties en were not applicable in practice (table 2):

- The 'propositional parameters' dimension could not be applied since all items consist of more than one node. But this dimension was helpful as a base to determine the 'node relation' parameter of an item.
- The parameters along the 'cognitive complexity' dimension were too vague and not defined at an operational level to be applied consistently. Moreover, they implied a reformulation of the items, which was not acceptable.
- The 'layers of knowledge' dimension was felt to be a duplicate of the 'behavioural dimension'.

In relation to the other dimensions an inter-rater reliability was obtained > .8**. If there was discussion in relation to the categorization of a specific item along a dimension, discussion resulted in a consensus on the final evaluation of the item.

6.2 Data screening

After screening and reorganizing the items, the raw item scores were used as the base for calculating new subscores for each parameter along each dimension. To be able to compare the parameter subscores, mean % scores were calculated. This helped to define the data set in table 3. This table presents an overview of the mean % scores, the standard deviation and the minimum and maximum score in relation to each parameter on each dimension. The last column indicates the number of test items on which the calculation of the parameter scores is based. In association to most parameters, the score is based on a sufficient number of test items. Only the 'node relation' and the 'curriculum accent' dimension present some problems since the numbers of items is not sufficiently equilibrated over all the parameters of these dimensions.

Table 3: Descriptive data of all the parameters for each dimension

	m	σ	min.	max.	n
Economics	42.85	19.14	.00	100.00	154
Reporting	25.63	14.69	.00	83.33	18
Financing	26.37	15.63	.00	72.22	18
Organization	34.50	18.58	.00	83.33	18
Marketing	35.74	16.80	.00	83.33	18
Macro-economics	25.66	14.67	.00	72.00	25
Micro-economics	24.44	14.89	.00	80.00	25
Public finances	32.93	18.50	.00	90.91	11
Int. econ. affairs	26.91	17.99	.00	81.82	11
Behavioural sciences	17.38	17.80	.00	80.00	10
Level 1	30.64	12.63	.00	65.69	102
Level 2	22.31	15.91	.00	78.85	52
Quantitative	24.36	13.77	.00	73.33	15
General	28.20	12.80	.00	66.91	139
Reason	27.91	24.42	.00	100.00	2

Initiate	20.45	18.17	.00	87.50	8
Consequence	29.09	13.21	.00	72.41	58
Manner	39.38	23.77	.00	100.00	5
Property	26.93	12.65	.00	69.14	81
Know	29.82	15.37	.00	79.17	24
Insight	28.95	13.38	.00	71.72	99
Apply	22.67	11.96	.00	70.97	31
Facts	22.43	19.62	.00	100.00	6
Concepts	35.25	16.67	.00	80.95	21
Relations	30.84	14.57	.00	75.00	32
Structures	26.48	13.52	.00	78.57	56
Methods	24.11	12.92	.00	76.92	39
K Ident.	29.22	16.18	.00	80.00	15
K Concept.	36.75	19.85	.00	90.00	10
Epistem.	28.02	13.39	.00	72.62	84
Logical	30.27	17.67	.00	86.67	15
Implem.	22.40	11.89	.00	70.00	30
Propos 1	27.18	12.63	.00	67.07	82
Porpos 2	13.63	6.08	.00	31.00	44
Propos 3	6.93	4.18	.00	22.00	28
With	29.92	13.45	.00	68.00	105
Without	39.32	19.67	.00	100.00	49
T.-graphical	28.77	13.25	.00	67.68	99
Textual	36.58	20.71	.00	88.89	9
T.-schematic	28.71	16.26	.00	71.43	14
T.-symbolic	22.47	12.68	.00	68.75	32

Considering the data in the table, some preliminary conclusions can be drawn in relation to the value of the dimensions and parameters to specify components of the PKS:

-The mean % scores for the specific parameters along a dimension show remarkable differences.

-For some parameters the number of items is low, which can have consequences for reliability.

-Also the σ values vary to a large extent, even when excluding the values based on a too small number of items. The high σ-values indicate that the score distribution does not represent a normal distribution of scores. The fact that the results are based on a prior knowledge state test are responsible for this large score-distribution.

-The fact that for all parameters/dimensions the minimum value is 0, indicates that the test has been difficult for the students. On the other hand, some student are able to show their mastery of economics by obtaining relatively high mean % scores.

Further screening of the data reveals:

-There are no missing data.

-The large standard deviations, mentioned above, might imply that the distribution of the scores is not normal (Skewness, Kurtosis) and that the distributions are influenced by outliers. The Kurtosis-values vary between -.64 and .99. Skewness-values vary between .08 and .88. Both measures suggest acceptable distributions of the data. To check multi-variate normality, box-plots of the data have been screened. They reveal there are extremes and outliers, but their number remains restricted. Moreover, outliers are expected (given the fact we measure mastery of the PKS) and considered part of the particular distribution of our data. They are properly part of the population from which we intend to sample.

-Analysis of the correlation matrix reveals high r-values, but never > .8 which implies that all distinctive variables are non-redundant and do not measure comparable issues[1]. Multi-collinearity is therefore not a problem.

-Homogeneity of variance-covariance matrices is a necessary assumption if we compare the subscores along the different dimensions. Calculation of the Cochrans C and Bartlett-Box F and their respective significance-levels, indicate that this assumption is not violated.

6.3 Validation Procedure

When validating the different knowledge profile dimensions, two approaches will be adopted:

-an analysis of the extent to which the parameters along the dimensions give information about different components of the PKS;

-an analysis of the discriminatory power of the dimensions to make apparent the PKS differences between student sub-populations.

In part 7 and 8 of this chapter, the first approach is adopted, whereas in part 9, the profile dimensions are used to compare a variety of student-subpopulations. In view of the first approach, a distinction is to be made between hierarchical and non-hierarchical dimensions. Hierarchical dimensions imply consecutive intercorrelations between the parameters. This can be evaluated by applying multiple linear regression techniques. Non-hierarchical dimensions imply low correlations between the dimension parameters. This is to be evaluated by analysis of the correlation matrix.

7 The value of the non-hierarchical dimensions

7.1 Economics subdomains dimension

As described in part 5 of this chapter, the economics-subtopics reflect a rather practical subdivision of the economics domain in view of educational purposes. It is therefore to be foreseen that the mean % scores for the different economics subjects do correlate to a certain extent (all r-coefficients are statistically significant at the

[1] Tabachnik & Fidell (1990, p.87) use a limit of r > .9 to distinguish redundant variables.

1% level)[2].

[2] When indicating significance levels in the next part of this text ** indicates $p < .01$ and * indicates $p < .05$.

Table 4: Correlation Matrix subtopic-dimension

	1	2	3	4	5	6	7	8	9
1. Report	1.00	.63	.58	.56	.43	.48	.38	.35	.44
2. Finance		1.00	.68	.63	.55	.53	.51	.50	.52
3. Organ			1.00	.68	.49	.48	.47	.45	.48
4. Market				1.00	.54	.52	.50	.48	.50
5. Macro					1.00	.66	.57	.57	.51
6. Micro						1.00	.48	.55	.52
7. Public							1.00	.56	.46
8. Internat.								1.00	.47
9. Behav.									1.00

The data in table 4 confirm our expectations. The different economics subtopics are intercorrelated. If we concentrate only on the r-coefficients > .6, we get the following picture:

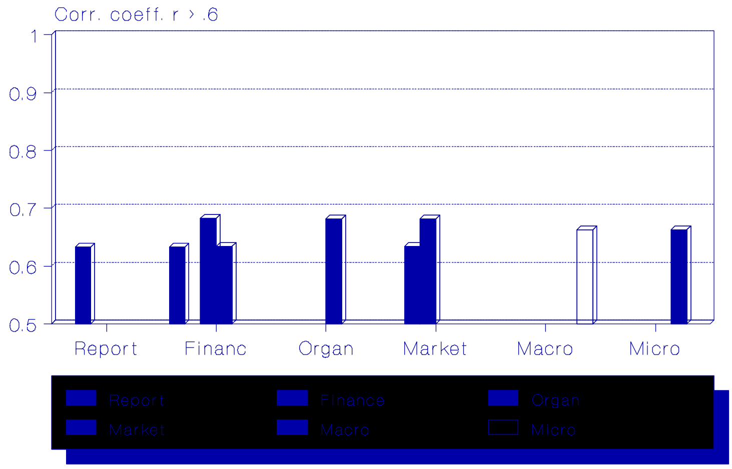

Figure 2: Correlation structure Economics subdomains dimension

The six following subtopics are highly intercorrelated: 'reporting, financing, organization, marketing, macro-economics and micro-economics'. This can be due to high degree of content-links between the economics subdomains or the fact that mastery of a particular subdomain explicitly builds on the mastery of another domain. This is particularly important in relation to the subdomain 'financing' and 'marketing'.

If we execute a linear regression analysis[3] by entering the mean % scores of the

[3] If regression analysis is executed, also the residuals are analyzed in order to detect violations of regression assumptions (linearity, normality of the distribution of the dependent variable, constant variance). If

subdomains 'reporting, macro-economics and organization' stepwise in the regression equation, we find that up to 50.3 % of the variance in the 'financing' is accounted for by these three subdomains. In a comparable way, stepwise regression from 'organization and financing' on 'marketing' reveals that up 51.6 % of the variance in the 'marketing' mean score is accounted for by the scores for 'organization & financing'.
This interdependence can be taken into account when setting up learning activities or guidance initiatives for students.

An alternative interpretation of these correlation coefficients links the mastery of the different economics subdomains to the previous experiences of students with the particular subtopics in their secondary education. A subpopulation of the starting university students might have got an introduction to certain economics domains, such as micro- and macro-economics and reporting (accountancy). The subsequent high scores for these subtopics in the PKS test are therefore expected to correlate to a certain extent.

7.2 Curriculum level dimension

In the prior knowledge test, the test items were subdivided into two course level sets: 'level1' grouping items that evaluate mastery at the first year level and 'level2' that evaluate the mastery at the second year level.
As table 3 shows, the mastery of both the level1 (m = 30.64) and the level2 (m=22.31) items is restricted. As expected, the 'level2' total subscore is the lowest indicating the higher difficulty level of this set of items. Content experts indicate that mastery of level1 items is to a certain degree related to mastery of 'level2' items since part of the 'level2' course content builds on the mastery of 'level1' items. This is confirmed by linear regression analysis. This analysis indicates that the 'level2' mean % score of the students helps to explain 38.2 % of the variance in the 'level2' mean % scores[4].

7.3 Curriculum accent dimension

A generally accepted subdivision in economics is splitting up the domain into general economics and quantitative economics. Our PKS test, does only reflect this subdivision to a very limited extent since only 15 items can be classified as quantitative economics questions whereas 139 items concentrate on the mastery of general economics knowledge. It is therefore predictable that the test results along this dimension will not be very useful. Analysis of the mean scores (m_{quant} = 24.36; $m_{general}$= 28.20) and standard deviations (σ_{quant} = 13.77; $\sigma_{general}$ = 12.80) in table 3

violations are detected, this will be reported.

[4] When executing linear regression analysis, basic assumptions have been checked. No violations have been detected, for instance the distribution of the residuals is normal.

indicate that the students master the two types of economics knowledge to a comparable extent. Moreover, the intercorrelation between both measure is rather high (r =.62**). The subdivision between quantitative and general economics is therefore not able - to a sufficient extent - to separate knowledge components.

7.4 Node relation dimension

The correlations clearly indicate that most parameters along this dimension can be clearly separated from each other. The 'property' parameter in highly correlated with the 'initiate' (.64**) and the 'consequence' (.85**) parameter. It should be noted that

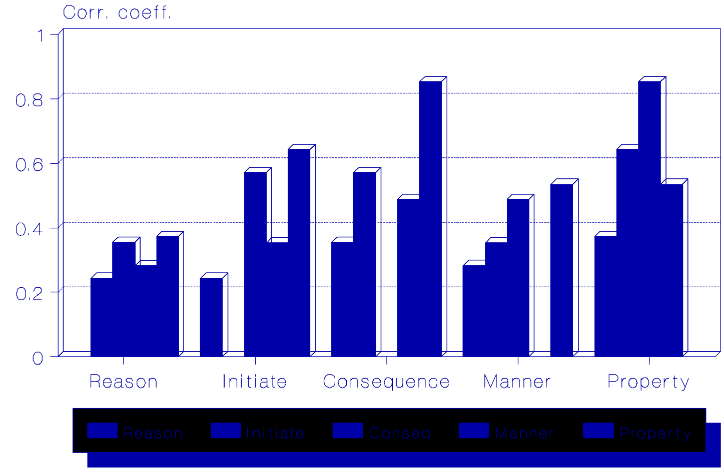

Figure 3: Node relation dimension - correlation structure

7.5 Number of propositions din

Since the number of propositions in test items can be considered as a measure comparable to difficulty levels, a gradual decrease in the mean scores for the three variables on this dimension is expected.

The data in table 3 confirm our expectations ($m_{propos1}$=27.18; $m_{propos2}$=13.63; $m_{propos3}$=6.93). Analysis of the correlation matrix reveals striking results (all r-coefficients are statistically significant):

Table 5: Correlation matrix: number of propositions

	1	2	3
1. < 5 propositions	1.00	.77	.80
2. > 4 < 10 propositions		1.00	.76
3. > 9 propositions			1.00

The results in table 5 can be interpreted as follows: students who are able to solve items, consisting of > 9 propositions (3) are able to solve items consisting of a number of proposition > 5 < 10 (2) and are certainly able of solving items consisting of < 5 propositions (1). This is confirmed by a linear regression analysis: the scores for items with > 9 propositions explain 64.3 % of the variance in the scores of items with < 5 propositions. If we add the scores for items with > 4 < 9 propositions to the regression equation, up to 70 % of the variance is accounted for.

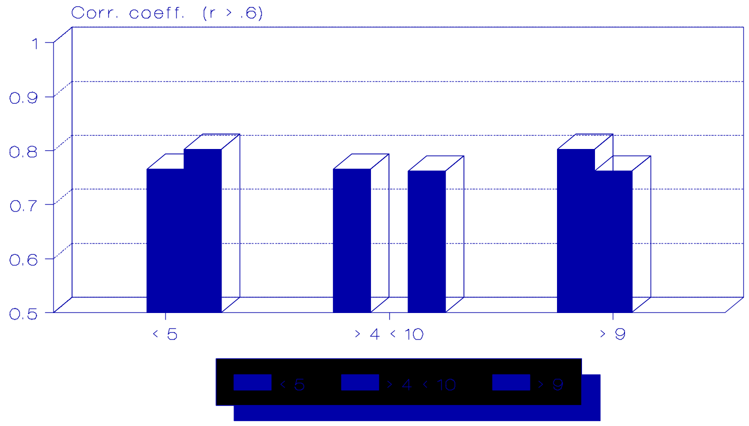

Figure 4: Intercorrelation between mean % scores for 'number of propositions'

7.6 Information level dimension

As explained in the theoretical part of this text, items with a stem have in common that the spatial and logical distance between the basic information and the question part of the item is greater than in items without a stem. To a certain extent, this dimensions can be compared to the 'number of propositions dimension'.

It is expected that items with a stem are more difficult than those without a stem. This is confirmed by the data in table 3 when comparing the mean % scores and σ ($m_{with\ stem}$ = 29.92; $m_{without}$ = 39.32; $\sigma_{with\ stem}$ = 13.45; $\sigma_{without}$ = 19.67). This is also confirmed by the analysis of the correlation matrix. Both variables are highly correlated ($r = .82^{**}$). This can be explained as follows: the students being able to solve items with a stem, are to a very high extent able to solve items without a stem.

Linear regression confirms this hypothesis: the mean % scores for the items with a stem help to explain 68 % of the variance in the scores for items without a stem.

8 Validating hierarchical dimensions

Up to 5 dimensions have been identified as being 'hierarchical' (table 1). As explained in part 6.3 hierarchical dimensions imply consecutive intercorrelations between the parameters along the dimensions. This can be evaluated by applying multiple linear regression techniques. Beforehand, some additional remarks are to be made because we isolate profile dimensions from the economics subdomain dimension. In validating the hierarchical nature of the dimensions abstraction is therefore made of the subject-matter on which the items are based. There is as a consequence no control on interaction effects between the dimensions and the economics subdomains dimension. If e.g. items evaluating the mastery of 'micro-economics' at the 'insight' level are compared to items evaluating the mastery of 'finance' at 'application' level and the latter economics subdomain is much more complex and difficult than the former, the analysis results in relation to the hierarchical structure of 'insight' and 'apply' along the behavioural dimension is disturbed.

8.1 Behavioural dimension

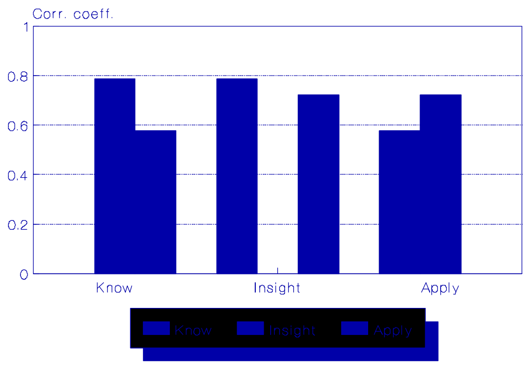

Figure 5: Behavioural dimension - correlation structure

Figure 5 clearly indicates features of the hierarchical nature of this dimension, although the highest correlations are observed in relation to items measuring the mastery of items at insight-level (insight-apply = .72**; insight-know = .79**). The latter fact is consistent with earlier validations of this kind of dimension (see the taxonomic classification of Bloom: knowledge - comprehension - analysis - synthesis and evaluation). Keeves (1988, p.346) refers e.g. to a validation by Ebel, Hill and Horn.

The three lowest levels, 'knowledge, comprehension and application', are comparable to our three first behavioral levels and are found to be hierarchical indeed. Higher up in the hierarchy a branching takes place.

If we extend our analysis to a regression analysis in which we evaluate the interrelation between the three consecutive behavioural levels, we get the following picture:

Table 6: Regression analysis behavioural dimension

Independent variable	Dependent variable	% variance explained
Know	Insight	61 %
Know*, Insight	Apply	52.2 %

When entering the parameters 'know' and 'insight' in the regression equation, the contribution of the parameter 'know' is considered too low[5] and therefore excluded from the regression equation.

The results in table 6 help to confirm the assumptions about the hierarchical nature of this dimension since the preceding parameter(s) always help to explain the major part of the variance in the mean % scores for the subsequent parameter along the dimension.

8.2 Content dimension

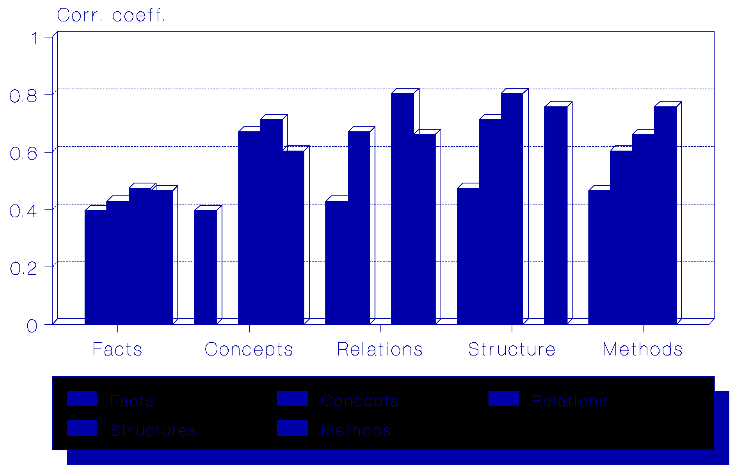

Figure 6: Content dimension - correlation structure

In the picture (figure 6) it can clearly be seen that the correlation coefficients increase along the consecutive parameters of this dimension. This can be considered as a first indicator to support the hierarchical nature of this dimension. Only in relation to the parameter 'methods' there is a minor decrease in the r-values.

[5] The limit of PIN=.05 is reached.

Table 7: Regression analysis content dimension

Independent variables	Dependent variables	% variance explained
(F)acts	Concepts	15.5% F
(F)acts - (C)oncepts	Relations	47.9% C F
(F)acts - (C)oncepts - (R)elations	Structures	71.0% R C F
(F)acts - (C)oncepts - (R)elations - (S)tructures	Methods	59.6% S F C R

The data in table 7 are very consistent. At the consecutive levels, the hierarchy between 'facts - concepts - relation and structures' is respected. In the second column the letters and the order in the letters refer to the pattern in which the independent variables help to explain the variance in the dependent variable. This order always respects the supposed hierarchy.
Only at the methods-level, there is a deviant structure. This suggests that 'methods' might be of a more general and complex nature than 'facts, concepts, relations and structures'.

8.3 Epistemological dimension

As explained earlier, this dimension can be considered as a combination of the behavioral and content dimension. The structure is especially based on the assumption that the lower behavioral parameters are rather linked with the lower content parameters (e.g. knowledge of a fact) and that the higher behavioral parameters are rather linked with the higher content parameters (e.g. application of a theory).

This mixture of two hierarchical dimensions might impose difficulties in terms of their interaction. The correlation matrix (figure 7) is e.g. not helpful to recognize - at first sight - the hierarchical structure.

There are some high correlations between the lower and higher parameters along this dimension. But nevertheless, the highest correlations are observed when the distance a high and another parameter along the dimension is small (implem-logical) and the lowest correlations are observed when the distance between a high and another parameter is large (e.g. 'implem - kident).
A regression analysis in which the parameters are evaluated in terms of their explanatory power for the consecutive parameters gives us the following picture:

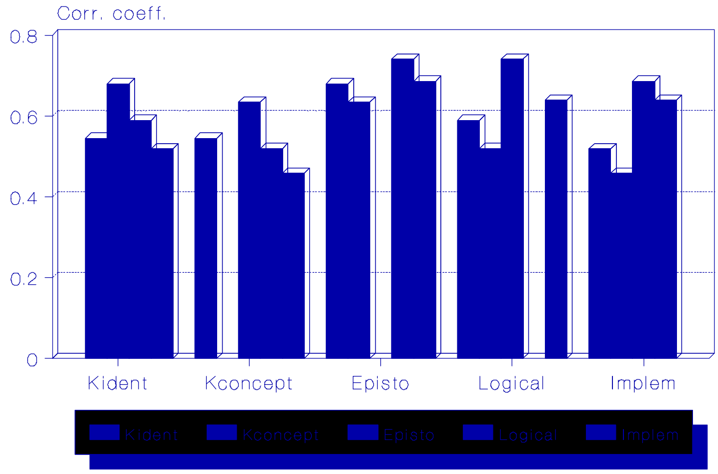

Figure 7: Epistemological dimension - correlation structure

Table 8: Regression analysis - epistemological dimension

Independent variables	Dependent variable	% expl. variance
(Ki)dent	(Kc)oncept	29.7 %
(Ki)dent - (Kc)oncept	(E)pisto	56 % Ki Kc
(Ki)dent - (Kc)oncept - (E)pisto	(L)ogical	56 % E Ki
(Ki)dent - (Kc)oncept - (E)pisto - (L)ogical	(I)mplem	50.7 % E L

The regression analysis confirms the former preliminary conclusions. The hierarchy is disturbed but still present. The proportion of the variance in the consecutive parameters is always explained by the former parameters with the closest distance along the dimension. Moreover, the proportion of explained variance is > 50 %. Lower level parameters are sometimes excluded from the regression equation.

8.4 Representation level dimension

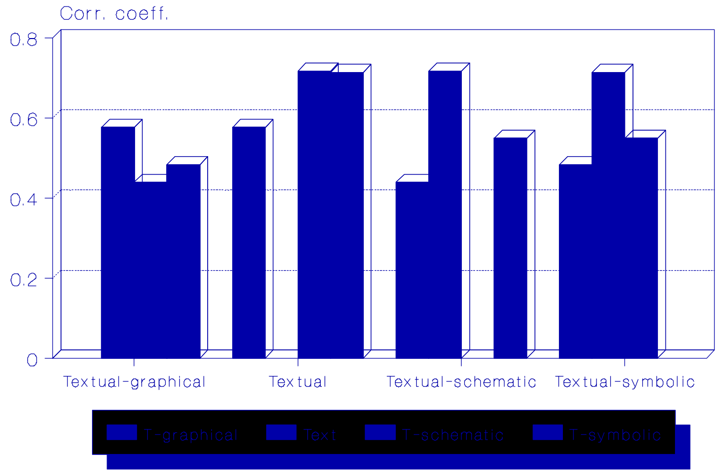

Figure 8: Representation level dimension - correlation structure

Figure 8 reveals high intercorrelations between all parameters and the 'text' parameter. This is to be expected since the test items are never based solely on concrete, schematic or symbolic information. The items are a priori based on textual information which is supported, enhanced, enriched or supplemented with information of another nature. Nevertheless it is remarkable in table 3 that the mean % scores for the subsequent parameters decrease from 36.58 % to 22.47 %, suggesting higher difficulty levels for the items based on higher order representation levels.

Table 9: Regression analysis representation level dimension

Independent variables	Dependent variable	% expl. variance
(C)oncret	(T)ext	33.0 %
(C)oncret - (T)ext	(Sc)heme	51.0 % T
(C)oncret - (T)ext - (Sc)heme	(Sy)mbol	51.6 % T C

Table 9 confirms the expected disturbance of the hierarchical nature by the predominance of the textual representation mode in most items. The 'text' parameter is always the parameter responsible for the highest proportion of the variance in the subsequent parameter.

9 The discriminatory power of the knowledge profile dimensions

In this last part of this chapter, we will try to test the applicability of knowledge profiles. To validate the use of the 10 knowledge profile dimensions, their power to

make explicit differences in the PKS between subpopulations of students has been determined[6]. Three subdivisions in student populations have been researched:
- difference(s) between students studying in differing university contexts.
- difference(s) between students with low and high levels of PKS.
- difference(s) between economics and law students, studying the same course.

For an extensive description of these three studies and the plots of the student knowledge profiles, we refer to Wagemans, Valcke and Dochy (1991a), Valcke and Dochy (1991b) and Dochy and Valcke (1991a,c). We will recapulate the main results, i.e. of the profile analyses, shortly below.

When comparing knowledge profiles from different student populations, a comparison of the mean results for the different parameters along a dimension is not possible using univariate statistics. A univariate analysis of variance (see Dochy and Valcke, 1991c) does not take into account the intercorrelations between the different parameters along the profile dimensions. These intercorrelations are important (although not making the specific variables redundant) and can be explained at the theoretical level as clarified elsewhere (see part 5 of this chapter). A multivariate analysis of variance is needed to refine our analysis and to look for more conclusive information about the differences in prior knowledge state between ES and LS. A multivariate analysis can take these intercorrelations into account. Profile analysis[7] is an extension of multi-variate analysis and is especially appropriate and helpful to evaluate the parameter structure in relation to each profile dimension when comparing subpopulations. Several tests are available in profile analysis. Of principal interest - for our purposes - is the 'parallelism' test which help to answer the question whether the profiles of two subpopulations are parallel or not. At the theoretical level, also the 'flatness' test might be relevant, since this test controls the similarity of responses for the different parameters along one dimension, independent of groups or subgroups. An answer to this question helps to support the validity of the different dimensions since the results indicate whether or not the dimensions/parameters are helpful to specify differences in the mastery of different components of the prior knowledge state. A profile analysis will be performed on the complex of parameters in relation to each dimension.

The grouping variable is different for the three analyses: 'university', 'level of PKS' and 'diploma type'. SPSS-PC$^+$ MANOVA was used for our profile analysis.

For each of the profile analyses, a control of underlying assumptions was done. Profile analysis implies that specific assumptions about the quality of the research data are met (no missing data, comparable sample sizes, (multivariate) normal distributions, no outliers, homogeneity of variance-covariance, multicollinearity).
- data screening to reveal missing data.
- control of sample sizes.

---

[6] research population n=627 (see 6.1, this chapter)

[7] When using profile analysis as a substitute for univariate repeated measures ANOVA, the parallelism test is the test of interaction.

- the evaluation of the homogeneity of variance-covariance matrices is based on the Cochrans C and the Bartlett-Box F test.
- to evaluate assumptions about multivariate normality, boxplots of the mean submeasures for each dimension are screened.
- multicollinearity is tested with the Bartlett test of sphericity.

In general we can conclude that assumptions are met in order to execute a profile analysis on the research data available. More information on this subject can be found in our publications already mentioned.

9.1 The difference(s) between students studying in differing university contexts

Students have - in the Dutch context - the option to choose for a variety of university contexts and educational approaches at university level (e.g. problem centred approach, distance education, etc). A relatively new development (since 1985) in this perspective is the provision of 'open and distance university education' by the Dutch Open University (OU). The question can be put forward whether this new university setting is just another higher education institution enriching the variety of already existing provisions or whether the Open University answers the need of (a) specific student population(s); e.g. second chance, older students, female students, handicapped people, foreign students, post-university students? A way to look for answers to this question is to analyze - by interviews, questionnaires, etc. - demographic variables of the actual student population of the OU.

Another approach goes beyond these surface variables and analyzes in more detail the PKS of the students opting for the OU, making use of our knowledge profiles. The logical research question, which results from this approach is whether the PKS of the students, opting for this study context is different from students studying at regular universities. Making use of the 10 knowledge profile dimensions, an analysis has been executed on research data of student studying economics at the Dutch Open University (OU) and students studying economics at the University of Limburg, Maastricht (UL) where a problem centred approach towards education is adopted.

The main hypothesis of this analysis is: OU students and UL students are different in terms of parameters along the knowledge profile dimensions.

Profile analysis: parallelism test

Table 10 presents an overview of the results for the parallelism test. This helps to answer the question whether the two different student groups have parallel or non-parallel profiles. This is commonly known as the test of parallelism and is the primary question addressed by profile analysis. Figure 9 represents the knowledge profiles of both university populations for the content dimension:

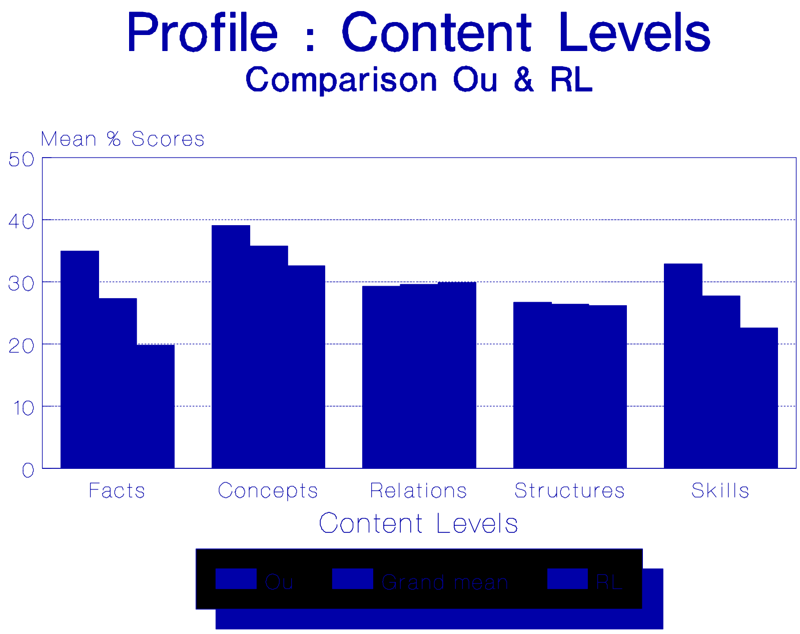

Figure 9: Content knowledge profile

In relation to each profile dimension, Wilk's Lambda (λ) was calculated and p-level determined. In the results table, Wilk's λ is not reported in relation to four dimensions (marked with *). This is because these dimensions only contain two variables; in these cases a test of significance for Hoteling's T2 using the unique sums of squares was calculated, checking the interaction of the independent variable (university) and the two dependent variables on the specific dimensions.
The data in table 10 are helpful to detect specific significant differences between OU- and UL-students. The intermediate conclusion of non-significant differences between OU-students and UL-students, based on analysis of the overall economics-score (F= 3.75, p_F= .054), can be revisited by the refined breakdown of the profile analysis results. There are 7 knowledge profile dimensions which are helpful to illuminate significant differences between both student populations.

Table 10: Profile analysis data for the parallelism test

Profile dimension	Wilk's λ or F	p
Economics subdomains	.59	.00
Curriculum level*	44.24*	.00
Curriculum accent*	0.13*	.72
Node relation	.78	.00
Behavioural	.89	.00
Content	.67	.00
Epistemological	.75	.00

Information level*	0.00*	.98
Number of propositions	.98	.16
Representation level	.83	.00

Profile analysis: discriminant analysis

A further analysis of non-parallel profiles can help to identify those parameters along the specific dimensions that contribute most to the differences between the two subpopulation (UL and OU). In table 11, the results of this discriminant analysis are reported. Wilks λ can in this context be interpreted as the proportion of variability not explained by the group differences. In the fourth column of the table, we derived from this value the proportion of variability explained ((1-λ) * 100) by the group differences resulting from the independent variable 'university'.

Table 11: Results of the discriminant analysis

Profile dimension	Wilk's λ	p_{λ}	$\%_{explained}$
Economics subdomains	.54	.000	46 %
Curriculum level*	.80	.000	20 %
Node relation	.75	.000	25 %
Behavioural	.85	.000	15 %
Content	.63	.000	37 %
Epistemological	.71	.000	29 %
Representation level	.80	.000	20 %

The discriminant analysis can be extended by calculating structure coefficients[8] to determine the discriminatory power of the separate parameters along a knowledge profile dimension. The results of this analysis are found in table 12.

Table 12: Overview of structure coefficients

Dimension	Parameter	Structure coeff.
Economics subdomains	Reporting	.49
	Financing	.30
	Organization	.40
Curriculum level	Level2	.64
Node relation	Initiate	-.76
	Property	-.29
Behavioural	Apply	.74
Content	Factual	.51
	Methods	.45
Epistemological	Implem. An.	-.51
	Kident	-.40

[8] Since the subvalues on each profile dimension are highly intercorrelated, we cannot use raw or standardized discriminant function coefficients. The highly correlated variables "share" the discriminants weights. It is safer to base our interpretation on the structure coefficients which are less likely to be influenced by these intercorrelations.

Representation Level	Textual-symbolic	-.44
	Textual-graphical	.42

A structure coefficient indicates the correlation between a parameter and the discriminant function. High values indicate important discriminant effects (only the relevant ones are given). Table 11 demonstrates that some parameters have structure coefficients up to > .5 . If we combine these results with the dimensions relevant to a statistical significant degree to differentiate between populations, we can conclude that the two university populations can especially be differentiated along the following two knowledge profile dimensions: the content level dimension and the epistemological level dimension.

Profile analysis: flatness test

Is the mastery of the prior knowledge state, as defined by the parameters along a dimension different, independent of the groups (a within-subjects main effect)? In other words, do students master the prior knowledge state in a similar way as defined by the different parameters along a dimension? This question is especially relevant for parallel profiles, since in non-parallel profiles at least one parameter is not flat; nevertheless also the results in relation to non-parallel profiles are reported.
If the flatness test is non-significant, then the profiles are not helpful to clarify or detect differences in the mastery of different components of the prior knowledge state. The results of the flatness test are therefore also of relevance to determine the validity of the knowledge profile dimensions. The results of the flatness test are found in table 13. For each dimension Wilks λ has been calculated, with the exception of the three dimension where only two parameters are available along the dimension; there the F-value is reported (marked with *).

Table 13: Results of the flatness test in profile analysis

DIMENSION	Wilks λ or F	p_λ or p_F
Economics subdomains	.39	.000
Curriculum level*	34.14*	.000
Curriculum accent*	23.36*	.000
Node relation	.69	.000
Behavioural	.83	.000
Content	.63	.000
Epistemological	.58	.000
Number of propositions	.19	.000
Information level*	111.54*	.000
Representation level	.74	.000

All dimensions result in non-flat knowledge profiles. This implies that all dimensions are helpful to identify a specific structure in the mastery of the prior knowledge state. Following this structure, the mastery of certain components[9] of the prior knowledge state is better than for other components.

Conclusion

Although the overall prior knowledge mean % score for the economics domain is not significantly different between the OU- and the UL-student population, a univariate analysis of the knowledge profiles helps to enlighten obvious and significant differences in relation to components of prior knowledge (see Wagemans, Valcke and Dochy, 1991a).
These results are confirmed and reinforced by the results of the profile analysis (multivariate analysis of variance). The results of the profile analysis (parallelism test and discriminant analysis) help to reveal specific and significant differences between the profiles of both student populations. Up to seven of the ten knowledge profile dimensions prove to be of relevance (table 10). Especially the 'content dimension' and the 'epistemological dimension' are helpful to describe these differences.
A further extension of our profile analysis (flatness test) helps to induce further evidence to support the validity of the knowledge profile dimensions, since all profiles are non-flat. This implies that all dimensions are helpful to identify a specific structure in the mastery of the PKS.

9.2 The difference(s) between students with high and low levels of overall PKS

The overall test score of students with low and high levels of the PKS can easily be compared. The student group can be subdivided into two groups, consisting e.g. each of 50% (or 25%) of the students with scores just lower than the mean (Low PKS; LPKS) and a group with 50% (or 25%) of the students with scores just higher than the mean score (High PKS; HPKS). This means that for the 25% groups we used the middle quartiles. It is obvious that - when comparing these subpopulations - significant differences are observed in the overall test scores ($F_{25\%}$= 706.4 $p_{F25\%}$= .000; $F_{50\%}$= 970.7 $p_{F50\%}$= .000). But, a more refined analysis of the test scores, making use of the ten knowledge profile dimensions can help to clarify in a much more specific way the small or large differences between students with low and high levels of the PKS.

Making use of the data gathered among students, studying at the Dutch Open University and at the University of Maastricht, a comparable statistical analysis as

[9] The concept "components" refers to this subpart of expertise that can be isolated in connection to a specific parameter along a knowledge profile dimension.

described in part 9.1 was executed. Some problems were encountered in relation to basic statistical assumptions about the research data when splitting up the population into two groups consisting of 50%-50%. Therefore, only the data gathered among the research population consisting of 25% of the students with mean % scores higher and 25% of the students with mean % scores lower than the mean were used in the profile analysis.

The main hypothesis of this analysis is: students with high and low levels of prior knowledge are different in terms of parameters along the knowledge profile dimensions. Figure 10 represents the knowledge profile of the two subpopulations in relation to the representation level dimension.

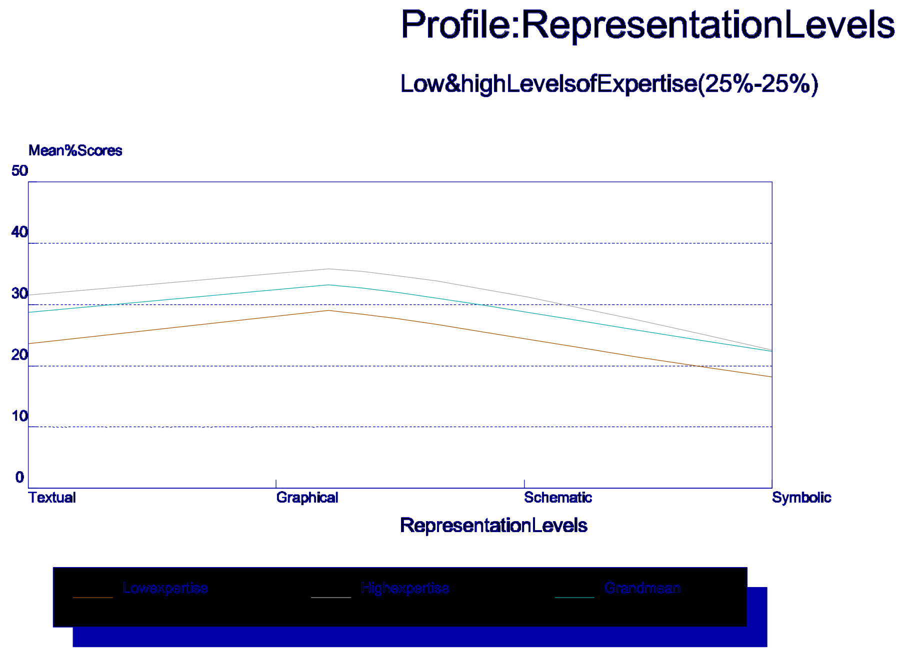

Figure 10: Representation level knowledge profile

Parallelism test

Table 14 presents an overview of the profile analysis results in connection to the parallelism test. This helps us to answer the question whether the two different student groups (LPKS and HPKS) have parallel or non-parallel profiles. In relation to each profile dimension, Wilk's Lambda (λ) was calculated and p-levels determined. In the results table, Wilk's λ is not reported in relation to three dimensions (marked with *). As mentioned earlier, this is because these dimensions only contain two variables; in these cases a test of significance for Hoteling's T2, using the unique sums of squares, was calculated checking the interaction of the independent variable (LPKS and HPKS) and the two dependent variables on the specific dimensions.

The data in table 14 help to detect specific significant differences in knowledge profiles of LPKS and HPKS students. Significant differences are reported in relation to the following dimensions: curriculum level, curriculum accent, epistemological level, number of propositions, information level and representation level. Therefore, the hypothesis stated can be confirmed to a very large extent. Since the results indicate non-parallel profiles, a discriminant analysis and the calculation of structure coefficients are relevant.

Table 14: Results of the parallelism test in profile analysis

DIMENSION	Wilks λ or F	p_λ or p_F
Economics subdomains	.96	.232
Curriculum level*	20.94*	.000
Curriculum accent*	536.17*	.003
Node relation	.97	.041
Behavioural level	.94	.659
Content level	.98	.140
Epistemological	.25	.000
Number of propositions	.80	.000
Information level*	53.55*	.000
Representation level	.96	.006

Discriminant analysis: structure coefficients

Which parameter(s) along the dimension does or do contribute most to the differences in profiles. Table 15 presents the results of the discriminant analysis. In the second and third column a new value for Wilks λ is reported and its significance level[10]. In the fourth column, we derive from λ ((1 - λ) * 100) the proportion of variability that is explained by the group differences based on the independent variable (LPKS and HPKS).

Table 15: Results of the discriminant analysis

DIMENSION	Wilks λ	p_λ	% $_{explained}$
Curriculum level*	.27	.000	73 %
Curriculum accent*	.27	.000	73 %
Epistemological	.25	.000	75 %
Number of propositions	.26	.000	74 %
Information level*	.26	.000	74 %
Representation level	.27	.000	73 %

Since this extension of our profile analysis indicates significant λ, the analysis can be continued by calculating structure coefficients to determine the discriminatory power of the separate values for each dimension parameter. The results of this further analysis are reported in table 16. Only the most relevant structure coefficients in relation to each dimension are given.

Table 16: Overview of most relevant structure coefficients

[10] Of course this further analysis is limited to the dimensions resulting in non-parallel profiles.

Dimension	Parameter	Structure coeff.
Curriculum level	Level1	.37
Curriculum accent	General	.92
Epistemological	Epist. an.	-.52
Number of propositions	< 5 prop.	-.47
Information level	With stem	.70
Representation level	Textual	-.68

Profile analysis: flatness test

Is the mastery of the prior knowledge state as defined by the parameters along a dimension different, independent of the groups (a within-subjects main effect)? In other words, do students master the prior knowledge state in a similar way as defined by the different parameters along a dimension?
The results of the flatness test are found in table 17. For each dimension Wilks λ has been calculated, with the exception of the three dimensions where only two parameters are available along the dimension; there the F-value is reported (marked with *). All dimensions result in non-flat knowledge profiles. This implies that all dimensions are helpful to identify a specific structure in the mastery of the prior knowledge state. Following this structure, the mastery of certain components of the prior knowledge state is better than for other components, as can be seen in the univariate analysis (see Valcke and Dochy, 1991b).

Table 17: Results of the flatness test in profile analysis

DIMENSION	Wilks λ or F	p_λ or p_F
Economics subdomains	.33	.000
Curriculum level*	2415.11*	.000
Curriculum accent*	26.46*	.000
Node relation	.50	.000
Behavioral level	.60	.000
Content level	.42	.000
Epistemological	.49	.000
Number of propositions	.05	.000
Information level*	2760.67*	.000
Representation level	.46	.000

Conclusion

All the knowledge profile dimensions help to differentiate - following the results of the univariate analysis of variance - between LPKS and HPKS students (Valcke and Dochy, 1991b). All differences between both student populations are significant, whether we consider the 25%-25% or the 50%-50% group. The reader might conclude that the dimensions have not been helpful to detect very specific contrasts between LPKS and HPKS students. But, the specification of the 25%-25% group has proven to be useful since more distinctive differences between low and high performers could be detected than comparing the 50%-50% group. Expectations about the knowledge profiles - based on our theoretical considerations - have been largely confirmed (e.g. hierarchical nature, subsequent difficulty levels, etc.)[11]. The results of the profile analysis - only executed for the 25%-25% group since basic assumptions to execute multivariate analysis of variance were violated - help to detect specific significant differences in the following knowledge profiles of LPKS and HPKS students: curriculum level, curriculum accent, epistemological level, number of propositions, information level and representation level. The results of the flatness test indicate that all dimensions result in non-flat knowledge profiles. This implies that all dimensions are helpful to identify a specific structure in the mastery of the PKS.

9.3 The difference(s) between economics and law students, studying the same course

'Diploma type' has been regularly used as an independent variable in our research (Dochy and Bouwens, 1990b; Dochy, Valcke and Wagemans 1991). This research body was helpful to conclude that 'diploma type' is not useful as a relevant indicator of the PKS since conflicting results are obtained from the consecutive studies. But the question endures in the context of this text about knowledge profiles. Are economics (ES) and law students (LS) for instance - notwithstanding non-significant differences in their overall PKS - different when looking at their knowledge profiles?

The main hypothesis of this analysis therefore is: ES an LS are not different in terms of parameters along the knowledge profile dimensions. In the present analysis, the test results of law students and economics students, studying at the Dutch Open University were used for further analysis. Figure 11 represents the knowledge profiles of the economics and law students for the content dimension.

[11] For more details: see Valcke and Dochy (1991b).

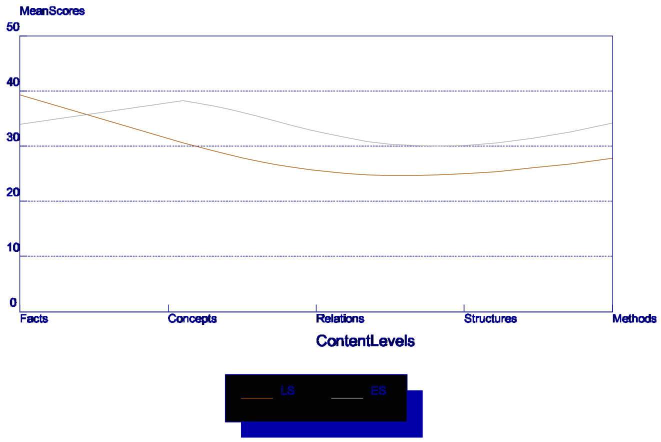

Figure 11: Content knowledge profile

Profile analysis: parallelism test

Table 18 presents an overview of the analysis results in relation to the parallelism test. In relation to each profile dimension, Wilk's Lambda (λ) was calculated and p-levels determined. Hoteling's T^2 is reported in relation to three dimensions (marked with *) for reasons mentioned earlier. For these cases the F-value and p_F-value are reported.

The data in table 18 are not helpful to detect specific significant differences between ES and LS. The profiles of ES and LS are parallel. The conclusion of non-significant differences between ES and LS, based on analysis of the overall economics-score (F= 2.12, p_F= .15), cannot be revisited by a more refined analysis, based on the 10 dimensions. The only, slightly significant F-value is obtained in relation to the content dimension (p=.025). Since the results indicate no significant differences at this level, a more elaborated profile analysis or the calculation of structure coefficients etc. is not relevant. As a consequence, the hypothesis stated cannot be confirmed.

Table 18: Results of the parallelism test in profile analysis

DIMENSION	Wilks λ or F	p_λ or p_F
Economics subdomains	.81	.061
Curriculum level*	3.03*	.086
Curriculum accent*	1.78*	.187
Node relation	.93	.288
Behavioural	.99	.659
Content	.86	.025
Epistemological	.98	.811
Number of propositions	.98	.453
Information level*	1.67*	.201
Representation level	.94	.197

Profile analysis: flatness test

The results of the flatness test are found in table 19. For each dimension Wilks λ has been calculated, with the exception of the three dimension where only two parameters are available along the dimension; there the F-value is reported (marked with *).

Table 19: Results of the flatness test in profile analysis

DIMENSION	*Wilks l or F*	*p_l or p_F*
Economics subdomains	.38	.000
Curriculum level*	241.16*	.000
Curriculum accent*	3.74*	.057
Node relation	.76	.000
Behavioural	.95	.151
Content	.61	.000
Epistemological	.59	.000
Number of propositions	.28	.000
Information level*	178.23*	.000
Representation level	.87	.017

With the exception of the 'curriculum accent' and 'behavioural level' dimension, all dimensions/parameters help to differentiate in the mastery of components of the prior knowledge state. The non-significant results for the 'curriculum accent' dimension can be explained by the restricted number of items, measuring the mastery of general economics and B.A. The differences in the mastery of the prior knowledge state along the 'behavioural level' dimension is indeed very small and as expected non-significant.

Conclusion

The results of the univariate analysis of variance when using the knowledge profiles dimensions are not helpful to identify specific and significant contrasts between both student-groups (see Dochy and Valcke, 1991c). This affirms our earlier research findings showing that 'student type' is not a relevant 'indicator' of the PKS (Dochy, Valcke and Wagemans, 1991; Valcke and Dochy, 1991a). The non-significant difference between both sub-populations in relation to the overall economics-score is confirmed. But on the other hand, we perceive a consistent trend that ES perform better than LS. The results of the multivariate analysis of variance (table 18) are also not helpful to detect specific significant differences between ES and LS. The profiles of ES and LS are parallel. The intermediate conclusion of non-significant differences between ES and LS, based on the univariate analysis cannot be revisited by a more refined analysis, based on the 10 dimensions. The only, slightly significant F-value is obtained in relation to the content dimension (p=.025). The results of the flatness test help to indicate that - with the exception of the 'curriculum accent' and 'behavioural level' dimension- all dimensions/parameters help to differentiate in the mastery of components of the

PKS.

10 Conclusions

In this chapter we have introduced knowledge profiles as being graphs of scores of a group or individuals on a prior knowledge state test. We defined several dimensions on which knowledge profiles can be based and tried to look beyond the subject-matter level. To validate the different knowledge profile dimensions, the parameters have been related to the DS KST items and we analyzed the extent to which the parameters along the dimensions give information about the components of the PKS. The non-hierarchical dimensions suggest that they do relate to different components of the prior knowledge state. This is further supported by the profile analyses (flatness tests) when comparing different populations. The other five dimensions reveal to be indeed hierarchical, although to a different extent.

Further analyses have been looking at the discriminatory power of the dimensions to detect the PKS differences between several student populations.
Although we did not expect differences between ES and LS as a results of the ANOVA, this could not be revisited by a more refined profile analysis. ES and LS do not significantly differ, even not in their knowledge profiles. UL en OU students do not differ in their mean % score, but profile analyses could reveal significant differences on the knowledge profiles. Also for students of the LPKS and HPKS groups (middle quartiles), differences are found in their knowledge profiles. Promising knowledge profile dimensions for differentiating are certainly curriculum level dimension, representation level dimension, epistemological dimension, content dimension, economics subdomains dimension, and further the behavioural dimension, curriculum accent dimension, information level dimension and the number of propositions dimension. Nevertheless, their value is different relying on the group or in this case the goal of the comparison between groups.

This chapter is of importance since we succeeded in defining and operationalizing a new and promising approach towards the analysis of prior knowledge. It is foreseen that in situations where there are significant differences between the PKS of specific subpopulations, the profile dimensions are helpful to detect and dissect the strengths and weaknesses of the students involved. This might be a promising starting point for differentiated diagnostic and guidance approaches.