Science & Education (2007) 16:291–311
DOI 10.1007/s11191-006-9021-2

Understanding the Role of Measurements in Creating Physical Quantities: A Case Study of Learning to Quantify Temperature in Physics Teacher Education

TERHI MÄNTYLÄ and ISMO T. KOPONEN
Department of Physical Sciences, University of Helsinki, Helsinki, 00014, Finland
E-mail: terhi.mantyla@helsinki.fi

Abstract. Learning to understand the content and meaning of physics' concepts is one of the main goals of physics education. In achieving this understanding, the creation of quantities through quantitative measurements, or rather through quantifying experiments, is a key process. The present article introduces a didactical reconstruction for understanding the construction of the meaning of physical quantities from a network point of view, where the quantities are part of networks and the quantifying experiments build up these networks. As a practical example, we discuss how the quantity temperature is constructed in an instructional unit designed for student teachers and what the learning outcomes are.

Key words: didactical reconstruction, measurements, teacher education, temperature

1. Introduction

In physics education, *conceptual understanding* refers to the understanding of the content and meaning of concepts, with an emphasis on qualitative understanding. Part of conceptual understanding is to learn to understand 'how we know what we know'. In physics, knowledge of 'how we know' goes back to question of how concepts acquire their meaning and empirical support. This is inherently related to measurements, which transform the concepts into measurable *physical quantities*. The formation of quantities through quantitative experiments or measurements is not often seen as a part of conceptual understanding, or at least it is treated as a non-problematic part of it. The purpose of this paper is to discuss the advantages that can be gained by focusing on the role of quantitative (or quantifying) measurements. This is done in the context of physics teacher education, because the teachers in particular need to be able to answer the question of 'how we know what we know', and moreover, be able to reflect this understanding in their teaching.

In reaching the goal of better conceptual understanding, the history and philosophy of science (HPS) can serve us in many valuable ways; the conceptual analysis of physics history helps us to regenerate the knowledge of physics by answering the questions: how did we come to believe what we believe, and how did we discover what we know (compare with Chang 2004, pp. 236–240). However, our purpose is not to produce historical reconstructions, but instead to use HPS as a starting point for developing and designing suitable didactical solutions, which can be called didactical reconstructions for teaching (Izquierdo-Aymerich & Adúriz-Bravo 2003). For this purpose we introduce here an epistemological reconstruction for understanding the construction (or rather, reconstruction) of physical quantities. Our approach introduces quantities as part of the networks of other quantities and laws, where quantifying experiments are seen as having a central role in building up the network and in determining its structure.

As an application, we discuss the (re)construction of temperature as a measurable quantity. This quantity was selected for the present case study for several reasons. Temperature is one of those physical quantities which is used daily. It is easily measured using thermometers and thus we see its measurement as quite unproblematic. Nevertheless, this simple concept seems to pose many learning difficulties for university students (cf. Carlton 2000; Taber 2000; Cotignola et al. 2002; Meltzer 2004). Moreover, textbooks quite often connect temperature straightforwardly to the average kinetic energy of particles, and so they reduce temperature to mechanical quantities and 'explain' it through the microscopic atomic model. This, however, does not help in understanding how, after all, temperature as a macroscopic concept and quantity is formed. Moreover, this 'reduction' too easily leads to the oversimplified ideas of temperature devoid of macroscopically meaningful content (cf. e.g. Meltzer 2004). Therefore, in the case of the question how do we know what we know about temperature, this oversimplified microscopic model and mode of thinking becomes very problematic and may act as a barrier to further learning. In understanding temperature as a macroscopic quantity, it is essential to understand how it becomes a measurable quantity in the first place. Consequently, the didactical reconstruction discussed concentrates on temperature as a macroscopic quantity. In many respects, the reconstruction parallels the historical development of temperature (for a detailed exposition see e.g. Chang 2004), but its intention is not to be an entirely historically authentic reconstruction. Instead, the history is interpreted from the point of view of modern conceptions, because the goal, after all, is to teach physics, not the history of physics. Nevertheless, we believe that this reconstruction is not only a useful tool for learning, but that it also conveys a more correct view of the role of measurements in the production of scientific knowledge than most standard textbook presentations or teaching solutions based on them.

2. Theoretical Background: Reconstruction and Networks

The history of physics provides a good source for understanding how different concepts and quantities have been constructed and abstracted. A conceptual analysis of historical development is needed, even in this case, where the purpose is to produce teaching solutions fostering the development of modern conceptions rather than giving an authentic picture of historical developments. The uses of historical analysis for better conceptual understanding is well exemplified in the work of Chang (2004), who has given an account of the role of conceptual and philosophical analysis of physics history with the goal of understanding knowledge creation and the justification of knowledge claims. He shows how closely the ways to measure are connected with the development – or according to Chang, the 'invention' – of the concept, and how much more delicate this process is than a straightforward 'operationalisation'. The message of Chang's analysis for our work is that in order to know 'how we know what we know' we need to pay attention to the practical methods of making the quantities measurable in the process of 'inventing' concepts. In the justification of the concept, there is a question about reaching a certain 'experimental closure' where the invented concept helps to make sense better of the experimental results and measurements. Chang's work is connected with the history and philosophy of science, but the basic themes introduced in it about the role of measurements parallel quite well the role of measurements and quantitative experiments in teaching and learning physics. Also in this case we can see concepts as the outcomes of the process where experiments and theoretical abstraction processes are intertwined and where theory becomes transformed in course of developing measurement methods. By relying on this viewpoint of quantitative measurements, we have designed a didactical reconstruction for teaching the construction of physical quantities. These kinds of didactical reconstructions can be designed within a broader framework, which we have called 'generative experimentality' (Koponen & Mäntylä 2006). The reconstruction does not try to repeat the historical experiments as such; instead it reconstructs the process of knowledge creation from the current point of view of physics, and attempts to retain those steps of the historical process which are crucial for understanding knowledge creation and which can be recognised through the conceptual analysis of historical developments.

The aspect of generative experimentality of interest is related to the process of what is often called the 'operationalisation of concepts', but here it is understood in a broader sense, not only making the existing theoretical concepts measurable but also creating or inventing every concept through a process of measurement (cf. Chang 2004, pp. 241–243). There the qualities are transformed into quantities through quantifying experiments and the experiments then assign the empirical meanings to the quantities and laws. According to our reconstruction, the building of the meaning of quantity always involves a chained set of experiments and measurements, in which each new experiment builds on the results of the previous ones. Quantities thus become networked and their meanings connected, and a structured network, where quantities and laws are the nodes, is produced. Taking into account the role of quantifying experiments and quantitative measurements in creating this connected net of quantities (and laws) equals giving attention to the very skeletal features of physics knowledge structures; from this notion derives our conviction of the importance of such experiments in teaching physics.

The network view of quantities is meant to make the interconnectedness of quantities (and concepts, of course) explicit, which is quite often only implicitly referred to in traditional teaching. Astonishingly this notion provided is little used in designing teaching solutions for physics education (some exceptions are, however, by Reif 1987, 1995; van Heuvelen 1991; Bagno et al. 2000) although it is well established in many philosophical accounts. For example, Thagard proposes concepts as complex structures, like frames or scaffoldings, where a special role is given to interrelations between concepts and where 'rules connected to concept are parts of them as well concepts are part of the rules' (Thagard 1992, pp. 29–30). This is also the viewpoint advocated by Giere, who notes that 'there are underlying principles which create a network of causal and explanatory links which hold individual concepts together and provide connections with related concepts' (Giere 1999, p. 105). Similar notions are also found in works concerned with cognitive aspects of learning, and, for example, diSessa and Sherin (1998) note that concepts 'get their meaning by participating in a web of relations with other concepts'. The network view is therefore meant to explicate the notion that concepts cannot be defined semantically or in isolation from other concepts. Moreover, the network view of concepts also makes it possible to maintain the concepts as structures open to development, which is not only an important aspect of successful physics teaching but also a necessary condition for the progress of physics itself.

The above notions of the concept network are directly related to our design principle for the graphical representations used to display the networks of quantities. In these network representations (NRs), each node corresponds to a quantity or law and each link corresponds to a relationship between quantities. The quantifying experiments have a central role when different quantities and laws are connected; they also assign order to concepts, and therefore the links are directed ones. From the viewpoint of learning and teaching, the advantage of this view is that such structures can be meaningfully analysed as a network of nodes. In particular, the graphical NRs are easily adapted in teaching and learning and their structure can be used to indicate the student's learning during the teaching sequences. The NRs externally resemble concept maps developed by Novak (Novak & Gowin 1984), or some aspects of graphic organisers introduced by Trowbridge and Wandersee (1998); they are two-dimensional node-link representations. However, the principles of producing the graphical NRs introduced here distinguish them from other types of graphical NRs.

3. Reconstruction Contextualised: Temperature

The didactical reconstruction of the development of temperature as a quantity was guided (but not constrained) by historical development (for historical development, see Erlichson 2001; Chang 2004), though ultimately constructed taking into account the modern conceptions of temperature. Based on the general scheme behind such reconstructions (for details, see Koponen & Mäntylä 2006), the phases of development of temperature were divided into three stages as follows.

1. *Level of qualities*: Temperature, T_{Sensory}, is connected with the sensory experience of warmness. Variations in the *degree of warmness of bodies* are distinguished as well as the formation of temperature equilibrium when bodies are in contact or fluids are mixed. The formation of temperature equilibrium leads to a conception of temperature as an *intensive property* of the system; i.e., temperature does not depend on the size of the system. This is observed through the notion that temperatures of fluids or bodies at different temperatures never add up, instead, the temperature is between the extremes. The idea of temperature as a state variable and intensive quantity is based on the changes in state (freezing, melting, and evaporating). The formation of the quantity of heat as an *extensive property* of a system (i.e. the quantity of heat depends on the size of the system) was also discussed: a larger substance needs larger amounts of heating (e.g. by burning gas or by electrical heating). This difference between intensive and extensive properties related to the experience of warmness acts as a starting point to distinguishing the temperature and the quantity of heat as different quantities characterising different properties of the system.

2. *Level of quantities and laws*: The quantifying experiments of temperature in terms of thermal expansion of liquids, liquid thermometers and scales are introduced. There are several possibilities for relating the changes in volume of the liquids to the changes in their 'degree of heat', e.g. by mixing known amounts of the same liquid with different degrees of heat. In this way, the 'ordinal scales' where comparison of degrees of heat is possible can be transformed to 'cardinal scales' through changes in volume. The introduction of the cardinal scale allows discussing the magnitudes between differences, and then it also becomes possible to perform calculations (for more detailed discussions, see Chang 2004, pp. 86–87). This is achieved by assuming a *linear proportionality* between these changes which finally allows constructing the temperature, T_{Liquid}, through measuring the changes in volume only. However, upon closer inspection, through comparing the behaviour of different substances, the assumption of linear proportionality turns out to be untenable. In order to obtain better and more reliable ways to measure the temperature, the *empirical gas laws* (based on measurements using liquid thermometers) and the regular invariances contained within them are considered as a better possibility. Experiments with different gases show that now the assumed linear proportionality between changes in the mechanical quantities of volume and pressure is nearly proportional to temperature changes when measured with the scale T_{Liquid}, but systematic deviations occur when different thermometers are used. Nevertheless, the results suggest that empirical gas laws, and even better, *ideal gas law* which abstracts and generalises the empirical laws, can be taken as a basis for a new gas thermometer scale. This now provides a new way to operationalise temperature, which is more reliable. Moreover, the possibility of an absolute scale with a common reference point is noted. The idealised nature of this new scale and its dependence on the model of ideal gas is analysed. The idea of pushing idealisation further to ideal gas, raise the level of abstraction, but temperature is still tied to the substance although now through an idealised model of ideal gas, which is realised in practice as a diluted real gas.
3. *Level of structured theory*: In order to emancipate the concept of temperature from the substantial world (liquid or gas) and abstract it further, the axiomatised theory, or theory of principles, is needed (cf. Chang 2004, pp. 173–154 and pp. 183–186). The general principles of thermodynamics and ideal Carnot cycle as a basis for *defining* absolute temperature are discussed from the viewpoint of how theory guides the construction of the concept. The *conserved quantities*, energy and entropy, are needed in order to reach a substance-independent concept of temperature (Erlichson 2001; Chang 2004 pp. 183–186). However, even

now one needs a model of how, in principle, under ideal conditions, the operationalisation could take place. This idealisation is an ideal heat engine, working with an *ideal reversible* cycle. When an *ideal* heat engine operates between temperature sources at different temperatures, the necessary requirement is that the *amounts of heat* exchanged are in proportion to the temperature changes. This is based on the conservation of energy and complete reversibility (i.e. the idea that the initial stage can be recovered). The operation of an ideal heat engine forms the basis for the *absolute temperature scale*. (The temperature, T_{Absolute}, is defined through this assumed strict proportionality.) Afterwards, it can be shown that this coincides with the ideal gas scale (which also, of course, is an abstraction). However, the temperature is now free from any substance, and it is based only on the conservation laws and the general idea of reversibility. From this viewpoint, the conserved quantities *energy* and *entropy* are fundamental. Temperature (or inverse temperature) becomes a measure of how the *change* in internal energy is connected with the change in entropy. From a modern vantage point, this can be seen as the basis of Kelvin's scale (and it is now this more modern interpretation rather than Kelvin's original which is taken into account in the reconstruction). The next step of abstraction would be temperature understood as reduced to microscopic theory, and as an emergent, macroscopic quantity based on a probabilistic conception of many body systems' properties. This ultimately reduces such a temperature as T_{Stat} to quantities related to mechanics (and thus, to statistical mechanics). Then the negative absolute temperature also becomes meaningful. Temperature, understood in this way, differs greatly from temperature at the level of macroscopic laws.

As we noted previously, the didactical reconstruction needs to deviate from the actual historical course of events in several respects. In particular, the position given to the empirical laws comes from a retrospective (and ahistoric) viewpoint; historically, the possibility to make the empirical results definitive presupposed the development of many of the theoretical ideas which only in retrospect can be seen as abstracted from these empirical laws (cf. discussion in Chang 2004). In addition, a reinterpretation of Kelvin's way of defining the absolute temperature scale is utilised, because it is the way that temperature is understood in statistical mechanics and consequently addressed in teaching at the advanced level today. We think that this is not only justified but also necessary for the successful learning of physics. Therefore, the order of conceptualisation as introduced in the reconstruction is motivated more from the point of view of the current knowledge structures of physics – which, after all, should be the goal of the learning process – than from the point of view of authentic history. Nevertheless, HPS has been used to recognise the crucial cognitive breakthroughs needed to establish temperature in all three stages in its development as included in the reconstruction.

3.1. TEACHING SEQUENCE

The students who were taking part in the teaching sequence attended a course entitled *Conceptual foundations of physics*, which is a half-year course for third year physics teacher students (for more details about the course, see Koponen et al. 2004; Koponen & Mäntylä 2006). As a background, they had standard introductory-level physics courses, including thermodynamics, like all other physics students. Therefore it is not expected that there be any fundamental differences between them and other students. The initial conceptions they have about temperature and problems with it are expected to be quite similar to those reported in the research literature; temperature being not only connected and explained but also 'reduced' to atomistic motion and kinetic energy, with only a vague idea of how temperature in principle becomes a measurable quantity. This expectation, although not thoroughly researched, was confirmed by initial class discussions with students. In summary, they tend to think that temperature is a measure of kinetic energy that is measured using a thermometer. This conception is a good starting point, but hardly enough for prospective teachers.

The implementation of the instructional unit started with a lecture (approx. 45 min). The lecture contained the aspects of thermometry and thermodynamics stated above, but without the explicit division into three different levels or 'different temperatures'. Throughout the instruction, temperature was discussed as a 'single' quantity, in the way it is usually done in any traditional thermodynamics course. Moreover, the subject content (required physics) has been taught in an introductory thermal physics course earlier, so the teaching during this sequence was a more run-through-like reiteration of already learned contents. In the course, however, the ascending levels of concept formation were discussed from general philosophical and epistemological points of view. Also the role of measurements in that process was discussed at a general level.

After the lecture, the task of the students was to produce individually the initial NRs about the quantitative development of temperature, and to recognise the essential experiments, models and theory needed in that development. The technical aspect of drawing NRs (boxes, lines and directionality of lines, meaning of arrows and two-headed arrows etc.) were familiar to the students, and they had earlier produced NRs about mechanics and heat and energy, for example. The way of producing these NRs was also based on the reconstruction, so the students had experience in the principles of constructing the NRs. However, the context was different in those earlier cases than in the case discussed here. Therefore, we believe that the differences between initial representations and final ones are not simply ascribed to a situation wherein during the teaching sequence students learned to express ideas in the form of graphs, as they are expected to display them, but instead say something about the organisation of the content knowledge itself that they are representing by using the NRs.

In the exercise session (90 min), the individually made initial representations were further developed collaboratively in study groups of two or three students. Students were asked to discuss their initial representations and to start constructing the final NR on the basis of the group members' initial NRs. Basically the students chose one initial NR as a basis and combined with it the ideas from other NRs, and these ideas were together developed further. The instructor followed the groups' work and discussion and, if needed, helped the students forward. At the end of the exercise sessions there was a short instructor-led summary discussion. The students produced the final NRs on the basis of their own ideas but their work was facilitated through the lectures, exercises and collaborative work in the small study groups. The students were asked to complete the final NR about the quantitative development of temperature as homework and return it to the instructors.

In order to see what the advantages of our reconstruction are, we formulated the following research questions, of how students do the following during their learning process:

1. make reference to the quantitative experiments in the case of temperature,
2. express the various steps involved in temperature's construction, and
3. evaluate the usefulness of the reconstruction.

These research questions are answered in what follows on the basis of empirical data gathered during the teaching sequence and analysed using a qualitative, interpretative research approach, as will be explained next. During the teaching sequence, collaborative learning methods were used; however, the focus of the research was on the outcomes of the teaching sequence, on the learning results of individual students.

4. Results

The research data to be discussed here was gathered during autumn 2001 from the course *Conceptual foundations of physics* mentioned above. The data consists of the students' initial and final NRs and interviews. In order to clarify how the students understood the construction of the quantity temperature, two researchers analysed their initial and final NRs. NRs are suitable for our purposes because it is known that a set of successive NRs reflects changes in understanding (Trowbridge & Wandersee 1998, p. 116). Altogether there were 24 students. They produced the initial NRs independently and in the case of the final NR got to choose whether to do it in a small group of two or three students or alone. So there were 24 initial and 14 final NRs (from which 8 NRs were produced individually, four in a group of three students and two in a group of two students). Furthermore, we decided to interview five students because we were interested in how well our analyses of NRs represent the interviewed students' thinking and to see in more detail how the students utilised the reconstruction.

4.1. ANALYSIS OF NETWORK REPRESENTATIONS

We examined how the different concepts appeared and were used in NRs. We concluded that the concepts (with their interconnections) in the NRs could be classified first into three main categories reflecting the three levels of abstraction introduced in Section 3: level of qualities, level of quantities and laws, and level of structured theory. The sub-categories were formed based on the concepts in the NRs, and our background knowledge of subject content (see Section 3) was also used in the forming of sub-categories. The analysis was based on the following task given to the students: to quantitatively develop temperature while recognising the essential experiments, models and theory needed in that development. Some of the sub-categories are desirable for successful learning (e.g. thermal expansion of liquids or solids, empirical gas laws, ideal gas law), while others are an ambiguous use of concepts. The categories of concepts are presented in Table I. By ambiguous use, we mean that in students' NRs there are concepts which do not relate to the given task; i.e., there is no way to infer how the concept relates to the quantitative development of temperature. We tried to minimise the possibility that this ambiguity is due to an improperly set and understood task by elaborating the instructions and stressing the necessity to express the ideas in an organised way, and if needed, to explain and justify the connection displayed in the drawings. There are also other possible categories; for example, categories related to entropy of ideal gases or to the temperature of radiation and its operationalisation. However, because these categories did not appear in the responses, they are not present in our categorisation.

Table I. Frequencies of categories of concepts in students' initial network representations (Initial) and final NRs (Final); the number of students is in parentheses

Category of concepts	Initial ($n=24$)	Final ($n=14$)
Ambiguous use of energy (heat, radiation)	46% (11)	–
Level of qualities, T_{Sensory}		
Sensory experience of warmness	42% (10)	14% (2)
Temperature equilibrium	4% (1)	7% (1)
Changes in state (freezing, melting, evaporating)	29% (7)	7% (1)
Thermal expansion (phenomena)	67% (16)	93% (13)
Level of quantities and laws		
Thermal expansion of liquids or solids, T_{Liquid}	42% (10)	100% (14)
Empirical gas laws, T_{Gas}	17% (4)	100% (14)
Connection to the mechanical quantities	38% (9)	100% (14)
Ambiguous use of volume	29% (7)	–
Ambiguous use of pressure	58% (14)	–
Ambiguous use of temperature scale	33% (8)	7% (1)
Level of structured theory		
Ideal gas law, T_{Gas}	33% (8)	100% (14)
Macro theory, T_{Absolute}	13% (3)	86% (12)
Micro theory, T_{Stat}	4% (1)	93% (13)
Ambiguous use of entropy	58% (14)	7% (1)

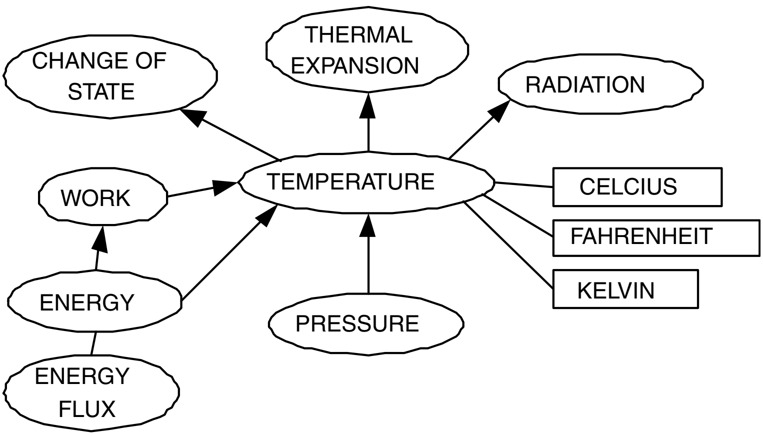

Figure 1. Typical example of a student's NR in the Centralised category. The NR is redrawn from one student's NR; the concepts have been translated from Finnish; the arrangement and shapes of boxes and arrows follows the student's original NR.

For example, Figure 1 shows one NR drawn by a student (redrawn here, but retaining all essential features and the original layout). On the basis of interpreting the expressions (names and words) and their connections in

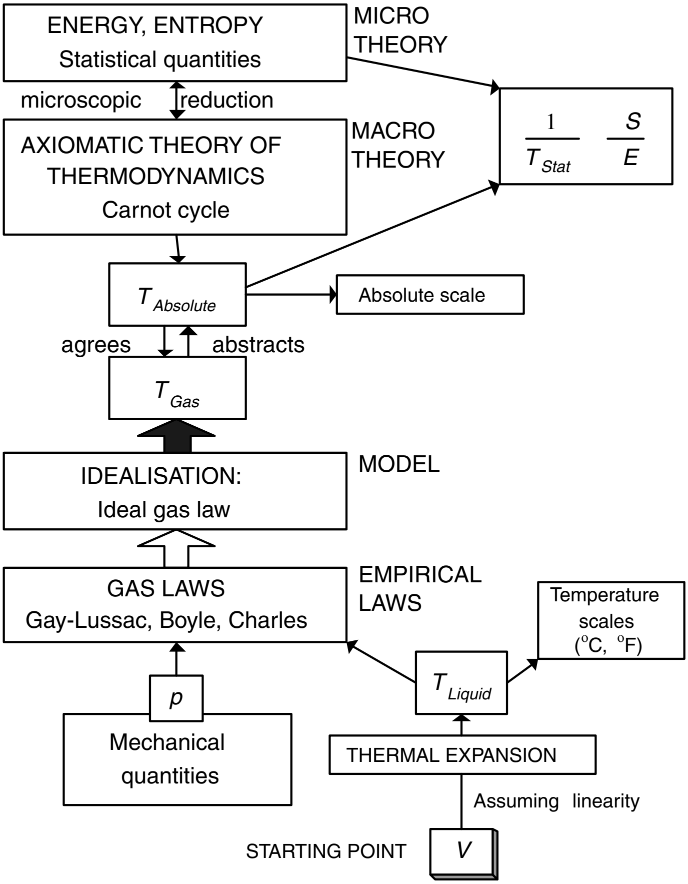

Figure 2. Typical example of a student's NR in the Hierarchical category. The NR is redrawn from one student's NR; the concepts have been translated from Finnish; the arrangement and shapes of boxes and arrows follows the student's original NR.

the NR, the following categories were recognised: *Ambiguous use of energy (heat, radiation), Changes in state (freezing, melting, and evaporating), Thermal expansion (phenomena), Ambiguous use of pressure, Ambiguous use of temperature scale*. Similarly, Figure 2 shows another NR drawn by an another student, and from it the following categories were recognised: *Thermal expansion (phenomena), Thermal expansion of liquids or solids, Empirical gas laws, Connection to the mechanical quantities, Ideal gas law, Macro theory and Micro Theory*. It should be noted that data was acquired in accordance with qualitative, interpretative methodology, and therefore was not intended for such quantitative analysis where a number of nodes and links etc. are calculated and tabulated as is sometimes done in the analysis of concept maps. In our case, what matters are the frequencies of each category picked out by the interpretative analysis. These frequencies are given in Table I.

The comparison of the frequencies of sub-categories in the initial and final NRs in Table I show that: (1) The ambiguous role of quantities (e.g. picked up from mechanics) disappeared, and instead (2) quantities displayed in NRs found a specific, restricted and physically motivated role in the network. The most important result is that now (3) the evolving meaning of temperature through all three stages is well displayed in NRs. This is seen in all cases although not all groups managed to represent the hierarchy of quantities properly.

Next we analysed the overall structure of the NRs, because the structure was expected to correlate with the content. The initial and final NRs were categorised based on how temperature was presented in the representations (the structure of the NR) and connected to other quantities or concepts. Four different categories of structure were formed: *fragmented*, *centralised*, *mixed* and *hierarchical*. There are differences within these categories; e.g., in the number of concepts in the representations and the complexity of representations, but these categories represent well the basic differences in structures between the NRs. The categories of structure are presented in Table II.

Fragmented. The NRs in this category had no structure. Temperature was connected to quantities such as pressure and volume, but there was no directionality in the links between the quantities showing how to proceed, or the connection between the quantities was not physical. In most cases, the NRs also included unnecessary concepts. On the basis of these NRs, it was difficult to see how the quantity temperature develops and what the essential steps in constructing the temperature quantitatively are. Clearly, the omission of quantitative experiments is one reason for the fragmented nature of the NRs. This is simply due to the fact that without the quantifying experiments there is no physically meaningful process to connect the elements in the NRs.

Centralised. The NRs had a simple structure with the temperature at the centre of the representation, as can be seen in Figure 1, where a typical

Table II. Frequencies of categories of the NRs' structure in students' initial NRs (Initial) and final NRs (Final); the number of students is in parentheses

Category of structure	Initial ($n = 24$)	Final ($n = 14$)
Fragmented	33% (8)	–
Centralised	54% (13)	–
Mixed	13% (3)	14% (2)
Hierarchical	–	86% (12)

example of this category is shown. Other concepts, such as volume, pressure and entropy were connected to temperature, but in a simple way without expressing how the connection between concepts might be established. Also the directionality of links between the concepts was, in most cases, rather obscure and often not physical. In the student's NR in Figure 1, for example, an arrow points from the temperature to the thermal expansion, which is in a loose way a reasonable connection; nevertheless, from the point of view of the task, the direction should be the other way around: thermal expansion gives us the first method of measuring temperature. As well, in this class of NRs, the lack of a well-defined role for the quantitative experiments can be noted.

Mixed. In this category, the structure of the NR was similar to the structure of NRs in the fragmented or centralised categories, but now there are hints of a hierarchy, and more attention is paid to the quantitative experiments and measurements. The phases of constructing the temperature quantitatively can be now recognised, which is an aspect missing from the two previous categories.

Hierarchical. In the hierarchical representations there is a clear organisation of structure, as seen in Figure 2 where a typical example of this hierarchical category is presented. In these representations the quantitative experiments and their role in the development and generalisation of temperature is clearly displayed. In these kinds of NRs the different 'hierarchical' steps are also recognised and represented clearly. Moreover, there are significantly more model-, law- and theory-type attributes than in the other categories. In the hierarchical representations there are several specified temperature concepts equipped with a sub-index that explain the position of each temperature concept in the network. In each case, where the operationalisation of the quantity is well represented, the role of the quantitative experiments forming the basis for this operationalisation is also well represented.

Finally, within these four categories of structure, the concepts included in the NRs were classified into qualitative (e.g. sensory experience of warmness, observations, and different phenomena) and quantitative (e.g. quantities, laws, models, theories, quantitative experiments) ones. The frequencies of qualitative and quantitative concepts within the categories of structure are given in Table III. For example, the NR in Figure 1 has both qualitative (e.g. the change of state, radiation) and quantitative (e.g. pressure, work) concepts, while the NR in Figure 2 has only quantitative concepts. The tendency to include qualitative concepts in NRs is strong in the first three categories (fragmented, centralised and mixed) and decreases in the last category, the hierarchical. It should be noted that although the qualitative concepts are often correct, the form in which they are represented makes them unnecessary from the viewpoint of quantitative

Table III. Frequencies of qualitative and quantitative concepts in students' initial and final NRs within the categories of structure of the NR

Category of structure	Initial		Category of structure	Final	
	Quantitative	Qualitative		Qualitative	Quantitative
Fragmented (8)	63% (5)	88% (7)	Fragmented (–)	–	–
Centralised (13)	77% (10)	100% (13)	Centralised (–)	–	–
Mixed (3)	100% (3)	100% (3)	Mixed (2)	100% (2)	100% (2)
Hierarchical (–)	–	–	Hierarchical (12)	17% (2)	100% (12)
All (24)	75% (18)	96% (23)	All (14)	29% (4)	100% (14)

development. For example, noting the connection of the sensory experience of warmness – which by itself is, of course, correct – does not yet give meaningful ways to measure temperature quantitatively.

To conclude, when examining Table II it is noteworthy that most of the initial NRs belong to the first two categories of structure, the fragmented (33%) and the centralised (54%), while most of the final NRs belong to the hierarchical category. In the initial NRs only 13% belonged to the mixed category and there were no representations in the hierarchical category. Most of the initial NRs included both qualitative and quantitative concepts. The common feature of these initial NRs in the fragmented and centralised categories is that temperature is just denoted in them, and different concepts are related to or connected with it somehow, so there is neither any conceivable physical meaning nor any physically meaningful operation to make the interconnections between concepts. Owing to the fact that there is no such structure, it is not possible to conceive temperature as a measurable and physically well-motivated quantity. In these cases, the connections between concepts are established on the basis of the most obvious appearances (e.g. warmness, expansion) or model-like attributes (e.g. change in volume, specific heat), which, however, are left as rather ambiguous.

Of the final NRs only two were categorised as mixed (14%) and the rest were categorised as hierarchical (86%) (See Table II). The final NRs are very different from the initial ones, as the examples in Figures 1 and 2 show. In Figure 2 there is an evident level of increasing abstraction (upwards in the case shown), which clearly indicates that the idea of evolving quantity is understood. The tendency to relate temperature to the models, laws and theory, i.e., to quantitative concepts, increases (100%), while the tendency to relate the concepts directly and ambiguously to the phenomena decreases (29%) as can be seen from Table III. The simple structure that was observed in the initial NRs disappears. In addition, the connections now include or make it possible to design the quantifying experiments and operationalise the quantity. This aspect was missing from the initial NRs in the fragmented and centralised categories.

4.2. INTERVIEWS

The lecturer of the course selected five students for interviews, with the aim of obtaining a representative sample of the students. Three of the students had physics as their major subject and two had it as a minor. Three were female, and two were male. The interviewed students represented different 'performance levels' of students based on their previous performance during the course. One student had constructed the final NR individually and the rest of the interviewed students had constructed their final NR in groups of two or three students. Two of the interviewed students were members of the same group. The interviews were semi-structured and the interview plan was based on the research questions. A preliminary plan was first tested by interviewing one student, and based on experiences acquired in this preliminary interview the interview plan was further modified. The interviews lasted approximately 45 min each. All students were asked basically the same questions, but the interviewer also used spontaneous questions not included in the plan in order to clarify or further probe into the students' responses. Students were encouraged to think aloud about the themes of the interview.

In the interview, students were asked to explain first their initial and then their final NRs, and at the end of the interview they had to compare the representations and think about how the representations reflected their thoughts. The students also had to evaluate their learning process during the teaching sequence. The interviews were videotaped and the pertinent verbal and non-verbal events on the videos were transcribed, focusing on writing down verbatim the natural discourse between the student and the interviewer. Two researchers read the transcripts of all five interviews several times and the categories for recurrent ideas were established. Both researchers first established the categories independently and then compared the categories and discussed them until a consensus was reached. Then the responses were classified according to these categories.

On the basis of the interviews, we evaluated our interpretations about the students' NRs. In the interviews, the students explained their NRs in their own words, and, when needed, specified questions were asked. When explaining their NRs, the students were free to say if they were unable to present their ideas satisfactorily using the NRs or if the NR differed from their views. They could also complete their thoughts about the topic represented in the NR verbally. The interviews confirmed that the initial NRs described the students' thinking at that time and that the NRs' interpretation was consistent with the students' ideas in these five cases. Also in the case of the final NRs, the interviews confirmed that our interpretation of them was quite correct. It was common to all the interviewed students that they could explain their NRs easily with respect to those parts which started from thermal expansion and ended at the level of the ideal gas law. However, temperature at the level of theory (excluding the ideal gas law) was difficult for them to explain. In many cases, the students failed in explaining their NRs and the relationships between the concepts at this level especially in the case of micro theory. This indicates that the interviewed students had understood the existence of, and the need for this level of theory in the construction of temperature as a quantity, but that they had not understood the concepts involved and the connections between the concepts in depth. The interviewed students raised the point that they would have needed more material from which to summon up the related knowledge before the exercises. Clearly, this brings forward the limitations of the NRs, because, after all, they are tools for organising fragmented knowledge, and for that, the basic requirement is the existence of those pieces of knowledge to be organised. We understand this point, hence in the future, the physics contents as well must be discussed in a more thorough way and be integrated into the use of NRs.

5. What was Gained by Using the Reconstruction?

Before the teaching sequence described above, we had discussed with the students our reconstruction at the general level and its uses in other contexts (e.g. mechanics, and heat and energy). Here we have applied it in a narrower context: in the case of one quantity. When we compare the initial and final NRs, the results show that the students needed more support for the utilisation of the reconstruction than the broad and general discussions in the lectures provided. Although the reconstruction is helpful in organising the existing knowledge, it alone does not improve the understanding of clearly inadequately acquired ideas. This is seen, for example, in cases of the role of entropy and internal energy in defining the temperature concept and the relation of temperature as a macro variable to the microscopic explanations. This simply means that attention needs to be paid to the physics contents also, and that learning tools such as graphical representations are not enough. This clearly shows the natural limitations of such tools. For future development the crucial question is whether or not the students can generalise the ideas discussed in the teaching sequence and apply them in other cases as well. On the basis of preliminary results for more complex situations, like induction law, the answer seems to be positive, but further research is needed.

An interesting notion contained in the results is that learning to use the reconstruction and recognise the detailed role of quantitative experiments seems to take place only when the general ideas are contextualised and applied in detail. One difficulty students had was that they were not familiar with thinking about quantities as evolving or process-like structures. Although they had previously learned how the construction of new quantities is always based on already existing quantities – through quantifying experiments – they still did not realise that this process can be continued and that it extends and generalises the meaning of quantity.

Most of the final NRs were hierarchical (Table II) and the different levels of reconstruction are also recognisable in them, as in Figure 2. One purpose of the interviews was to find out how the students evaluated the usefulness of the reconstruction in producing the NRs and what, in their own opinion, were the advantages of it for their learning. It can be inferred from the students' responses, that the students really learned to distinguish different abstraction levels in their final representations and that they recognised the different roles of measurements with respect to different concepts and quantities. For example:

> 'This is clearly the region of mechanics... more empirical-level things... empirical gas laws are here; it is perhaps the closest connection to experimentality... [W]hen one comes to the ideal gas law, it begins to be on a higher [level of hierarchy]... and the absolute temperature, it is a rather high-level concept, and finally we have these axiomatic theories of thermodynamics... they are [at] the highest [level of the hierarchy].'

The reconstruction also helped students to think in an organised and logical way by giving them a structured means to approach the subject content. Students stated that, for example:

> 'It is a kind of scaffolding. I automatically start to think what concepts are interrelated... and in which order I represent things.'

> 'I have also in my mind right now such a map in which I know where and how to connect these [quantities] to each other.'

> 'Learning physics is like climbing upwards step by step, and every step is needed. This [making NRs] is useful because it organises thinking and one easily recognises in what step something is missing. It is possible to build a whole structure of what one has learned.'

There are also responses which clearly confirm that the students have understood the underlying principles and truly formed a new and better understanding of the subject matter, for example:

> 'You can interrelate different quantities to each other and understand where you need them; earlier they have perhaps been just a mishmash in your mind.'

> 'It has been nice to build the concept hierarchy... obscure connections become more definite and probably some new connections emerge.'

> 'The topics [physics contents] discussed in this course are familiar from previous [introductory] courses, but now it becomes a bit clearer how different things [concepts] are related. In previous courses a thing [concept] has been named... and then we have done some calculations.'

The interviews thus support the conclusion that the reconstruction has helped students to learn the topic under discussion, and in addition, has helped them to achieve a more general level of thinking about physics. Moreover, the students were clearly aware themselves of the positive effect and utility of the reconstruction. Especially the last quotation brought up the fact that the topics discussed already have been taught during previous courses, but have remained a disconnected set of concepts. The reconstruction helped students to relate and organise knowledge in a certain way, which they felt was useful and effective in comparison with other types of learning methods they had experienced (lectures, writing, simple group working). The interviews also showed that the collaborative part of constructing the NRs, with its reaching of a consensus, was important for their learning.

6. Conclusions

We have introduced a didactical reconstruction for understanding the construction of physical quantities by using a network point of view, where the quantities are part of networks and the quantifying experiments build up the networks. The reconstruction used here is based on the history and philosophy of science and uses the history and philosophy of science as a resource. As a practical example, we discussed how student teachers used the reconstruction in constructing the quantity temperature during an instructional unit designed for them. We evaluated the utility and advantages of the reconstruction by collecting empirical data (network representations, interviews) from the students' learning process.

From the research results given here, the following conclusions can be drawn (with numbers referring to the research questions):

1. Students learn to understand the role of quantifying experiments in the process of constructing quantities. This is seen from the development of the NRs. The final NRs had experiment-based connections between concepts that were missing from most of the initial NRs.
2. Students learn to understand that a quantity is an evolving process. For example, temperature just exists in most of the initial NRs, but the final NRs have several temperature quantities, which show the evolving meaning of temperature.

3. Students learn to integrate quantities as part of a network. This is a sign that students learn to utilise the reconstruction in supporting their learning, or rather, what is to be learned.

The results given in this paper show that by using a suitable didactical reconstruction, teacher students can be supported in their process of understanding the structure of physics knowledge. In this process, making NRs, where the reconstruction helps students to organise their knowledge, seems to be advantageous. However, equally – taking into account that our student group consisted of third year students who already had taken a course in thermodynamics and should have been familiar with the ideas discussed here but apparently were not – we can think of the reconstruction also as an aid or a tool for grasping what *should have been* learned. On the other hand, the approach we have introduced here also has natural limitations. Paying attention to the role of quantifying experiments helps students to think about physics in an organised way, partly because the experiments in the framework of the reconstruction make it possible for students to understand the different steps and processes involved in the construction of physics knowledge. In summary, the students learn to describe *how they know what they know*, which is a clear indication of achieving a better conceptual understanding.

Acknowledgements

We would like to thank Prof. Jari Lavonen for helping us in gathering and analysing the data, and also for many inspiring and helpful discussions concerning our work.

References

Bagno, E., Eylon, B.-S. & Ganiel, U.: 2000, 'From Fragmented Knowledge to a Knowledge Structure: Linking the Domains of Mechanics and Electromagnetism', *American Journal of Physics* **68**, S16–S26.

Carlton, K.: 2000, 'Teaching about Heat and Temperature', *Physics Education* **35**, 101–105.

Chang, H.: 2004, *Inventing Temperature: Measurement and Scientific Progress*, Oxford University Press, Oxford.

Cotignola, M.I., Bordogna, C., Punte, G. & Cappannini, O.M.: 2002, 'Difficulties in Learning Thermodynamic Concepts: Are They Linked to the Historical Development of this Field?', *Science & Education* **11**, 279–291.

diSessa, A. & Sherin, B.L.: 1998, 'What Changes in Conceptual Change?', *International Journal of Science Education* **20**(10): 1155–1191.

Erlichson, H.: 2001, 'Kelvin and the Absolute Temperature Scale', *European Journal of Physics* **22**, 325–328.

Giere, R.N.: 1999, *Science Without Laws*, The University of Chicago Press, Chicago.
Izquierdo-Aymerich, M., Adúriz-Bravo, A.: 2003, 'Epistemological Foundations of School Science', *Science & Education* **12**, 27–43.
Koponen, I.T. & Mäntylä, T.: 2006, 'Generative Role of Experiments in Physics and in Teaching Physics: A Suggestion for Epistemological Reconstruction', *Science & Education* **15**, 31–54.
Koponen, I.T, Mäntylä, T. & Lavonen, J.: 2004, 'The Role of Physics Departments in Developing Student Teachers' Expertise in Teaching Physics', *European Journal of Physics* **25**, 645–653.
Meltzer, D.: 2004, 'Investigation of Students' Reasoning Regarding Heat, Work, and the First Law of Thermodynamics in an Introductory Calculus-based General Physics Course', *American Journal of Physics* **73**, 1432–1446.
Novak, J.D. & Gowin, D.B.: 1984, *Learning How to Learn*, Cambridge University Press, Cambridge.
Reif, F.: 1987, 'Instructional Design, Cognition and Technology: Applications to the Teaching of Scientific Concepts', *Journal of Research in Science Teaching* **24**, 309–324.
Reif, F.: 1995, 'Millikan Lecture 1994: Understanding and Teaching Important Scientific Thought Processes', *American Journal of Physics* **63**, 17–32.
Taber, K.: 2000, 'Finding the Optimum Level of Simplification: The Case of Teaching about Heat and Temperature', *Physics Education* **35**, 320–325.
Thagard, P.: 1992, *Conceptual Revolutions*, Princeton University Press, Princeton.
Trowbridge, J.E. & Wandersee, J.H.: 1998, Theory-Driven Graphical Organizers, in J.J. Mintzes, J.H. Wandersee and J.D. Novak (eds.), *Teaching Science for Understanding: A Human Constructivistic View*, Academic Press, San Diego.
Heuvelen, A.Van: 1991, 'Learning to Think Like a Physicist: A Review of Research-based Instructional Strategies', *American Journal of Physics* **59**, 891–897.