

Using propositions for the assessment of structural knowledge

Nick J. Broers
Maastricht University

Journal of Statistics Education Volume 17, Number 2 (2009), www.amstat.org/publications/jse/v17n2/broers.html

Copyright © 2009 by Nick J. Broers all rights reserved. This text may be freely shared among individuals, but it may not be republished in any medium without express written consent from the author and advance notification of the editor.

Key Words: Conceptual understanding; Concept mapping; MPM.

Abstract

It is well known that meaningful knowledge of statistics involves more than simple factual or procedural knowledge of statistics. For an intelligent use of statistics, conceptual understanding of the underlying theory is essential. As conceptual understanding is usually defined as the ability to perceive links and connections between important concepts that may be hierarchically organized, researchers often speak of this type of knowledge as structural knowledge. In order to gain insight into the actual structure of a student's knowledge network, specific methods of assessment are sometimes used. In this article we discuss a newly developed, specific method for assessing structural knowledge and compare its merits with more traditional methods like concept mapping and the use of simple open questions.

1. Introduction

Teaching introductory statistics to students with a non-mathematical background is a difficult task for multiple reasons. Quite often, the statistics course is a mandatory part of the curriculum and students are not very motivated to engage in the task of learning this subject. In addition, the conspicuous presence of formulas in the statistics textbooks are often experienced as intimidating and can easily lead to avoidance behavior on the part of the student. Only with the exam looming ahead are some students prepared to open the books and try to make sense of it all. As statistics is a cumulative subject, growing in complexity from the first week onward, such a strategy is likely to end up in failure.

There is, however, another reason why teaching statistics is difficult. This has to do with a divergence of the learning goals set by teachers and those set by the students. To many students, statistics is all about doing calculations and making proper use of formulas. It may be we are trying to teach our students to consider some research question, to ponder on the correctness of the research design that was chosen to answer the research question, to think about appropriate statistical models, to check on the validity of their underlying assumptions, and to attempt to derive as much information from the data as possible.

Yet, many students fixate themselves on the production of numbers and on the conclusions that may be derived from comparing p-values to significance levels.

As many teachers experience to their dismay, students are often quite capable of doing the sums correctly and of making sound decisions on significance, but in spite of the emphases that may have been placed during the course, many students have difficulty in understanding what the numbers mean, why these particular numbers are appropriate in a given context, and why seemingly suitable alternatives should be ignored on this particular occasion. What the students lack, is a conceptual understanding of the statistical theory that underlies the numbers they are using for their decisions.

Conceptual understanding can be defined in different ways (e.g. [Hiebert and Lefevre, 1986](#); [Lampert, 1986](#); [Huberty, Dresden & Bak, 1993](#); [Broers, 2002a](#); [Kaplan 2006](#); [Budé, 2007](#); [Schau and Mattern, 1997a](#)) but in all cases this type of knowledge is contrasted with mere propositional knowledge, which is taken to be isolated knowledge of definitions, principles and basic ideas ([Huberty et al., 1993](#); [Broers, 2002a](#)). In agreement with the current interpretations of conceptual understanding, we can say that as soon as knowledge become interrelated or structured, propositional knowledge becomes conceptual understanding. Because such interrelationships involve a mental organization of links between concepts, [Schau and Mattern \(1997a\)](#) preferred to speak of connected understanding. Apart from showing links between concepts, a knowledge network that is indicative of conceptual understanding usually also displays an hierarchical structure, with higher level concepts subsuming more elementary ones. Such mental organizations of knowledge are therefore alternatively described as structural knowledge ([Jonassen, Beissner and Yacci, 1993](#)).

In statistics, a lot of practical knowledge is of a procedural nature. The application of a statistical model for the purpose of data analysis, for instance, involves the choice of a correct model on the basis of the chosen research design, the checking of assumptions and the drawing of conclusions on the basis of relevant statistics. Many of the steps in this process can be carried out as an array of automated steps and decisions. It is only by ensuring that our students develop a conceptual understanding of the rationale underlying these steps and decisions, that we can prevent them from acquiring cookbook knowledge, rather than statistical knowledge.

A well known example pertains to the elementary use of inferential statistics. Early on in an elementary statistics course, students learn about the concept of the significance level and how it should be used to decide on the fate of the null hypothesis. Of course, as the course progresses, students gain additional knowledge on this matter that is meant to put such practice into perspective. For instance, they will learn about confidence intervals, effect sizes and power. Students that succeed in integrating this additional knowledge with the more basic knowledge on significance levels and hypothesis testing, will refrain from drawing simple conclusions on the basis of p -values like " $p = .04$, so the theory works" or " $p = .08$. What a pity, it was all a waste of time". Being able to shift one's attention from such p -values to, for example, sample size considerations is a healthy indication of the existence of structural knowledge.

As [Kelly, Sloane and Whittaker \(1997\)](#) have stressed, in order to be able to draw conclusions on the conceptual understanding of students, we must find ways of directly assessing the existence of such structural knowledge. Assessment tasks that require students to do work on statistical problems that can be solved by automated strategies will not permit us to draw conclusions on their true understanding of the subject. And without conceptual understanding of statistics, demonstrated knowledge is not likely to last for long or to be transferred to different problem tasks that superficially seem different from the known ones, but which are structurally identical (so-called far transfer problems. See [Paas, 1992](#)).

Over the years, many different assessment methods have been developed for the express purpose of uncovering the structure of a knowledge network (see for an extensive overview [Jonassen et al., 1993](#)). Sometimes such knowledge has been elicited indirectly, like having learners generate free word association lists on a given knowledge domain (e.g. for determining the structure of a students' mathematics knowledge, see [Geesling and Shavelson, 1975](#)), or by having students rate the similarity between a number of pairs of concepts, all related to a given knowledge domain (see [Brown and Stanners, 1983](#), for an example related to psychology). Apart from such indirect methods for the assessment of structural knowledge, various methods have been developed which explicitly tap and describe the structure of a students' knowledge network. By far the most familiar of these is the use of concept mapping.

1.1 Concept mapping for assessment purposes

[Figure 1](#) shows an example of (part of) a concept map that was used for instructional purposes and later, in a slightly

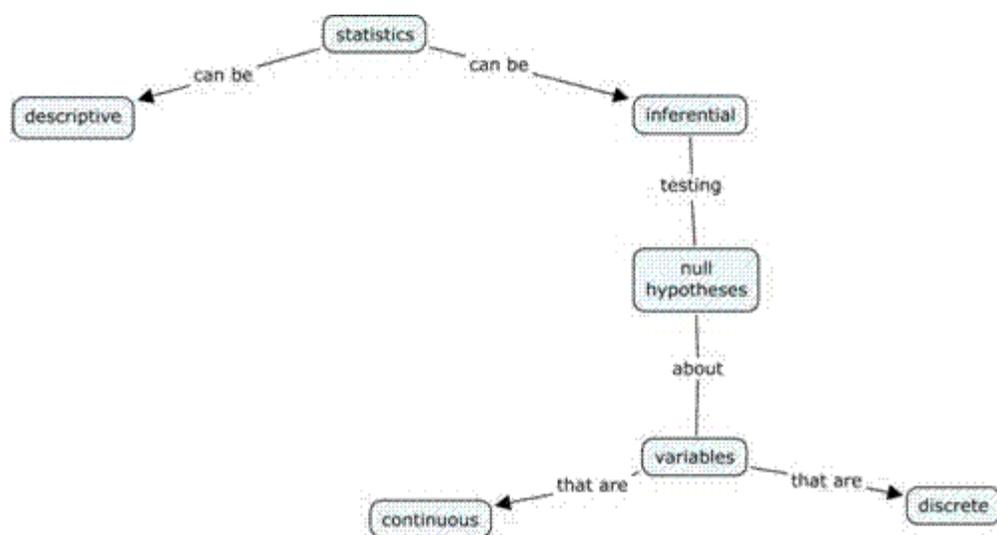
modified form, for assessment purposes (taken from [Schau and Mattern, 1997a](#)). The distinctive features of concept maps are: concepts that are placed in ovals, connecting lines or arrows showing relationships between these concepts, and small commentaries alongside the arrows that describe the nature of the relationship between the two connected concepts. In the context of concept mapping, a concept – link – concept triplet is called a proposition and these propositions are considered to be the basic units of meaning.

The use of concept mapping has gained widespread popularity as an aid in instruction, for instance for the purpose of structuring the discussion around some topic or knowledge domain. Apart from its role as a tool for instruction, concept mapping is also used for assessment purposes. There are multiple ways in which concept maps can be used for assessment (see for an extensive overview [Ruiz-Primo and Shavelson, 1996](#)), but the most often used alternatives are free mapping with concepts provided and a Select-And-Fill-In the map method (SAFI for short).

In free mapping with concepts provided, a student is given a list of concepts that are pertinent to a given knowledge domain, and is asked to draw a concept map showing how these concepts interrelate. In the SAFI variant, a concept map like in Figure 1 is provided, but with a number of ovals being empty. The student is given a list of concepts that should then be placed in the empty ovals. This latter method was actually used by [Schau and Mattern \(1997b\)](#) for the assessment of conceptual understanding of statistics.

So far, the use of concept mapping for the purpose of assessment has not matched its popularity as an aid to instruction. Partly, this is due to the fact that concept mapping is a

Figure 1 – Example of a concept map



translation of structural information in a psychological knowledge space into the structural information of a graph. As such, it involves a transformation of knowledge that was gained and structured in a verbal (or sometimes mathematical) sort of way, into a graph displaying relationships between abstract symbols. For such a transformation to yield a reliable and informative display of structural knowledge, a lot of training is necessary, and some students never get the hang of it. ([Schau and Mattern, 1997b](#)). In addition, for a display of structural knowledge pertaining to a given domain, the number of possible relationships quickly becomes sizeable. When, say, 20 concepts are provided, the number of possible links equals $\binom{20}{2}$ or 190 relationships. Not all of these possible links will provide meaningful

relationships, but for a student to adequately decide on which relationships are, and which relationships are not relevant, will take up a lot of time. It is likely that a student will omit a number of relationships, not because he or she does not know them, but because the large number of possible links makes it very likely that a subset of them will be overlooked. When concept maps are used as an aid to instruction, for instance as a tool for structuring discussions in a group, this would not be a problem. However, when concept maps are meant to reveal to us the structure of a students'

knowledge network, this fact obviously threatens to undermine the validity of the assessment.

For these reasons [Schau and Mattern \(1997b\)](#) advocate use of the SAFI alternative instead. It is much easier and takes less time for students to become acquainted with. In addition, unlike free mapping this assessment task is easy to score objectively. However, a major drawback against this method is that the knowledge structure that we wish to determine in the mental organization of the subject, is already implanted by the instructor. The student merely fills in concepts in the blank spaces, the actual relationships are already preset. Research has shown that for this reason, the validity of SAFI as a measure of structural knowledge is disputable ([Ruiz-Primo, Shavelson, Li and Schultz, 2001](#)).

1.2 The method of propositional manipulation (MPM)

An instruction method that has been proposed as an alternative to concept mapping, and that focuses on propositions rather than on concepts, is called the method of propositional manipulation, or MPM for short. ([Broers, 2002b](#); [Broers, 2007](#)).

In this method, the study material is decomposed into a finite collection of constituent propositions. Subsequently, students are presented with tasks in which they build arguments on the basis of prescribed subsets of these propositions. The building of such arguments, it is believed, forces the student to engage into guided self-explanation. Guided, because the argument has to be built on the basis of the provided subset of (mostly about six) propositions. For an example, consider the MPM assignment in [Box 1](#). Note that in this example, the propositions are replaced by study questions. The answers to these questions form the propositions they have to use in the argument.

A complete description of this method as an aid to instruction can be found in [Broers \(2002b\)](#). At our university, such MPM tasks are provided on a regular basis to keep the students constantly aware of the sort of knowledge they are expected to gain in following the course. It prevents them from overly concentrating on the computational aspects of statistics, or from memorizing procedural steps in a rote fashion, without reflection on their underlying rationale.

The potential of MPM as an instructional method has been explored elsewhere. Research has shown this method to be beneficial for furthering both propositional knowledge ([Broers and Imbos, 2005](#)) and conceptual understanding ([Broers, Mur and Bude, 2005](#)). In the present research project we wanted to examine more closely the potential of MPM as an instrument of assessment. To bring this potential more closely into focus, we explored the qualities and limitations of this assessment method, relative to the use of concept mapping as an instrument for assessment.

In comparison to concept mapping, MPM may be expected to have a number of advantages. First, a subject like elementary statistics is mainly built on the basis of verbal propositions, complemented by a number of mathematical expressions. Although we do

Box 1 — Example of an MPM assignment for instructional purposes

Statement: If we make a Type II error, this implies that the probability distribution we used to determine the p-value of our test statistic was not an adequate model of the empirical reality.

Instruction: construct an argument that shows the above statement to be either true or false, and use the answers to each of the following questions in your argument:

Study questions to be used:

- What is a test statistic?
- What conditional probability distribution are we working with, when testing a null hypothesis?
- The value of the test statistic is reported with a corresponding p-value. What is meant by this p-value?
- Why is this p-value a conditional probability?
- What is a significance level?
- What is a Type II error?

not know how the brain physically stores the knowledge it receives, we do know that this knowledge is primarily transferred in the form of verbal propositions, and the most convincing proof of comprehension would be for the student to explain her knowledge verbally to someone else. When doing an MPM task, a student has to puzzle on the links that will meaningfully connect the propositions to the statement that has to be shown true or false. This cementing together of the propositions into an argument requires some training, but the eventual argument, when correctly delivered, takes up the form of a verbal explanation. Its form of expression therefore seems close to the way that the student has actually organized his or her knowledge in the mind. Unlike concept mapping, where the subject has to translate his knowledge organization into a graph, MPM allows the student to give a more direct, less transformed display of his or her structural knowledge. In comparison to concept mapping therefore, less training seems necessary to enable students to do this kind of assessment task in a way that will be informative to the instructor.

Second, an assessment of structural knowledge should tell us whether relevant concepts were properly understood (propositional knowledge) and whether important interrelationships between these concepts are properly comprehended (thus indicating connected or conceptual understanding). The graphical display in the form of concept maps is very concise and therefore often lacking in information on propositional knowledge. The arrows that connect concepts tend to provide space for only very brief comments, like "is", "causes", "assumes", et cetera. This typically results in propositions like "Sampling distribution ... is ... probability distribution". This could summarize a wealth of correctly organized knowledge, but this simple statement of a relationship could equally well hide huge misconceptions. The sparsity of information provided by the graph prevents us from determining what is actually the case.

By contrast, we expect MPM to be far less equivocal than concept mapping. MPM forces the student to explicate her propositional knowledge, and to relate the various propositions in a convincing way so as to deliver a verdict on the true-false statement. Because of its richness in verbal content, such an argument, we assume, will give a far less ambiguous overview of structural knowledge. In addition, such arguments can easily betray the existence of certain misconceptions existing in the knowledge network of the subject.

To see whether the assumed advantages of MPM tasks over concept mapping actually hold, we conducted an empirical study. Since the emphasis of the study would be on an examination of the written content of the MPM tasks and on a qualitative analysis of the concept maps, we opted for a small group of students. Although the small sample size will not permit sweeping generalizations, we do expect it to yield meaningful insights into the potential and into the possible shortcomings of MPM, relative to the use of the concept map for assessment purposes.

2. Method of the study

2.1 Participants

Fourteen second year psychology students, who all had a fine record on the two first year statistics exams (covering descriptive statistics and inferential statistics, respectively), volunteered to participate in this study. Their willingness to participate stemmed from their interest in the subject of statistics; no financial reward was offered. Each subject was asked to do several tasks which were presented on three different occasions, with one week intervals in between. Amongst the tasks were the construction of an MPM task and a corresponding concept mapping task. Four of the 14 subjects only did either the MPM or the concept mapping task, and the results of these students have for that reason been omitted from the present analysis.

2.2 Material

1. Open question (OQ)

All students received the following open question:

"Is the following statement true or false?:

To determine whether an estimator is unbiased or not, you have to study its sampling distribution.

Explain your answer."

This open question corresponds to the true-false statement in the MPM task, but in the OQ case the students are free to respond in any way they like. This open question was given to permit a sort of baseline for the amount of information that is yielded by the true-false statement per se, without the MPM assignment structure.

2. Partial concept map (PCM)

We asked all subjects to create a concept map on the basis of the following seven concepts: estimator, random variable, parameter, probability distribution, sampling distribution, unbiased, expected value. Usually, when requested to draw a concept map, the number of concepts provided tends to be larger, and we therefore speak of a partial concept map, because it is constructed on the basis of a sample of the full range of concepts that would be provided for the purpose of drawing the complete concept map. The seven concepts chosen for the PCM correspond to the seven propositions that figure in the MPM task.

The participants could be expected to be acquainted with these seven concepts and their interrelationships, as these had all figured prominently in the first year lecture on elementary testing and estimation theory. At the same time, since the students had not touched on these statistical topics for over six months, part of it might have been forgotten. So, some variation in structural knowledge with regard to this small sample of concepts might be expected.

By focusing on this limited collection of only seven concepts, we could determine beforehand which relationships could be revealed by the subject. In fact, there are $\binom{7}{2}$ possible combinations (or in the present context: relationships),

but not all of these 21 relationships are meaningful. There is no direct relationship between 'sampling distribution' and 'unbiased', for instance, nor is there a direct relationship between 'random variable' and 'parameter'. Apart from non-existing relationships, a number of relationships are more general duplicates of specific relationships. For instance, take 'A random variable has a probability distribution' and 'An estimator is a random variable'. From this it follows automatically that 'An estimator has a probability distribution'. Every relationship that goes for a random variable, also goes for an estimator, being a specific instance of a random variable. Since the probability distribution of an estimator is called the sampling distribution, we would want our students to demonstrate knowledge of the relationship between 'random variable' and 'probability distribution', but not necessarily between 'estimator' and 'probability distribution'. On the other hand, a demonstrated link between 'probability distribution' and 'sampling distribution' would be essential. In addition, a few relationships were considered less important because they received little emphasis in the first year lecture on this topic (e.g. 'A random variable has an expected value'. That lecture focussed on the estimator and its sampling distribution, and so the idea of an expected value was connected to estimators in specific, rather than to random variables in general).

On the basis of such considerations, we ended up with 7 relationships that we believed could be expected to be observed in the structural knowledge network of the informed student. These are listed in [Box 2](#).

Box 2 — The seven relationships that we wished to see demonstrated in the PCM (as well as in the MPM)

R1: An estimator is used to estimate the value of a parameter

R2: An estimator is a random variable

R3: A random variable has a probability distribution

R4: An estimator has a sampling distribution

R5: A sampling distribution is a probability distribution

R6: The expected value is the mean of the sampling distribution

R7: An estimator is unbiased if its expected value equals the value of the parameter

We would look for the presence of these seven relationships in the concept map, in the MPM task, and in the Open Question.

3. MPM

Apart from the concept map, students also had to do an MPM task involving the same concepts and therefore possible relationships. The MPM task opened with the following true-false statement: "To determine whether an estimator is unbiased or not, you have to study its sampling distribution". Next, the assignment was: "Construct an argument showing the above statement to be true or false, and in your argument make use of the answers to the following questions:

- what is an estimator?
- what is a random variable?
- what is a parameter?
- what is a sampling distribution?
- what is a probability distribution?
- what is an unbiased estimator?
- what is meant by the expected value?

2.3 Procedure

The 10 subjects who had participated all the way, were presented with the three tasks on three consecutive weeks. In the first week, all subjects did the open question. In addition, the subjects were provided with examples of concept maps and written instruction of how to construct one, and with an example of an MPM task and a written instruction of how it should be worked on. In the second week, about half of the students did the PCM, the remaining group did the MPM. In the last week, the students who had done the PCM in week 2 now did the MPM, and those who did the MPM in week 2 now did the PCM. In their work on the tasks, the students were permitted to consult the textbook of the first year course (Moore and McCabe's "Introduction to the practice of statistics") to refresh their memory on the subject.

It is important to stress that our objective was not the determination of the actual level of structural knowledge of the individual participants, but rather on the quality of the information that concept maps and MPM type tasks would yield in this respect.

3. Results

Because of the small number of subjects, the application of significance tests is not very meaningful. In the discussion of the data, the emphasis is on qualitative information.

To get a clear idea of what the material looks like, we shall now first look at the output presented by subject #8. In [Table 1](#) we first look at the profile of this subject, in terms of the relationships that he or she has demonstrated by doing the OQ, the PCM, and the MPM respectively. Information on the coding of the PCM and the MPM can be found in [Appendix 1](#) and [2](#), respectively. As can be seen, OQ scarcely gives any information at all, and PCM and MPM are in complete agreement as to the presence or absence of relationships.

Table 1 – Relationships demonstrated by subject #8, according to PCM, MPM and OQ (a '+' means 'relationship indicated as present', '-' means 'relationship indicated as not present').

	Relationships						
	R1	R2	R3	R4	R5	R6	R7
PCM	-	+	-	+	-	-	+
MPM	-	+	-	+	-	-	+
OQ	-	-	-	-	-	-	+

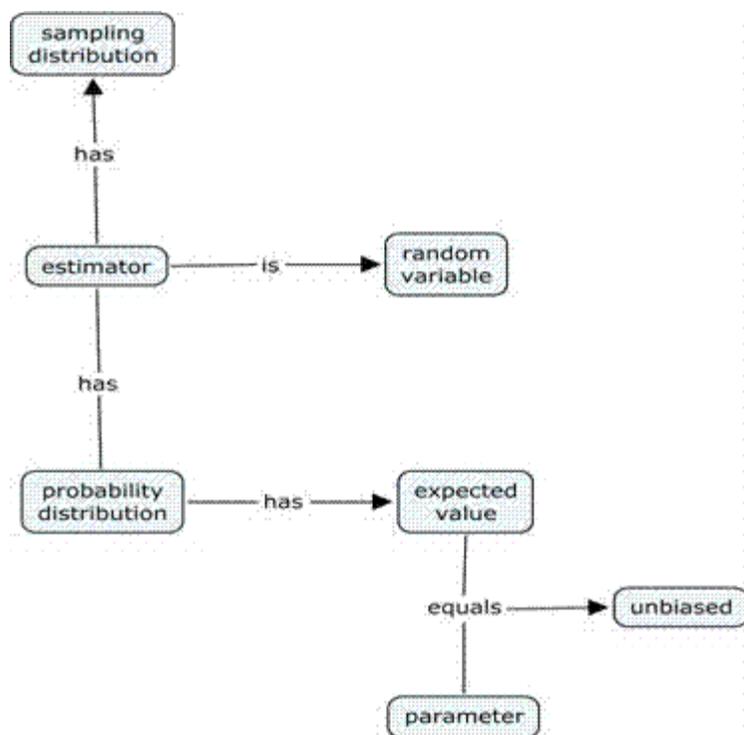
In [Box 3](#), we see how this subject has responded to the Open Question. As we can see, the response to the open question, although basically correct, yields very little information. Only a single relationship is stated (R7), and no additional propositional knowledge is conveyed (the opening sentence is merely a tautology). The sparsity of information in the OQ is perhaps not surprising, because arguably none of the other relationships is necessary for providing an answer to the question.

Box 3 — Response to Open Question by subject #8

Correct, an unbiased estimator should not contain any sort of bias. You can check this by determining whether the expected value equals the parameter (the true population value).

Next, we consider [Figure 2](#), which displays the PCM for this subject. Technically speaking, the concept map is alright. All lines are directed ¹ and commented. (On average, the 10 subjects produced 6.9 links. 4 out of 10 subjects produced at least some – on average 44% - links that were undirected. In addition, 4 out of 10 subjects produced at least some – on average 45% - links that had no comment). The first thing that is apparent is the conciseness of the information provided by the concept map. Just a few lines suffice to express a lot of structural information. However, the information, although structural, is far from clear or unequivocal. First of all, the graph does not really convey a lot of information on the knowledge structure of the subject. For instance, we cannot tell to what extent the subject really understands the concepts that he has used. The concept of sampling distribution is known to be very complex and difficult to many students. All we can tell from the map is that the subject has not displayed any relationships involving this concept, except for the link with the concept of estimator. Partly, this may reflect shallow knowledge on the part of the subject. But even if the subject has a fairly rich conception of the sampling distribution, the method of concept mapping, especially with so few concepts provided, does not offer much opportunity to demonstrate the richness of his knowledge. Against this position one might perhaps argue that more concepts should have been provided, thereby enabling a richer graphical display of the meaning of a concept like sampling distribution. However, as more concepts are provided, the number of possible links quickly increases and with that the probability that structural knowledge fails to be displayed.

As to the issue of equivocality, if we look at the lower right part of the figure for example, we may be tempted to assume that the subject testifies to correct knowledge concerning the definition of an unbiased estimator. The subject probably means to say that as a random variable, an estimator has an expected value which shows up as the mean of the sampling distribution. If this expected value equals the true value of the parameter, the estimator is said to be unbiased. But by concluding that this is in fact what the subject means to say, we are projecting our own structural knowledge onto the fuzzy

Figure 2 – PCM by subject #8

structure in the graph of the subject. It is not there in so many words and we cannot be sure that the subject has correctly organized his knowledge. If so, then why is there no explicit connection between probability distribution and sampling distribution? Why no direct connection between expected value and estimator or random variable? What exactly is "equals --> unbiased" supposed to mean? The information in the PCM is not rich enough to allow us to clearly determine what the subject means to convey. Because we cannot tell for sure, we are tempted to give him the benefit of the doubt and to assume that he means the right thing. But that remains an interpretation.

Although less so, the PCM suffers from the same structural weakness as the OQ. Because the subject is not directed in any way to express his structural knowledge (other than that he is to restrict himself to the seven concepts provided), the presented relationships may be a mere random subset of all relationships that the subject actually knows. It may be that if we had explicitly encouraged the subject to focus on specific relationships, these would have shown up in a concept map. None of the links that are portrayed in Figure 2 are incorrect, yet most of the relationships we would like to see demonstrated are not in the PCM. Is the subject unaware of these relationships, or does he know them but for some reason has not drawn them in the map?

Finally we look at the corresponding MPM task made by this subject (Box 4).

Box 4 — MPM by subject #8

To be able to say something about a population you should draw a sample and repeat this process in infinite succession. You then get the sampling distribution. The mean of the sample is an estimator. The value of the estimator is determined by chance: if you draw a new sample the estimator will take up another value (which means that an estimator is a random variable). When we consider the probability distribution of the sample mean (we then get all possible \bar{x} values, together with their associated probabilities), we can determine the expected value of the sample mean, by multiplying each \bar{x} value with its associated probability, and by adding up all the products. In the sampling

distribution of \bar{X} , the expected value is equal to μ . μ is the population mean and therefore a parameter (a value that describes the population). For the estimator to be unbiased, its expected value should equal the parameter. Since you can only say something about the population via use of the sampling distribution, you should, in order to determine whether an estimator is unbiased, do so via the sampling distribution. The statement is therefore correct.

Recall that the MPM task explicitly instructs the students to make use of the answers to seven questions ('What is an estimator?', 'What is a parameter?', etc.). It is therefore not surprising that the MPM is rich in providing detail on propositional knowledge as well as on knowledge of relationships. In Box 4, we find correct propositions on the concepts of parameter, random variable, probability distribution and an unbiased estimator. However, the proposition on estimation is left implicit, the information on the sampling distribution is a bit vague and no definition of an expected value is given (just a procedure for its computation).

As for relational knowledge, the relationship between estimator and random variable is explicitly and unequivocally described, as is the relationship between an unbiased estimator, its expected value and the value of the parameter in the population. R4 ('An estimator has a sampling distribution') is actually stated as 'An estimator has a probability distribution'. The subject then goes on explaining that on the basis of this probability distribution we can compute the expected value. Next, he states that in the sampling distribution of \bar{X} , the expected value is equal to the population mean. In this statement, he explicitly relates the concept of sampling distribution to that of estimator (R4). But does he fully understand the nature of this relationship? I.e., does he comprehend R5? Can we infer from the sequence of statements that he is aware that the probability distribution of \bar{X} and the sampling distribution of \bar{X} are one and the same? It seems plausible, but nowhere does the subject explicitly state this equality. Only if he had done so, could we unequivocally decide that knowledge of R5 had been demonstrated.

Other remarks in the MPM suggest that the subject is not totally confident in his knowledge of the sampling distribution. The opening statement will be puzzling to the general reader. In fact, it reflects the content of a lecture on sampling distributions that the subject had attended in his first year of study. In the lecture, it was said that if we wish to use the sample mean as an estimator for the population mean, we need to know how this estimator 'behaves' if we were to draw an infinite succession of equally sized samples from the same population. Only by using knowledge of the expected value and the standard error of the estimator can we turn the statistic into an informative measure. It is this point that was made in the lecture, that we now find translated in the first two sentences of the MPM. And again, in the last part of the argument, where the subject states 'Since you can only say something about the population via use of the sampling distribution...etc.'. Of course, put in this way, the statement becomes obscure and it suggests that the subject does not fully comprehend the role of the sampling distribution in inference.

3.1 General findings

We have just seen that for subject #8 PCM and MPM completely agreed on the presence or absence of relationships. However, in most cases the two assessment methods tended to diverge somewhat. [Table 2](#) gives an overview of the number of relationships that were uncovered by PCM, MPM, and the Open Question (OQ). Since for each assessment method the relationships were determined for the same 10 subjects, the diverging numbers clearly indicate that none of the methods succeeds perfectly in assessing structural knowledge that is actually there. Specifically, evidence for R3 has been found for nearly all subjects by PCM, but for only 2 subjects by MPM. On the other hand, R6 has never been detected by PCM, but relatively often by MPM.

[Table 3](#) shows to what extent PCM and MPM were consistent in their diagnosis of the presence or absence of relationships in the knowledge organizations of the subjects. For each separate subject, [Table 3](#) gives the number of relationships that were evaluated by both methods in a consonant manner (# cons.) and the number of relationships that were evaluated in a dissonant manner (# diss.).

Table 2 - Detection of relationships by MPM and PCM (cells give the number of subjects for whom the particular relationship was detected).

	R1	R2	R3	R4	R5	R6	R7
PCM	5	10	8	10	8	0	6
MPM	8	10	2	9	7	6	6
OQ	2	0	0	2	0	1	4

As we can see, in each case the number of consonant judgments is greater than the number of dissonant judgments, but the agreement will be boosted by the fact that in those cases where subjects are truly unaware of the existence of a particular relationship, the relationship will simply fail to turn up in any valid type of assessment. Furthermore, we see that in 3 out of 10 cases, the number of dissonant judgments nearly equals the number of consonant ones.

Table 3 – Number of relationships that were evaluated consonantly or dissonantly by PCM and MPM.

Subject	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
# cons.	5	5	4	6	7	4	5	7	4	5
# diss.	2	2	3	1	0	3	2	0	3	2

For most subjects PCM and MPM sometimes gave different verdicts on the presence of relationships. [Table 4](#) shows the number of times that MPM picked up a particular relationship that PCM did not, and vice versa. It is conspicuous that for six subjects, PCM registered the presence of R3 ('A random variable has a probability distribution') whereas MPM did not. On the other hand, MPM picked up the presence of R6 ('the expected value is the mean of the sampling distribution') in six cases where PCM did not pick up this same relationship. Looking at the MPM arguments of the six subjects for whom R3 was detected by PCM but not by MPM, we see that the absence of R3 derives from the fact that after these subjects had made a connection between 'random variable' and 'estimator', they next proceeded to relate the concept of estimator to other concepts like probability distribution and sampling distribution, whilst no longer minding about the concept of random variable. This is because they work towards the true-false statement that focuses on properties of the estimator. The MPM assignment does not require them to integrate into the argument every possible connection between the seven concepts, but only those relevant to the argument. As to the lacking of R6 in PCM, we saw in table 1 that R6 was never once detected by PCM, but that R7 was detected by PCM in six cases.

Table 4 – Number of times that a relationship was detected by MPM but not by PCM, and vice versa (maximum number equals the sample size of 10)

	Relationships detected						
	R1	R2	R3	R4	R5	R6	R7
MPM yes, PCM no	3	0	0	0	0	6	1
PCM yes, MPM no	0	0	6	1	1	0	1

Four of these six subjects had in fact demonstrated knowledge of R6 in their MPM. It seems therefore that in drawing a concept map, subjects may sometimes focus on key relationships and by doing so simply forget to include others.

3.2 Identification of misconceptions

When assessing structural knowledge, it is not only important to determine which relationships are correctly perceived by a subject, but also to pinpoint certain misconceptions that the subject may hold. That is, we also wish to determine the existence of possible erroneous links between concepts. In a number of times, the relationships that are specified by the subject will be obscure. That is, in these instances it will not be possible to determine objectively whether the subject is aware of the correct relationship or not. In the discussion of the MPM material of subject #8, we saw an example of this with regard to the relationship between sampling distribution and probability distribution. At other times, we will be able to determine unequivocally that a specified relationship is erroneous. [Table 5](#) gives us the number of obscure and erroneous relationships that were detected by PCM and MPM, respectively.

To get an idea of obscurity and error reflected by PCM, consider [Figure 3](#), which contains the PCM drawn by subject #3. The obscure link in this PCM is the statement that the expected value of the estimator equals the parameter. Does the subject purport to say that in that case we have an unbiased estimator? This interpretation seems contradicted by the right part of the concept map, in which we read, again somewhat cryptically (which seems connected to the format of concept mapping, in which concise

Table 5 – Number of obscure and erroneous relationships detected by PCM and MPM

	PCM	MPM
Obscure	1	8
Erroneous	2	10

comments are to be placed by the links between concepts) "estimator...this is....unbiased....approaches....parameter". This can only be erroneous and it suggests to us that in the left part of the PCM, the subject does not really comprehend that the notion of unbiasedness is linked to the expected value of the estimator. So the statement "estimator....expected value....is equal to....parameter" is probably reflective of erroneous structural knowledge, but it does take some interpretation to come to this conclusion.

Let us see what this same subject says on this matter in her MPM. We read:

"An unbiased estimator approaches the true value of the population very closely. This way the expected value of the estimator equals the parameter".

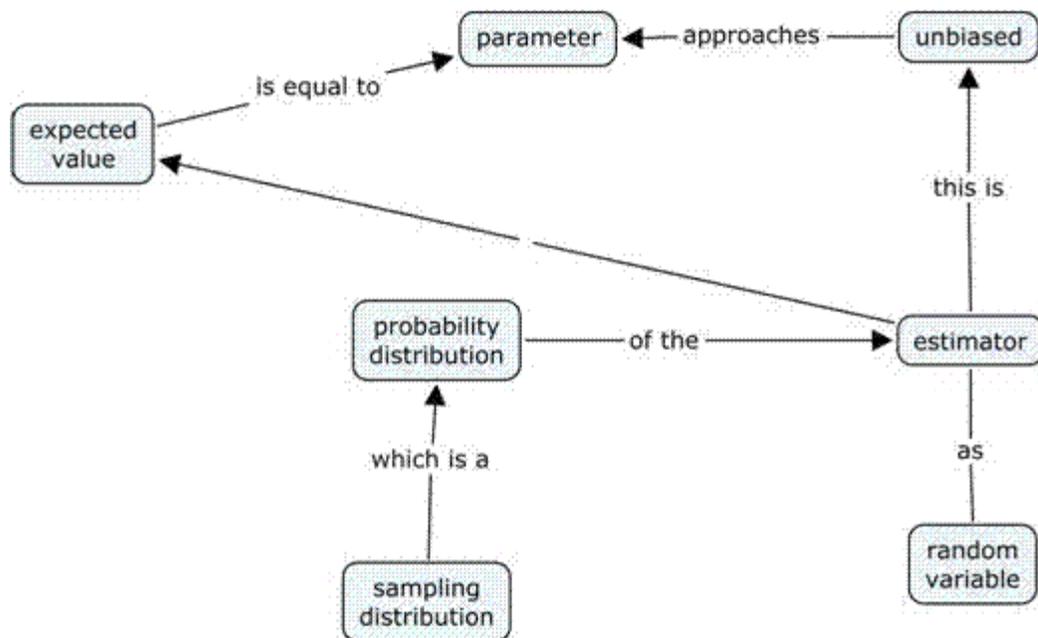
Interestingly, the actual wording in the MPM closely resembles the comments in the PCM, but the connecting words

"This way" makes it far more clear how the subject has actually organized her knowledge. It is now evident that the subject believes unbiasedness to reflect a mixture of dispersion (small standard error) and central tendency (expected value equals parameter value), which is wrong.

Scanning for other dubious links in the MPM of this subject, we come across the following obscure relationship:

"From the sampling distribution you can derive a percentage on the composition of the population, which reflects a probability distribution".

Figure 3 – PCM by subject #3



This is an incomprehensible statement which does not become any clearer in the context of the total argument. It suggests that although the subject understands that there is a link between sampling distribution and probability distribution (as displayed in the PCM, where the subject simply and seemingly correctly states that 'a sampling distribution is a probability distribution'), she does not fully comprehend the nature of that relationship. Because concept mapping asks us for concise comments, relationships can emerge that upon closer examination by a method revealing more detail, are shown not to have been properly understood.

The subject concludes her MPM with a statement which is clearly erroneous. Apparently, our subject believes unbiasedness to be somehow equated with normality:

"If the sample is large enough and drawn at random you can be sure that you have an unbiased estimator because the sampling distribution then becomes normally distributed".

Consideration of the MPM therefore gives a far more detailed picture of the knowledge organization than the PCM does. This fact seems to be reflected in the figures of table 5, which show that MPM reveals a lot more obscurity and errors than PCM does. Many of the obscure links that emerge in MPM tasks seem related to a difficulty of the subject to express his or her knowledge into words. Consider the following examples, taken from different subjects:

"A sampling distribution is when you compute the mean of the mean of different samples of the same size"

"All these estimators could be possible parameters"

"The probability distribution of this estimator, the distributional curve that allots probabilities and a certain area to each interval, is the sampling distribution"

"The sampling distribution is the distribution of all the values of all the samples that you could draw from

the same population"

Such wordings seem to reflect an intuitive, but not a firm and structurally embedded understanding of the concept of sampling distribution. A concept map is not likely to reveal such slippery understanding, when all a subject has to do is to draw a line between 'sampling distribution' and 'probability distribution', and comment on the link by adding 'is' to the arrow linking the two concepts.

4. Discussion and conclusions

In our effort to determine the level and organization of the statistical knowledge that a student has gained at the end of a course, we sometimes resort to the use of assessment methods that are specifically devised to probe structural knowledge. One of the most widely known of such methods is concept mapping. However, for the purpose of assessment, valid use of this method requires considerable effort in training students to use it properly, and it is acknowledged that some students never get the hang of it. In addition, concept maps tend to convey a very sketchy impression of knowledge organization and because of this, tend to provide information that can sometimes be very ambiguous. In this study, we explored the potential of an alternative method for assessing structural knowledge. We expected the MPM assessment method, as the alternative is called, to be easier to use by students (in other words, to require less training time than concept mapping) and to provide less ambiguous information. The latter expectation was based on the fact that, whereas the concept map requires a transformation by the student of verbally acquired information into an abstract graph, doing the MPM task does not require any such sort of transformation.

We believe the results of the study demonstrate that it is indeed not difficult to instruct students to produce an informative MPM response. All students had worked on the MPM assignment in a correct fashion, that is to say they all provided some sort of an argument and thereby some insight into structural knowledge. Of course, it helps even further if during the course students are confronted with MPM assignments as an aid to instruction. Indeed, in our own elementary statistics courses for psychology students we regularly make use of MPM assignments to help the students focus their attention on the relevant structural aspects of the theory. On the other hand, in the construction of concept maps nearly half of the students in the sample produced maps that were, to a certain extent, technically incorrect. That is, connecting lines were frequently undirected or remained without comment. This finding seems to underline the often noted fact that learning to draw an informative concept map requires a considerable amount of training.

Contrary to what we had at first expected, the number of obscure links was clearly greater in MPM tasks than in PCM tasks. However, upon reflection this is not a weakness of the MPM task but on the contrary a point of strength. Drawing a single line between two concepts in a concept map and adding 'is' as comment may suggest correct comprehension of a relationship, but of course this simple drawing yields very little information. In the MPM, the student is instructed to make his perception of a certain relationship explicit in the argument and it seems plausible that a fuzziness in the description of the relationship is indicative of a true fuzziness in the structural organization of the student. This interpretation is aligned to the familiar fact that true comprehension of a subject is shown by your ability to explain the material clearly to somebody else. In addition, we found that MPM succeeds in clearly exposing the existence of misconceptions on the part of the student, whereas the PCM rarely gives a clear indication of faulty knowledge. This is all the more surprising in view of the small number of concepts that needed to be included in the map. One would expect that such a small map would facilitate the exposition of faulty links in the structural organization of the student, but contrary to the MPM task, it almost never did.

Although we believe MPM to be promising as a method for assessment, there are a few drawbacks. In particular, while MPM yields rich information, not all the presented structural information is easy to qualify as either correct or incorrect. As we saw in the discussion of the material of subject #8, sometimes some interpretation is required to decide whether or not a given relationship has been demonstrated. However, this is equally true of concept mapping and even far more so when dealing with responses to open questions. And although it is true that sometimes some interpretation is necessary, in the majority of cases the MPM material permits a pretty clear cut decision on whether certain relationships between concepts have been grasped or not.

Apart from the fact that, like in the case of concept mapping, some interpretation on the part of the assessor will

sometimes be necessary, another limitation on the use of MPM concerns the amount of time needed by a student to actually do the task, or a couple of such tasks in a row. Like concept mapping, doing an MPM task thoroughly will demand some time on the part of the student. This demand will increase with each additional proposition that we would like to see integrated into an argument. For this reason, rather than turning the complete assessment into an array of MPM tasks, it is more practical to do a multiple choice test for overall assessment of the *level* of conceptual understanding, and to complement such a test with a couple of MPM tasks to gain insight into the structural organization of this understanding.

Multiple choice tests are easy to use and permit an efficient process of automated scoring. Various multiple choice items on the ARTIST website attest to the fact that this form of testing can be used for the assessment of conceptual understanding. However, although such tests do show us differences between students in their level of conceptual understanding, they do not show just how an individual student has structurally organized his or her knowledge. The test score can be considered as the fruit of the structural organization of knowledge: the better this structure, the higher the score will be. The test score does not tell us anything about the actual structure.

Not uncommonly therefore, in addition to a multiple choice test, instructors also provide their students with a number of open questions. The responses to these open questions, it is hoped, will give us a better insight into the way that the students have organized their knowledge. To this end, such open questions are usually presented with the instruction to "explain (or motivate) your answer". Unfortunately, this open question format usually evokes very idiosyncratic responses. The students have a lot of freedom in the way they wish to respond and different students may focus on very different aspects of knowledge. Consequently, two responses may both be correct (and motivated), yet differ widely in the amount or clarity of information that is provided. Our present research underlines this often encountered experience and shows that open questions rarely yield the sort of propositional or relational knowledge that we would like to see displayed.

Compared to open questions, MPM tasks are highly structured and this study has shown that as a consequence, they tend to yield the richness of information that we often look for in vain in the responses to open questions. In the MPM tasks, the instructor aims to examine the ability of the student to demonstrate insight into the relationship between key concepts or propositions. In elementary statistics, for example, the concept of sampling distribution is of overriding importance. Without good comprehension of this concept, subsequent knowledge of inferential statistics will lack a proper foundation. We may use MPM to examine to what extent this concept has been adequately understood by having it related to a number of other concepts or propositions of key importance. This will yield information that will not be easily captured by a multiple choice test, nor by an unstructured open question.

In the MPM task that was used in this study, all the questions that the students needed to integrate into their arguments were on definitions of concepts. However, questions in MPM tasks need not be restricted to the definitional sort. They could be on formulas (e.g. 'What is the range of values that the correlation coefficient can have?') on limitations of certain statistics ('In which cases will the mean not be a good measure of central tendency?'), on terminology ('Why is the p -value of the test statistic called a conditional probability?'), etc. After all, MPM is about relating propositions to each other, rather than strictly relating concepts, as in concept mapping.

In our own courses, we habitually analyze each separate lecture (and corresponding literature) into its constituent propositions. These propositions are on the elementary concepts, principles, ideas etc., that we wish to put across. For instance, an introductory lecture on hypothesis testing might focus on propositions such as "What is a test statistic?", "What do we mean by the sampling distribution of a test statistic?", "Why is this sampling distribution a conditional probability distribution?". Our experience is that, typically, a lecture can be decomposed in about 35 such propositions. Having an overview of such constituent propositions, makes it relatively easy to select a subset of these for the construction of MPM tasks.

Although MPM tasks are designed to probe the structural organization of conceptual knowledge, their use need not be restricted to a purely verbal exercise such as was used in this study. MPM tasks may for instance also be used to structure a discussion on a graph or table. Quite often, students restrict their inspection of graphics or tables to certain key features. For instance, in the ANOVA table a student may restrict attention to the p -value, and ignore information that could be used for reflecting on effect size. Or a student may be summarizing a distribution in a histogram by

reporting the mean and standard deviation, completely ignoring the fact that the presented distribution is strongly skewed or contains marked outliers. By coupling an MPM task to a graph or a table the student may learn to develop a broader focus in the inspection of such information. In the assessment phase, MPM can help to reveal whether or not the student is indeed able to use this broader perspective.

Endnotes

¹For the purpose of this article, the hand drawn concept map has been reproduced by Cmap Tools. In this program, all links run from top to bottom and no arrows are shown because the direction of the relationship follows from the vertical ordering of concepts. Only lines that run sideways or relationships that are directed from bottom to top will appear as explicit arrows in maps created with Cmap Tools.

Appendix 1: coding the PCM

1) A relationship between two concepts C1 and C2 has been demonstrated only by an explicit link between C1 and C2. The link should be directed, but if undirected may still be taken as evidence for correct cognition, if the comment placed by the link unambiguously suggests correct comprehension.

Insufficient would be:

EXPECTED VALUE ----- in -----SAMPLING DISTRIBUTION

Because it remains unclear whether the subject has comprehended the link between 'expected value' and 'sampling distribution'.

Sufficient would be:

ESTIMATOR ----- has ----- SAMPLING DISTRIBUTION

Because it is unlikely that the subject would have organized his knowledge so as to believe that a sampling distribution has an estimator.

2) A correctly directed link between two concepts C1 and C2 with a nonsensical comment will not be considered as prove of correct comprehension.

For instance:

ESTIMATOR ----- because of -----> RANDOM VARIABLE

Suggests improperly organized knowledge.

3) Concepts that are not linked will be assumed unlinked in the cognitive organization of the subject.

4) Concepts that are linked in either a directed way or an undirected way, but without any comment placed by the link cannot be assumed to be cognitively correctly organized. They may be, but the missing datum prohibits us from drawing any conclusion in this regard.

Application of the coding scheme to the PCM of subject # 8 (see Results-section):

Because of failing links, R1, R3, R5 and R6 have not been demonstrated. All links are directed and commented. R2 and R4 have thus been demonstrated unambiguously. The links between 'Expected value', 'parameter' and 'unbiased'

take some interpretation, however it seems reasonable to assume that the subject has correctly demonstrated knowledge of R7.

Appendix 2: coding the MPM

1) Correct comprehension of a proposition demands an unambiguously correct answer to the question that pertains to the proposition. If the answer is brief but correct, it will be taken as proof of comprehension. If the answer is elaborated, largely correct but as some points confusing, this will not be taken as an unambiguous demonstration of correct comprehension.

Sufficient would be:

"An estimator is a random variable"

Insufficient would be:

"An unbiased estimator approaches the true value of the population very closely. This way the expected value of the estimator equals the parameter".

This subject seems to contradict himself by stating both that an unbiased estimator will have an expected value equal to the parameter value, and by denying this same assertion in the preceding line.

2) Correct comprehension of a relationship demands an unambiguously correct wording of this relationship. If the wording is brief but correct, it will be taken as proof of comprehension. If the wording is elaborated, largely correct but as some points confusing, this will not be taken as an unambiguous demonstration of correct comprehension.

For example:

"A sampling distribution is a probability distribution" would be considered correct.

However, we would consider the following, more elaborate statement to be incorrect:

"From the sampling distribution you can derive a percentage on the composition of the population, which reflects a probability distribution".

This statement is so unclear that it is not even certain that the subject means to imply that a sampling distribution is a probability distribution.

3) A relationship that has been correctly stated will be taken as a correct part of knowledge organization, even though the subject may have shown incorrect comprehension of the propositions that were involved.

For example:

A subject states 'a sampling distribution is a probability distribution', but elsewhere in the MPM shows to have a faulty conception of what a sampling distribution actually is. In this case, knowledge of the proposition on the sampling distribution has not been demonstrated, but knowledge on the relationship between sampling distribution and probability distribution has been demonstrated.

Application of the coding scheme to the MPM of subject # 8 (see Results-section):

This subject states that the sample mean is an estimator. He does not say that the sample mean can be used as an estimator for the population mean. R1 has therefore not been unambiguously demonstrated.

R2 has been unambiguously worded.

R3 is not mentioned in the text.

R4 has been unambiguously worded ("In the sampling distribution of the sample mean, the expected value is μ "). However, R5 is not. The subject talks about the probability distribution of the sample mean, and next starts to discuss the sampling distribution of the sample mean. Whether or not he realizes that he is talking about the same, cannot be decided unambiguously on the basis of the text. This shows the importance of a clear instruction to the MPM tasks: if you believe two concepts to be related, you should say so explicitly.

R6 is not stated, instead the subject describes a computational formula.

R7 is unambiguously stated in this MPM.

References

- Broers, N.J. (2001). "Analyzing Propositions Underlying the Theory of Statistics". *Journal of Statistics Education*, [Online]. 9 (3).(<http://www.amstat.org/publications/jse/v9n3/broers.html>)
- Broers, N.J. (2002a). "Selection and Use of Propositional Knowledge in Statistical Problem Solving", *Learning and Instruction*, 12 (3), 323-344.
- Broers, N.J. (2002b). "Learning Statistics by Manipulating Propositions". *Proceedings of the Sixth International Conference on Teaching Statistics*, Capetown, South Africa.
- Broers, N.J. and Imbos, Tj. (2005). "Charting and Manipulating Propositions as Methods to Promote Self-explanation in the Study of Statistics". *Learning and Instruction*, 15 (6), 517-538.
- Broers, N.J., Mur, M.C. and Bude, L. (2005). "Directed Self-explanation in the Study of Statistics". In: G. Burrill & M. Camden (eds.) *Curricular development in statistics education*. (pps. 21-35). Voorburg, The Netherlands: International Statistical Institute.
- Broers, N.J. (2007). "Designing Open Questions for the Assessment of Conceptual Understanding". *Proceedings of the IASE Satellite Conference on Assessing Student Learning in Statistics*. Guimaraes, Portugal.
- Brown, L.T. and Stanners, R.F. (1983). "The Assessment and Modification of Concept Interrelationships". *Journal of Experimental Education*, 52, 11-21.
- Bude, L.M. (2007). *On the Improvement of Students' Conceptual Understanding in Statistics Education*. Unpublished doctoral dissertation. Maastricht University.
- Geesling, W.E. and Shavelson. R.J. (1975). "Comparison of Content Structure and Cognitive Structure in High School Students' Learning of Probability". *Journal of Research in Mathematics Education*, 12, 109-120.
- Hiebert, J. and Lefevre, P. (1986). "Conceptual and Procedural Knowledge in Mathematics: an Introductory Analysis".

In J. Hiebert (Ed.), *Conceptual and Procedural Knowledge: the Case of Mathematics* (pp. 1-27). Hillsdale, NJ: Erlbaum.

Huberty, C.J., Dresden, J., and Byung-Gee, B. (1993). "Relations Among Dimensions of Statistical Knowledge". *Educational and Psychological Measurement*, 53, 523-532.

Jonassen, D.H., Beissner, K. and Yacci, M. (1993). *Structural Knowledge: Techniques for Representing, Conveying, and Acquiring Structural Knowledge*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Kaplan, J.J. (2006). *Factors in Statistics Learning: Developing a Dispositional Attribution Model to Describe Differences in the Development of Statistical Proficiency*. Unpublished doctoral dissertation, University of Texas.

Kelly, A.E., Finbarr, S. and Whittaker, A. (1997). "Simple Approaches to Assessing Underlying Understanding of Statistical Concepts". In: I.Gal and J.B. Garfield (Eds.), *The Assessment Challenge in Statistics Education*. Amsterdam: IOS Press.

Lampert, M. (1986). "Knowing, Doing, and Teaching Multiplication". *Cognition and Instruction*, 3, 305-342.

Paas, F. (1992). "Training Strategies for Attaining Transfer of Problem-Solving Skill in Statistics: A Cognitive-Load Approach". *Journal of Educational Psychology*, 84, 429-434.

Ruiz-Primo, M.A. and Shavelson, R.J. (1996). "Problems and Issues in the Use of Concept Maps in Science Assessment". *Journal of Research in Science Teaching*, 33 (6), 569-600.

Ruiz-Primo, M.A., Shavelson, R.J., Li, M. and Schultz, S.E. (2001). "On the Validity of Cognitive Interpretations of Scores From Alternative Concept-Mapping Techniques". *Educational Assessment*, 7 (2), 99-141.

Schau, C. and Mattern, N. (1997a). "Assessing Students' Connected Understanding of Statistical Relationships". In: I.Gal and J.B. Garfield (Eds.), *The Assessment Challenge in Statistics Education*. Amsterdam: IOS Press.

Schau, C. and Mattern, N. (1997b). "Use of Map Techniques in Teaching Applied Statistics Courses". *The American Statistician*, 51 (2), 171-175.

Nick J. Broers, Ph.D.
Dept of Methodology and Statistics
Maastricht University
E-mail: Nick.Broers@stat.unimaas.nl

[Volume 17 \(2009\)](#) | [Archive](#) | [Index](#) | [Data Archive](#) | [Resources](#) | [Editorial Board](#) | [Guidelines for Authors](#) | [Guidelines for Data Contributors](#) | [Home Page](#) | [Contact JSE](#) | [ASA Publications](#)