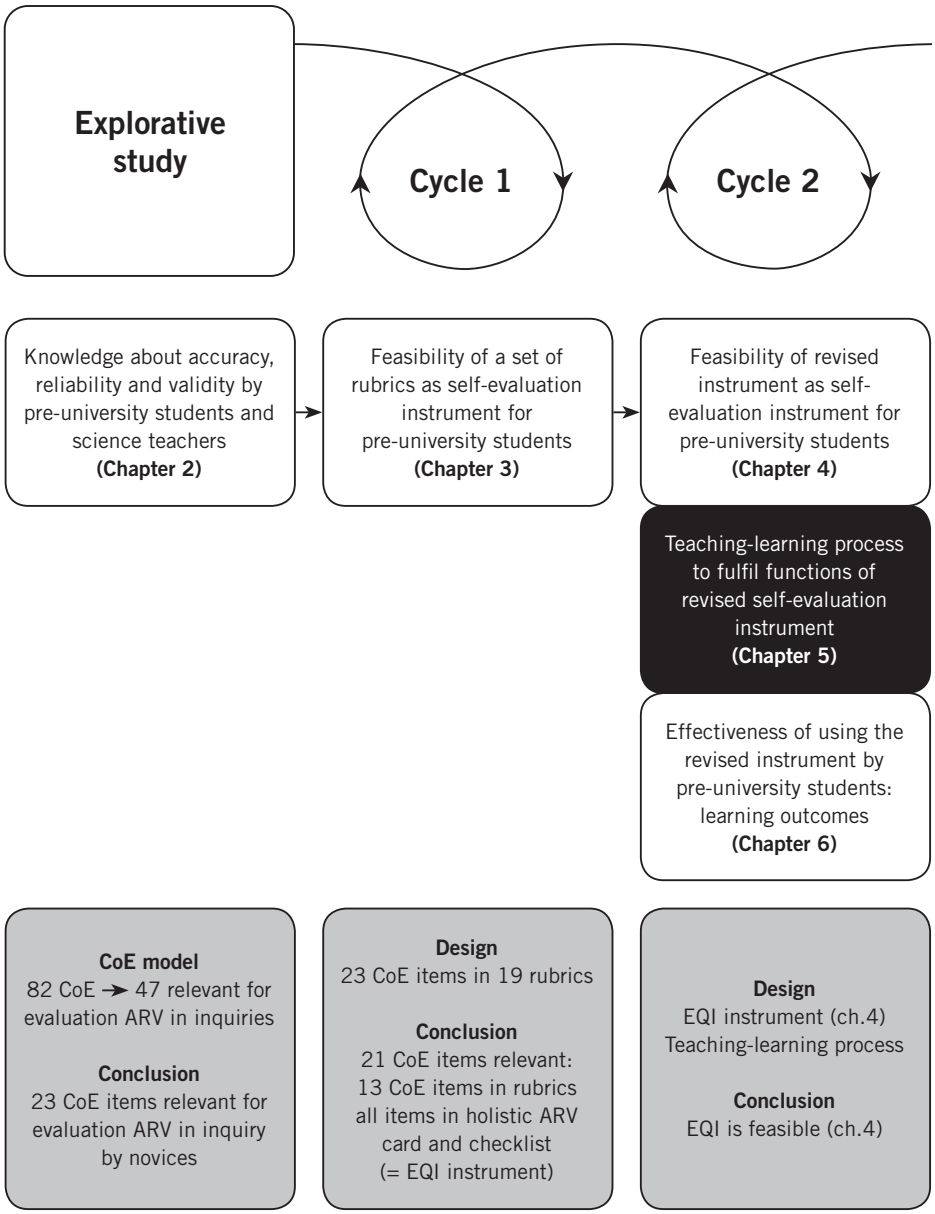

# TEACHING-LEARNING PROCESS TO FULFIL FUNCTIONS OF REVISED SELF-EVALUATION INSTRUMENT

# CHAPTER 5

This study focused on the design of a teaching-learning process for different inquiries in which pre-university science students learned to evaluate the accuracy, reliability and validity (ARV) of inquiries by using a feasible Evaluation of Quality of Inquiries (EQI) instrument (see Chapter 4). The design of the teaching-learning process was based on eleven hypothetical design characteristics that were identified from the literature. The EQI instrument had three intended formative feedback functions in the teaching-learning process: (1) self-evaluation of ARV by students, (2) support of self-evaluation by peers and teachers and (3) facilitation of transfer from one inquiry to another. Whether the designed teaching-learning process fulfilled the three intended functions was tested in class with 27 pre-university science students. From the analyses of audio and video recordings, observers' field reports, teacher's reflection reports, students' filled-out workbooks, the filled-out EQI instruments by students and teachers and students' responses in questionnaires and interviews, it was concluded that the teaching-learning process supported the fulfilment of functions 1 and 2. Function 3, regarding the facilitation of transfer, was fulfilled to some extent: students recognised analogies between the different inquiries and between scientific research and their own inquiries, but did not sufficiently apply knowledge on differences in evaluating ARV in different inquiries. The series of three successive inquiry units was probably too short to facilitate the transfer of all content of the EQI instrument from

This chapter is based on the article: S. A. W. van der Jagt, L. van Rens, H. H. Schalk, A. Pilot, & J. J. Beishuizen. Teaching-learning process for learning to evaluate the quality of inquiries at pre-university level by using a self-evaluation instrument. In preparation

one inquiry to another. Implications for the designed teaching-learning process are discussed, followed by revision of the eleven hypothetical design characteristics to fifteen design characteristics for a (more) appropriate teaching-learning process. Although some revisions are necessary, the students' performances in the teaching-learning process could be used to determine the effectiveness of the EQI instrument (see Chapter 6).

## 5.1 INTRODUCTION

The revised formal science curriculum for pre-university chemistry, biology and physics in the Netherlands (College voor Examens, 2013) requires that pre-university science students learn how accuracy, reliability and validity (ARV) in inquiries can be evaluated. Evaluation of ARV is a new aspect in the formal Dutch science curriculum. As it is a new aspect in the revised curricula for chemistry, biology and physics, all these curricula contain more or less the same examination levels regarding ARV, which means that evaluating the ARV in inquiries can be coherently taught in the various school science subjects. To meet these new examination levels, a feasible, effective and coherent teaching-learning process and accompanying educational materials have had to be developed, by which students can learn to evaluate the ARV of inquiries in different school science subjects.

In two preceding studies, a self-evaluation instrument was developed by which students could be supported in evaluating the ARV of inquiries. The content of this instrument was based on the concepts of evidence (CoE) model. This model describes 82 CoE that are important in building up evidence in scientific inquiries (Gott, Duggan, Roberts, & Hussain, n.d.). In the first research cycle on the design of the self-evaluation instrument (see Chapter 3), 23 CoE were assessed as relevant for pre-university science students in formative evaluations of the ARV of inquiries. Gott et al. (ibid.) described the CoE in general terms. As shown in Chapters 2 and 3, the CoE were translated to the level of Dutch pre-university science students and indicated as '(CoE) items' in the instrument. From the first research cycle (see Chapter 3), it was concluded that 21 items were feasible and 13 of them were described in twelve rubrics[5]. Moreover, in the design of the EQI instrument (see Chapter 4), all 21 items should be brought together in a checklist with which students could check whether they had addressed all relevant aspects when they conducted an inquiry. All 21 items should also be set out on an ARV card, a holistic overview tool, that gives the students an overview of how the 21 items are related to ARV in an inquiry. The complete instrument with rubrics, checklist and ARV card, named the Evaluation of Quality of Inquiries (EQI) instrument, can be found in Appendix C.

5 From the twelve designed rubrics, eventually the rubric 'Theoretical framework' was not used, because it was decided to provide the students with a theoretical framework instead of letting them write this framework themselves (as was done in the first research cycle).

The hierarchical descriptions of the various levels in each of the twelve rubrics were based on the Structure of Observed Learning Outcomes (SOLO) taxonomy (Biggs & Tang, 2007). This taxonomy has been proven to be supportive for university students in evaluating their actions at various points during the performance of an inquiry task (e.g. Hodges & Harvey, 2003; Levins & Pegg, 1993; Minogue & Jones, 2009). The SOLO taxonomy distinguishes five levels of increasing complexity: prestructural, unistructural, multistructural, relational and extended abstract. A prestructural level occurs when students use everyday language, for example, to describe accuracy: 'we did everything as precisely as possible'. At the unistructural level, students are able to describe one aspect, mostly by imitation of the language in the learning materials or the teacher's language. At the multistructural level, students describe relevant aspects without considering inconsistencies or possible relations between these aspects. At the relational level, students do consider inconsistencies and possible relations. When students are able to indicate how inconsistencies and related aspects fit in the research domain concerned, they perform at the extended abstract level. If the SOLO taxonomy is correctly worked out in an instrument, then it is expected that students can only achieve a certain level of performance when they have mastered all the previous levels of performance in the SOLO taxonomy (Biggs & Tang, 2007).

From testing the self-evaluation instrument in class, it was concluded that it is a feasible instrument to support pre-university science students in performing an evaluation of the ARV in inquiries of general science, biology and physics (see Chapter 4). As part of the second research study, this study focused on the design characteristics of a teaching-learning process of different inquiries in which pre-university students use the EQI instrument to evaluate the ARV in an inquiry.

## 5.2 THEORETICAL FRAMEWORK

The EQI instrument was designed to support students in learning to evaluate ARV from one inquiry to another. With regard to the intended practical use of the EQI instrument in class to learn to transfer the evaluation of ARV of inquiries, previous research showed that students could use the instrument for self-evaluation, and peers and teachers could support this ARV evaluation. Consequently, in the teaching-learning process, effective formative feedback on the ARV of an inquiry could be given when the same EQI instrument is used: (i) in the self-evaluation of students' inquiries, (ii) by peers and teachers to give feedback on the students' inquiries, and (iii) in inquiries in different school science subjects. This is related to the three intended formative feedback functions of this instrument in the teaching-learning process:

1) Self-evaluation of ARV by students
2) Support of self-evaluation by peers and teacher
3) Facilitating transfer of self-evaluation of ARV to different inquiries.

These three functions were supported by the principles of good formative feedback proposed by Nicol and Macfarlane (2006, p. 205) who conducted a review on formative assessment. They defined 'good feedback practice' *as anything that might strengthen the students' capacity to self-regulate their own performance*. They considered self-evaluation as an activity that contributes to self-regulation of students' learning processes. Self-regulation is a complex process for students to learn, due to the combination of cognitive, motivational and contextual aspects (Pintrich, 2000). By formulating their seven principles of good formative feedback, Nicol and Macfarlane (2006) gained insight into the various aspects of learning to self-regulate learning processes. In their words (p. 205), 'good feedback practice':

1) *helps to clarify what good performance is (goals, criteria, expected standards);*
2) *facilitates the development of self-assessment (reflection) in learning;*
3) *delivers high quality information to students about their learning;*
4) *encourages teacher and peer dialogue around learning;*
5) *encourages positive motivational beliefs and self-esteem;*
6) *provides opportunities to close the gap between current and next performances;*
7) *provides information to teachers that can be used to help shape teaching*.

The fifth principle, on the motivational beliefs of students and their self-esteem in using the EQI instrument, was addressed in the study about students' learning outcomes (see Chapter 6). The other six principles were part of the study described in this chapter and were expected to support the three functions of the EQI instrument in a teaching-learning process in which it was intended to teach students to self-evaluate their inquiry performance.

When students receive an instrument, such as the EQI instrument, to reflect on their actions in order to improve them, it is important that there is consistency between the intended functions of the instrument and the teaching-learning process in which the instrument is used. In the unfortunate case of a mismatch between instrument and teaching-learning situation, the learning materials will lose their function; the teaching-learning process will be ineffective and will therefore not lead to appropriate student learning outcomes (Ledford & Sleeman, 2000). Consequently, each of the three intended formative feedback functions of the EQI instrument must be appropriately elaborated in the teaching-learning process. From literature on teaching-learning processes at pre-university level (see the next sections for the literature references), hypothetical design characteristics (hDC) for the elaboration of each function in the teaching-learning process were identified. From these hDCs, indicators were derived to investigate the extent to which each of the three functions was fulfilled in the teaching-learning process. Table 5.1 presents an overview of the hDCs of the teaching-learning process with matching indicators for the fulfilment of the three intended functions of the EQI instrument in the teaching-learning process. The hDCs and their indicators are described in more detail below.

Table 5.1
*Hypothetical design characteristics and indicators for fulfilment of three intended functions of the EQI instrument*

| DESIGN CHARACTERISTIC | INDICATOR |
|---|---|
| **Function 1:** Self-evaluation of ARV | |
| **hDC1:** Learning tasks to learn concepts of EQI instrument | **1a** Students understand concepts in EQI instrument |
| **hDC2:** Explanation about relevance of self-evaluations and how to conduct these evaluations | **1b** Students use EQI instrument at intended points |
| **hDC3:** Supportive in changing ARV performance to higher levels | **1c** Students change their level of evaluating the ARV of inquiries |
| **hDC4:** Central activity of the students | **1d** Self-evaluation before peer and teacher feedback |
| **Function 2:** Support by peers and teacher | |
| **hDC5:** Peer feedback before teacher feedback | **2a** Peers give feedback before the teacher does |
| **hDC6:** First exchange of information between peers, then discussion | **2b** Peers suggest changes of ARV |
| **hDC7:** Feedback linked to criteria which students can head to | **2c** Teacher gives written feedback with EQI instrument |
| **hDC8:** Discussion between teacher and students about improving inquiry | **2d** Teacher gives oral feedback followed by discussion on changes |
| **Function 3:** Facilitating transfer | |
| **hDC9:** Analogies in different inquiries' learning materials | **3a** Students recognise analogies in different inquiries |
| **hDC10:** Scientific research examples to explain abstract terms on ARV | **3b** Students recognise analogies with scientific research |
| **hDC11:** Explanation of differences in evaluating ARV in specific subjects | **3c** Students apply knowledge on differences in evaluating ARV in different subjects |

### 5.2.1 Function 1: Self-evaluation

Fulfilling function 1, self-evaluation of ARV by students, first required learning tasks in the teaching-learning process in which students had sufficient opportunity to learn to understand the meaning of the concepts (CoE items) in the EQI instrument (hDC1, see Table 5.1). Only with this understanding were they able to perform their self-evaluations of the ARV of an inquiry adequately (Sadler, 1989). Secondly, the

students, as novices in evaluating the ARV of inquiries, needed to be provided with explanations as to the relevance of self-evaluating the ARV in various stages of an inquiry, and how such a self-evaluation can be conducted (hDC2) (Bransford, 2000). Thirdly, Andrade and Valtcheva (2009) concluded that the development of students' self-reflective practices can be increased when students are given guidance on how they can move from their present performance to higher levels of performance, for example, by using the descriptions from the rubrics in the EQI instrument (hDC3). Lastly, Nicol and Macfarlane (2006) concluded that students should have a central and active role during all evaluations in the teaching-learning process. They first have to perform an evaluation themselves, or at least have read the evaluation form, before they can apply it to the research of others. Consequently, in the teaching-learning process of this study, students constantly performed self-evaluation with the same instrument before a peer or teacher evaluated the ARV of the same inquiry (hDC4).

Hence, to see whether function 1 was sufficiently fulfilled, the enacted teaching-learning process was analysed on four indicators:

a) Students understand the meaning of the concepts (CoE items) in the EQI instrument.
b) Students evaluate the ARV of an inquiry with the EQI instrument at the intended points in the teaching-learning process.
c) Students change the ARV in their inquiry by using the EQI instrument.
d) Students perform a self-evaluation before the peer and teacher evaluations.

### 5.2.2 Function 2: Support of self-evaluation by peers and teacher

Function 2 of the EQI instrument relates to the support of self-evaluation by peers and teachers. The order 'first peers then teachers' was chosen because of Nicol and Boyle's (2003) conclusion that support from peers and peer discussions should occur in the teaching-learning process before the students get feedback from the teacher. In the starting phase of a student's learning process, peers, more often than the teacher, use language that matches the prior knowledge of other students, through which feedback can be effectively internalised by the receiver (hDC5).

Peer feedback requires an active participation of students in the teaching-learning process upon which peers can base their feedback in the role of critical friends and upon which discussion between peers can be stimulated (Lynch, Mannix McNamara, & Seery, 2012). Furthermore, according to Gagné, Yekovich and Yekovich (1993), when students are asked explicitly to discuss issues with peers, transfer can be increased, which can also help to fulfil function 3 on facilitating transfer. Nicol and Macfarlane (2006) concluded that the efficiency of peer feedback increases when exchange of information between groups precedes the discussion (hDC6). They defined discussion as a situation in which a group of students offer arguments for changes that should be made by

another group of students. The latter group can reject the arguments or ask for more specification. To optimise the peer discussion and to improve mutual comprehension, the students should all make use of the same terminology (Gibbs & Simpson, 2004).

However, novices insufficiently recognise errors and misconceptions in their own as well as their peers' work, which can lead to under- or overestimation of peers' performances in inquiries. Therefore, the support of a teacher, as an expert, remains a necessity in the teaching-learning process for novices to learn adequately about the application of evaluation criteria (Bransford, 2000; Sadler & Good, 2006). Schalk, Van der Schee and Boersma (2013) showed that students who received feedback from the teacher in an evaluative task regarding procedural understanding had better learning outcomes than students who did not receive teacher feedback. Schalk et al. (2013) suggested that the teacher needs to take the initiative in providing feedback and should not wait until the students ask for feedback. Bransford (2000) stated that the teacher should first give the students feedback, after which students can discuss the feedback with their teacher. It is not yet clear whether and when oral or written feedback from the teacher is most effective in teaching-learning processes, but in any case, both need to be linked to aims or criteria so that students can hear or see 'where they can head to' (Nicol & Macfarlane, 2006). In feedback, it is also necessary that the terminology used during discussions is clearly and similarly understood by both the students and the teacher (Gibbs & Simpson, 2004). Talking about changes helps students to understand the descriptions in rubric levels that the students have not yet properly reached. Consequently, a better understanding of the descriptions in the EQI instrument could help students to improve their capacity to self-regulate (Nicol & Macfarlane, 2006). The aforementioned guidelines were implemented by first letting teachers fill out the EQI instrument to evaluate the ARV of students' inquiries; then, the students compared the teacher's filled-out EQI instrument with their self-evaluation with the EQI instrument and attempted to understand the differences between both evaluations (hDC7). With support of the filled-out EQI instrument and by using the wordings from the instrument, the teacher can give oral feedback on the students' performances and suggested changes in the ARV of an inquiry (hDC8).

Hence, to see whether function 2 was sufficiently fulfilled, the enacted teaching-learning process was analysed with the help of four indicators:

a) Peers give feedback before the teacher gives feedback.
b) Students discuss the ARV of their inquiries with peers by using terminology from the EQI instrument.
c) The teacher gives written feedback to the students by using the EQI instrument.
d) The teacher gives oral feedback to the students, in which he or she discusses changes in the ARV of an inquiry.

### 5.2.3 Function 3: Facilitating transfer

Fulfilling function 3, facilitating the transfer of self-evaluation of ARV, means that, on the basis of mindful abstraction, students should be supported in learning to use the same instrument or tool in different situations (Salomon & Perkins, 1989). In order to accomplish this, previously acquired experiences in evaluating the ARV of an inquiry with the EQI instrument have to be 'translated' to fit a new inquiry. To support this 'translation', students should recognise the intended analogies (Van Oers, 1998b), for instance, between an inquiry in biology and an inquiry in physics, and analogies should be taught in different inquiries, explaining that evaluation of ARV with the EQI instrument could be performed in a similar way. Gagné (1985) concluded that recognising analogies (1) helps students to generalise knowledge that is used in both inquiries, (2) improves their self-evaluation ability and (3) stimulates knowledge storage in the long-term memory. To retrieve the stored information from the memory at appropriate moments, it is useful to elaborate identifying marks in the learning materials to help students recognise the analogies between two different inquiries. Important identifying marks include a coherent use of language and the same patterns in students' mental actions in different inquiries (hDC9) (Gilbert, Bulte, & Pilot, 2011).

Students should also be supported to recognise analogies in the use of (abstract) concepts (Ledford & Sleeman, 2000), such as the ARV of an inquiry, by tasks in the teaching-learning process in which the students evaluate the ARV of scientific research, followed by evaluating the ARV in the inquiries they perform themselves. Chen, Yanowitz and Daehler (1995) concluded that spontaneous transfer by students is facilitated when they are provided with scientific research examples wherein the abstract terminology is illustrated (external instantiation). When students are more familiar with the abstract terminology, transfer can also be increased by asking the students to produce their own examples (internal instantiation). The study presented in this chapter focused on the elaboration of external instantiation in the teaching-learning process, because the students were novices at using the EQI instrument in inquiries in different school science subjects and were not yet familiar with the terminology used (hDC10).

A third way to support the facilitation of transfer relates to showing differences in evaluating ARV in different inquiries, with the implication that, for example, within the teaching-learning process students should learn to use the concepts in the EQI instrument in a flexible way. A goal of teaching is to help students to understand the kinds of inquiries in which certain concepts apply and those in which they do not (Minstrell & Stimpson, 1996). As Gilbert, Bulte and Pilot (2011) put forward, transfer is more complex for students when the disciplinary domain is changed. Moreover, each domain has developed its own culture (Van Oers, 1998a) in which, for example,

ARV has its own specific meanings. This creates differences regarding how these concepts are applied in different scientific research contexts. To offer an example about differences in evaluating the reliability of an inquiry: when non-living objects are investigated, repetition of measurements is needed, whereas in research with living subjects a representative sample is needed (Mayr, 1996). Therefore, it is important to explain clearly to the students what the similarities and differences are in evaluating the ARV in inquiries in different school science subjects, especially when the disciplinary domain is changed during the teaching-learning process (hDC11).

Hence, to see whether function 3 is sufficiently fulfilled, the enacted teaching-learning process is analysed on three indicators:

a) Students recognise analogies regarding evaluating ARV in the different inquiries.
b) Students recognise analogies with evaluation of the ARV of scientific research studies. They apply this knowledge in evaluating the ARV of their own inquiries.
c) Students apply their knowledge about the differences in evaluating ARV in different inquiries to their own inquiries.

Based on the above described hypothetical design characteristics (see Table 5.1), a design of a teaching-learning process was made, in which it was expected that the three functions of the EQI instrument would be fulfilled when the process is enacted in class.

## 5.3 DESIGN OF TEACHING-LEARNING PROCESS

The designed teaching-learning process contained three successive inquiry units – general science, biology and physics. The students were expected to work in groups during the inquiry units, so that they could discuss the various tasks involved in the units. For each unit, a workbook with tasks for the students and a teacher's guide was written. The teacher's guide contained a description of the intended functions of the EQI instrument and the intended teacher instructions during the enactment of the inquiry units., Self-evaluation of the ARV was intended to take place at similar points in all three inquiry units. The students were provided with the complete EQI instrument. The teacher had to explain the use of the instrument to the students before they used it for the first time. The terminology in the instrument had to be elaborated by the teacher at the start of each unit. The student workbooks contained instructions for the students to use the EQI instrument at the intended points during the teaching-learning process.

### 5.3.1 Inquiry tasks

Each inquiry unit was constructed around an inquiry task in which students planned and performed their own inquiry, and reflected on it, because performing their own inquiries can help students understand the evidence used in inquiries (Gott &

Duggan, 1996; Hodson, 1999). In the first inquiry unit, general science, the subject of the inquiry task was the cooling rate of hot coffee. In the second, the biology unit, the inquiry task focused on different theories of taste on the human tongue. Third, in the physics unit, the inquiry task concentrated on improving the traffic situation on a dangerous crossing by, for example, collecting data about the speed and braking distance of cyclists. Figure 4.1 (p. 85) gives an overview of the successive inquiry topics in the three inquiry units.

For each inquiry unit a workbook with activities for the students was designed. All inquiry units were accompanied by a teacher's manual in which the intended use of the student EQI instrument as well as the guidelines for instruction to enact the unit were described (see http://hdl.handle.net/1871/38422). From a pilot study in class with the students who participated in the first research cycle, it was concluded that the inquiry tasks in the units were experienced as interesting by most of the students. This pilot study is described in Appendix D.

In all inquiry tasks, students first had to design an inquiry plan that was related to a given theoretical framework about the inquiry subject. Instructions in the workbooks helped students to formulate an inquiry plan (an inquiry question, a hypothesis and the planned method). They then performed a real-life pilot experiment to test the practicability of their described inquiry method. Using the results and experiences of the pilot experiment, the students were asked to adjust, if necessary, their inquiry plans. After this, the students used the checklist in the EQI instrument to check whether any further adjustments to the inquiry plan were needed. Next, the students were expected to use the relevant rubrics in the EQI instrument to evaluate the ARV of their inquiry plan and to decide whether the plan needed any further changes. Then, the students had to conduct the inquiries according to their plans, using the checklist in the EQI instrument to find out whether the inquiry needed repetition to strengthen evidence. After this, they had to process the obtained data into conveniently arranged results and then to formulate their conclusions. Finally, the students were expected to evaluate the ARV of their completed inquiries with the relevant rubrics in the EQI instrument and to write down their suggestions for changing the aspects concerning the ARV of their inquiry (hDC2, hDC3, hDC4). They could use the ARV card at any point during the inquiries. After each inquiry unit the students were explicitly asked to note additions and clarifications on their ARV cards.

### 5.3.2 Supporting tasks

To guide and support the students' performances in each inquiry task in the teaching-learning process, the inquiry units contained three types of supporting tasks:

i. Orientation tasks on the evaluation of ARV with the EQI instrument and on the meaning of the CoE items in the EQI instrument;

ii. Tasks on the enactment of the peer feedback with the EQI instrument;
iii. Tasks on the transfer of evaluation of ARV with the EQI instrument from one inquiry to another.

***(i) Orientation on ARV tasks***

The general science unit contained a task that was designed to introduce the ARV card, on which the meanings of all CoE items in the EQI instrument were explained to the students. The purpose of this orientation task was to make the students aware of the items that contributed to evaluating the ARV of an inquiry (hDC1). In addition, all units contained two tasks in which the students compared their evaluations with the teacher's evaluations. The teacher was expected to give written feedback with the same checklist and rubrics on the ARV of all student inquiry plans and to return these evaluations to the students in the next session. The teacher also had to give written feedback on the ARV of the completed inquiries (hDC7). The students were then asked to consider and explain possible differences between their self-evaluations and the teacher's evaluation. If necessary, the teacher clarified the meaning of the CoE items from the EQI instrument. The teacher was expected to give oral feedback along with the filled-out EQI instrument on the students' performances and to discuss changes in the ARV of an inquiry with all student groups in all inquiry units (hDC8).

***(ii) Tasks on peer feedback***

Each inquiry unit contained a task wherein students were expected to give feedback on the ARV of peers' inquiries before they received feedback from the teacher (hDC5) and to discuss the changes in the ARV of an inquiry (hDC6). The general science unit contained a task which featured a peer discussion between two small groups of students on the ARV of their completed inquiries on the cooling of coffee. The peers first had to evaluate the ARV in the inquiry of the other student group by means of the EQI instrument and by asking questions of their peers. Then they were asked to formulate suggestions for changes in the inquiry in order to improve the ARV; these changes were expected to be discussed between both student groups. During the biology unit, all the results from the student's inquiry on human taste were presented on a poster sheet after which each student group was asked to evaluate the ARV of at least three other peer groups' inquiries by using the EQI instrument. This peer evaluation was intended to be followed by formulating suggestions for changes in the ARV of inquiries of other groups and by a discussion on these issues between students. For the traffic situation inquiry in the physics unit, the peer feedback task was intended to take place as a discussion between two small student groups, as in the general science unit. The above described variety in the peer discussion procedures in the three units was chosen to keep students motivated to participate in peer discussions (Ledford & Sleeman, 2000).

***(iii) Tasks on transfer***

The analogies between the different inquiry tasks were expected to be demonstrated by using the same EQI instrument at similar points in all the inquiry units. The inquiry tasks contained the same instructions about the use of the EQI instrument and similar instructions were given to guide the students' mental actions. For example, every time students had to formulate a hypothesis as part of their inquiry plan, the instruction in their workbook was the same: 'Based on the theoretical framework, what relation do you expect to find? Work this relation out in a hypothesis.' The terminology used in the EQI instrument was applied to all inquiry and supporting tasks (hDC9).

Next to this, each inquiry unit contained a task supporting students in the transfer of evaluation of ARV in inquiries in different situations. The general science unit started with a task in which the students were expected to evaluate the ARV of a scientific research study on the use of probiotics in recovering patients with an acute inflammation of the pancreas. During the inquiry task, the students were asked to compare the strong and weak aspects of the ARV in this scientific research study with the ARV in their own inquiry plan on the cooling rate of coffee. In this task, the students were instructed to make use of the EQI instrument in evaluating the ARV of the probiotic research study (hDC10). In the biology unit, the students were expected to perform a task on the possible differences in use of relevant items in the EQI instrument about evaluating the ARV in one inquiry with living test objects as well as one with non-living test objects (hDC11). In the physics unit, the students had to work on a task regarding the accuracy of a measuring instrument and had to relate it to previous information on accuracy from the general science and biology units by means of the EQI instrument. The students were expected to use the general information about accurate measurements to evaluate the accuracy of their traffic situation inquiry plans (hDC10). Table 5.2 presents an overview of the content of the designed teaching-learning process with the inquiry tasks and supporting tasks in the three successive general science, biology and physics inquiry units related to the hDCs.

## 5.4 RESEARCH QUESTION

This study was carried out to determine whether the three functions of the EQI instrument regarding self-evaluation, support by peers and teacher, and facilitating transfer were fulfilled in the designed teaching-learning process. In this teaching-learning process, pre-university science students evaluated the ARV of inquiries in different school science subjects by using the EQI instrument. This research study was guided by the following research question:

> *To what extent does the designed teaching-learning process fulfil the three intended functions of the EQI instrument when the teaching-learning process is enacted in class?*

Table 5.2
*The teaching-learning process with different student tasks in the three inquiry units with the corresponding hypothetical design characteristics*

| Student tasks | Teaching-learning process | | | Corresponding hypothetical design characteristic (hDC) |
|---|---|---|---|---|
| | General science inquiry unit | Biology inquiry unit | Physics inquiry unit | |
| **Inquiry task** *orientation, conduct, inferring and reflection* | Cooling of coffee inquiry | Human taste inquiry | Traffic situation on a dangerous crossing inquiry | hDC1 hDC2 hDC3 hDC4 hDC9 |
| **Orientation tasks** | Introduction EQI instrument + Evaluation by teacher | Evaluation by teacher | Evaluation by teacher | hDC1 hDC7 hDC8 |
| **Peer feedback tasks** | Discussion on ARV with one other group | Discussion on ARV with at least three groups | Discussion on ARV with one other group | hDC5 hDC6 |
| **Transfer tasks** | Evaluation of ARV in microbiological research | Evaluation of reliability in research with living and non-living objects | Evaluating accuracy in relation to measuring instrument and representative sample | hDC1 hDC10 hDC11 |

## 5.5 METHOD

This study involved the same experimental situation as the previous study (Chapter 4), but with the use of different data to answer different research questions. The participants were the same, as were the materials and procedures.

### 5.5.1 Participants

Twenty-seven pre-university science students, aged 16 or 17, voluntarily participated in the second research cycle. These are the same students who participated in the study that was described in Chapter 4. They signed up to participate after receiving an e-mail about the study from their science teachers. All of the students studied biology, physics and chemistry at pre-university level at the same school. They all had experience with practical work in the science subjects, but they had never systematically evaluated ARV in an inquiry. The students received a small financial incentive after finishing the last inquiry session to reward them for the effort involved in participating in the study after regular school periods.

### 5.5.2 Micro design cycles

In this study, a design-based research approach, as described by Gravemeijer and Cobb (2006), was used. This approach was chosen to sharpen the design characteristics of a teaching-learning process wherein pre-university science students used the EQI instrument to evaluate the ARV of inquiries. Two *micro design cycles* (Gravemeijer & Cobb, 2006, p. 24) were undertaken to arrive at the design characteristics. Each microcycle contained a design phase and a phase of analysis of and reflection on the educational design.

This study presents the results of the phase of analysis of and reflection on the educational design in the second micro design cycle. In the preceding (first) micro cycle, the researcher made the first version of the educational design. Its expected feasibility in an educational setting – a pre-university science class – was discussed by a design team consisting of three experienced and qualified pre-university science teachers and four educational researchers. After reaching consensus on changes to the educational design, the design was further elaborated. Thereafter, the three teachers checked the design of the teaching-learning process in class. The designers' reflections on the results of the testing phase of the first micro cycle led to further consensus-based refinements of the educational design for the second micro cycle.

These refinements of the teaching-learning process and learning materials concerned two main points. First, the instruction by the teacher about the evaluation of ARV with the EQI instrument was described in more detail in the teacher's guide. Second, more identifying marks and analogies regarding the evaluation of ARV were added to the student workbooks for the inquiry units. After revision of the design of the teaching-learning process, three researchers independently analysed the adapted design to determine whether all intended functions of the EQI instrument could be fulfilled in the teaching-learning process. These analyses were discussed again in the design team, and if necessary the educational design was changed, until consensus was reached (Janesick, 2000).

### 5.5.3 Procedure

The students worked in twelve groups of two or three students on the three successive inquiry units. The composition of the groups was the same in all units. The students worked in seven afternoon sessions, each of three hours, over a time span of three months. The decision to depart from regular lesson periods of 50 minutes was made because the students' inquiries required more time. Moreover, interruptions that often happen in regular school periods were avoided. It should be remarked that, due to heavy snowfall during the last afternoon session of the physics inquiry unit, the student groups could not conduct their inquiries on the traffic situations, could not draw conclusions and did not evaluate the ARV of their completed inquiry. For the peer discussion about the ARV in the physics unit, the inquiry plans were used instead of the evaluation of the ARV of the completed inquiries.

The researcher – who holds a master's degree in biology and a degree in teaching, and has eight years of experience in pre-university biology teaching – was the teacher of the various inquiry units in the second research cycle. The fact that she was one of the designers of the EQI instrument avoided misinterpretation by the teacher of both the content of the instrument and the intended teaching-learning process. Moreover, having one teacher to teach all the inquiry units reduced the distortion of continuity that would result from three different teachers being involved in the teaching-learning process. Two precautions were taken in the research procedure to prevent the 'dual role' of teacher and designer becoming a limitation of the study. First, two independent observers were present in all afternoon sessions to observe whether the lessons were taught as intended and to check whether the teacher directly influenced the indicators in the teaching-learning process. Second, all collected data were independently analysed by two other researchers, as is described in more detail in the section on data analysis (see 5.5.5).

### 5.5.4 Data collection

In the data collection, triangulation of data sources was used (Yin, 2003). Table 5.3 presents the various instruments used to collect data in order to investigate the indicators for three design characteristics in the teaching-learning process.

Table 5.3
*Overview of the instruments for data collection regarding the different indicators*

| | Data collection instruments | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Indicator** | **Audio recordings** | **Video recordings** | **Workbooks** | **EQI instrument** | **Field reports** | **Reflection reports teacher** | **Questionnaires** | **Interviews** |
| **1a** Students understand concepts of EQI instrument | x | | | x | | | | x |
| **1b** Students use EQI instrument at intended points | | x | | x | x | | | |
| **1c** Students change ARV of inquiries | | x | x | | | | | |
| **1d** Self-evaluation before peer and teacher feedback | | x | | | x | x | | |
| **2a** Peers give feedback before teacher | x | x | | | x | x | | |
| **2b** Peers suggest changes of ARV | x | x | x | | x | x | | |
| **2c** Teacher gives written feedback with EQI instrument | x | x | | x | x | x | | |
| **2d** Teacher gives oral feedback followed by discussion | x | x | | | x | | | |
| **3a** Students recognise analogies in different inquiries | x | x | | | | x | x | x |
| **3b** Students recognise analogies with scientific research | x | x | x | | x | x | | |
| **3c** Students apply knowledge on differences in evaluating ARV | x | x | x | | x | x | | |

***Function 1: Self-evaluation of ARV***

One of the observers video-recorded the teaching-learning process. Relevant parts of the recordings about the self-evaluation of the ARV were transcribed by means of the multimedia qualitative data analysis programme Atlas.ti, version 6. The transcripts were divided into identification (ID) units. Each ID unit was labelled for a particular student group (n=12) to make it possible to create an overview of the performances of each student group for the data analysis. The other observer made observations and wrote a field report to record how the intended inquiry and supporting tasks regarding self-evaluation were enacted in class. Both observers were acquainted with the intended teaching-learning process and the intended use of the EQI instrument. All discussions between students in a group and between teacher and students concerning the moments of self-evaluation with the EQI instrument were audio-recorded and transcribed. After each afternoon session, the teacher wrote a reflection report in which she compared the enactment of the self-evaluation with the intended tasks in the educational design. The teacher's reflection reports were verified and signed by the observers.

After each afternoon session, the parts from the EQI instrument that were filled out by the students were collected. From the student workbooks of all three inquiry units, all changes that the students made concerning preparation, conduct, data handling and evaluation of the inquiries after using the EQI instrument were collected. These changes were visible in the workbooks because the students were instructed to use a pen or pencil in another colour whenever a change was made.

Furthermore, one week after the last afternoon session, all the students, in pairs, were involved in a one-hour interview (in total 14 interviews) undertaken by two researchers who had not been present in the seven learning sessions. In the interviews, the EQI instrument was at hand and the groups answered the following questions: Did the EQI instrument help you to evaluate the ARV in an inquiry? If no, why? If yes, how did it help you to guard the ARV in your inquiry during the afternoon sessions? All interviews were audio-recorded and transcribed. The transcripts were divided into ID units for each student.

***Function 2: Support by peers and teacher***

To determine the support provided by peers and the teacher, the same approach of video and audio recordings, as described under function 1, was followed. This also accounted for the method used by the observers and the teacher's reflection report. Immediately after each afternoon session, the students filled out a questionnaire containing, among others, the open question: How did the discussion between your group and the other group progress? Moreover, peers' suggestions for changes regarding the ARV of inquiries were collected from the student workbooks.

***Function 3: Facilitating transfer***

To determine the support of item transfer from the EQI instrument from one inquiry to another, the same approach – video and audio recordings, field reports, reflection reports from the teacher, questionnaires and interviews –, as described under function 1, was used. In the questionnaire, two open questions were relevant for the support of transfer. These questions were: (1) Can you explain whether the ARV card, checklist and rubrics were useful during this afternoon session? (2) Did you use the ARV card, checklist and/or rubrics to change your inquiry? If yes, which parts and/or descriptions were useful to you? Concerning the interviews, two questions were relevant to determine support for transfer. These questions were: (1) Can you indicate at what points you used the EQI instrument during the inquiries? (2) During which tasks in the three inquiry units did the EQI instrument support the evaluation of the ARV of an inquiry?

### 5.5.5 Data analysis

The data analysis was focused on the indicators in the teaching-learning process for the three functions of the EQI instrument. In the data analysis, a qualitative method was used (Denscombe, 2007). Whenever relevant (see Table 5.3), the analyses of different data sources were compared.

All data were independently analysed on the indicators by two researchers, with an overall proportion agreement of 72%. See Table 5.4 for the proportion agreement for each indicator and the average proportion agreements for all data collection instruments. All differences between both researchers' data analyses were discussed until consensus was reached (Janesick, 2000). As can be seen in Table 5.4, the data analysis of the written data sources (students' workbooks, EQI instrument, field reports, reflection reports and questionnaires) showed a higher level of agreement than the data analysis of the transcripts of the primary oral data sources (video and audio recordings in classroom and interviews). Analysing oral data sources is more difficult than analysing written data sources, because often a speaker stops in the middle of a sentence and starts to talk about something else he or she has in mind, while a person mostly completes his or her sentences when writing. Due to this, exact meanings are clearer in written data than in oral data sources (Polkinghorne, 2005). In the study described in this chapter, in the analysis of each indicator, a combination of oral and written data sources was used (see Table 5.3) and all differences between both researchers' data analyses were discussed until consensus was reached (Janesick, ibid.). Therefore, it was ensured that the moderate overall proportion agreement of 72% did not undermine the strength of the findings and the conclusions.

***Function 1: Self-evaluation***

To determine whether the teaching-learning process fulfilled function 1 of the EQI instrument, self-evaluation of the ARV of an inquiry with the EQI instrument, the data

was analysed on indicators 1a, 1b, 1c and 1d (see Table 5.1). First, the data of the ID units from the transcripts of audio, video and interview recordings, as well as the filled-out EQI instrument, were scored to determine the number of students who showed understanding of the CoE items in the EQI instrument (indicator 1a). Secondly, the filled-out EQI instrument, the field reports and the ID units of transcripts of the video recordings were used to score the number of student groups that used the EQI instrument at the intended points in the teaching-learning process (indicator 1b).

Table 5.4
*Proportion agreement (in %) of the independent data analysis by two researchers*

| | Percentage of proportion agreement | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Indicator** | **Audio recordings** | **Video recordings** | **Workbooks** | **EQI instrument** | **Field reports** | **Reflection reports teacher** | **Questionnaires** | **Interviews** | **Average** |
| **1a** Understanding of concepts | 31 | ■ | ■ | 94 | ■ | ■ | ■ | 72 | 66 |
| **1b** At intended points | ■ | 78 | ■ | 100 | 100 | ■ | ■ | ■ | 93 |
| **1c** Change of ARV | ■ | 48 | 82 | ■ | ■ | ■ | ■ | ■ | 65 |
| **1d** Before peer and teacher | ■ | 83 | ■ | ■ | 100 | 91 | ■ | ■ | 91 |
| **2a** Peers before teacher | 79 | 85 | ■ | ■ | 98 | 95 | ■ | ■ | 89 |
| **2b** Peers suggest changes ARV | 51 | 48 | 86 | ■ | 89 | 69 | ■ | ■ | 69 |
| **2c** Teacher written feedback | 83 | 85 | ■ | 100 | 100 | 100 | ■ | ■ | 94 |
| **2d** Teacher oral feedback | 39 | 69 | ■ | ■ | 81 | ■ | ■ | ■ | 63 |
| **3a** Recognise analogies(units) | 41 | 34 | ■ | ■ | ■ | 61 | 83 | 54 | 55 |
| **3b** Recognise analogies (science) | 43 | 39 | 78 | ■ | 67 | 74 | ■ | ■ | 60 |
| **3c** Apply differences | 39 | 35 | 85 | ■ | 63 | 73 | ■ | ■ | 59 |
| **Average** | **51** | **60** | **83** | **98** | **87** | **80** | **83** | **63** | ■ |

*Note. The proportion agreement is rounded off to whole numbers.*
*The total proportion agreement of 72% was calculated with all individual percentages.*

With regard to indicator 1c, the workbooks and ID units in the transcripts of the video recordings were analysed to determine the number of student groups that made changes after using the EQI instrument. The changes that were made by each student group were categorised as:

1) changes of aspects of ARV
2) other changes (e.g. changes of language), unclear changes or no changes.

Lastly, the field reports, the teacher's reflection reports and the ID units in the transcripts of the video recordings were analysed to determine the number of student

groups that evaluated the ARV of their inquiries with the EQI instrument in all inquiry units before they received feedback from peers and the teacher (indicator 1d).

***Function 2: Support by peers and teacher***

To determine whether the teaching-learning process fulfilled function 2 of the EQI instrument, support of self-evaluation by peers and teacher, the collected data was analysed on indicators 2a, 2b, 2c and 2d (see Table 5.1). First, the field reports, the teacher's reflection reports and the ID units in the transcripts of audio and video recordings were analysed to determine, for each inquiry unit, the number of student groups that discussed the ARV of their inquiries with peers before the teacher gave feedback (indicator 2a). Second, with regard to indicator 2b (peers suggest changes on ARV), the field reports, the teacher's reflection reports, the students' workbooks and the ID units in the transcripts of audio and video recordings were analysed. This analysis determined, for all three inquiry units, the number of student groups that:

1) mentioned CoE items from the EQI instrument
2) mentioned general aspects related to the quality of an inquiry/gave unclear suggestions/gave no suggestions.

Third, the field reports, teacher's reflection reports, ID units in the transcripts of audio and video recordings and filled-out EQI instrument were analysed to determine the number of student groups that received written feedback from the teacher at all intended points in the teaching-learning process (indicator 2c).

With regard to indicator 2d (teacher oral feedback), the field reports and the ID units in the transcripts of audio and video recordings were analysed to determine the number of student groups that received oral feedback from the teacher at all intended points in the teaching-learning process. Furthermore, all oral feedback was scored on:

1) focus on the change of evaluation of ARV of an inquiry
2) focus on aspects of an inquiry other than ARV.

***Function 3: Facilitating transfer***

To determine whether function 3, facilitating transfer of self-evaluation of ARV in inquiries in different school science subjects, was fulfilled in the teaching-learning process, the data was analysed on indicators 3a, 3b and 3c (see Table 5.1). First, the questionnaires, ID units in transcripts of audio, video and interview recordings and teacher's reflection reports were analysed to determine the number of students who showed recognition of analogies in different inquiries by means of using the EQI instrument (indicator 3a). Second, the field reports, teacher's reflection reports, transcripts of audio and video recordings and student workbooks were analysed to determine the number of student groups in the general science unit who applied knowledge about the ARV of the scientific probiotics research study to their own inquiry. In addition, the number of student groups in the physics unit that applied

knowledge about the accuracy of measuring instruments to their own inquiry plan was determined. Third, the field reports, teacher's reflection reports, ID units in transcripts of audio, video and interview recordings and student workbooks were analysed to determine the amount of student groups that appropriately applied knowledge about differences in evaluating the ARV in inquiries of different school science subjects in their respective personal biology and physics inquiry plans (indicator 3c).

### 5.5.6 Set norms for indicators

The norm for fulfilling the three functions of the EQI instrument was set at 80% of the students (indicator 1a and 3a) or student groups (other indicators) – this means at least 22 students or ten student groups whose data sources showed the indicators were concerned (Juran, Gryna, & Bingham, 1974).

## 5.6 FINDINGS

### 5.6.1 Function 1: Self-evaluation of ARV

***Indicator 1a: Understanding of content of EQI instrument***

All student groups performed the orientation task about the meaning of the items on the ARV card, as was shown in the video recordings. The field report for this afternoon session read: 'The students were very active during the ARV [orientation] task. They had fierce discussion about whether a described situation contributed to the accuracy, reliability or validity of an inquiry. Even with the reference model at hand, they could not end their discussions.'

The audio and video recordings showed that after the general science unit, nine student groups (75%) could explain the meaning of the CoE items in the EQI instrument. After the biology unit, ten student groups (83%) could explain this. From the analysis of transcripts of the interviews, it appeared that all 27 students (100%) were able to explain how the CoE items in the EQI instrument could be used for the evaluation of the ARV in an inquiry and how the ARV card, checklist and rubrics could be used in the evaluation of the ARV.

Five CoE items from the EQI instrument had to be explained by the teacher, in addition to the planned instruction. The teacher's reflection report and the audio and video transcripts revealed that during the seven afternoon sessions the teacher explained the item 'Hypothesis can be tested by inquiry method' in the EQI instrument to three student groups. Moreover, the items 'Sample is large enough' and 'Conduct control experiment' were explained to the whole class in the general science unit after it appeared that at least six groups could not explain these items. When the students indicated during the biology unit that they did not know how many persons to take as a sample in their human taste inquiries, the item 'Sample is large enough' was again explained to the whole class. Analysis of the audio recordings revealed that

in the human taste inquiry, five student groups interpreted the item 'Measuring instrument is calibrated' as: *'Yes … not really putting on zero, but asking the subjects of experiment to rinse their mouths is also a kind of putting on zero'*. The video recordings showed that the teacher thereafter explained this item in class. In the audio and video recordings of the physics unit, it was illustrated that all twelve student groups (100%) used the item 'Measuring instrument is calibrated' as intended in their inquiry plans on the traffic situation.

***Indicator 1b: Use of EQI instrument***

The filled-out EQI instrument, the field reports and the video recordings showed that all twelve student groups (100%) used the EQI instrument at all intended points in the educational design. A more detailed analysis showed that, in each of the three inquiry units, one or more student groups neglected the instructions in the workbook on the use of the EQI instrument at a particular point, for example, the evaluation of the ARV of an inquiry plan. In all cases, the teacher noticed this and pointed out the instructions to the students. Comparison of the inquiries showed that in the general science unit four student groups (33%) followed the instructions in the workbook by themselves, whereas in the physics unit eleven student groups (92%) followed the instructions.

Moreover, the video recordings showed that during the general science unit two student groups (17%) used the EQI instrument while conducting their own inquiry and analysing their inquiry results, in addition to the use of the instrument in checking and evaluating the completed inquiry. The video recordings of the biology unit showed that six student groups (50%) used the EQI instrument when they wrote their inquiry plan and six groups (50%) used the instrument while conducting and inferring their inquiry. While writing the inquiry plan of the physics unit, five groups (42%) made use of the EQI instrument before they were advised to check or evaluate their inquiry plan.

***Indicator 1c: Changing the ARV of an inquiry***

During the general science unit, six student groups (50%) adequately changed aspects regarding the ARV of their inquiry plan, as was shown by analysis of the changes made by the students in their workbooks after using the EQI instrument, and as seen in the video recordings (score category 1) (see Table 5.5, p. 121). Three student groups made changes that were not related to the ARV of their inquiry plan and three student groups did not make any changes (50%; score category 2). In the biology and physics units, the analysis revealed that ten student groups (83%) adequately adapted the ARV of their inquiry plans (score category 1) and two student groups (17%) made changes that were not related to ARV (score category 2). All twelve groups (100%) wrote general comments about the ARV of their completed inquiries in their work-

books after they had used the EQI instrument during all three units, but none of the student groups (0%) repeated or adjusted separate aspects of their inquiries following the evaluation to try to improve the ARV.

***Indicator 1d: Self-evaluation first***

From the field reports, the teacher's reflection reports and the video transcripts, it was found that in the general science unit, ten student groups (83%) evaluated the ARV of their completed inquiries before they received feedback from peers and the teacher. The other two groups first evaluated all parts of their inquiry with their peers, then filled out the EQI instrument, and finally got feedback from the teacher. In the biology and physics units all twelve student groups (100%) first evaluated their inquiries with the EQI instrument before receiving feedback from their peers and the teacher.

Table 5.5

*Findings (in %) regarding the indicators for the three intended functions of the EQI during the teaching-learning process in the three inquiry units*

| Indicators | | Percentage of students groups in: General science inquiry unit | Biology inquiry unit | Physics inquiry unit |
|---|---|---|---|---|
| **1a** | Students understand concepts of EQI instrument[a] | 75 | 83 | 100 |
| **1b** | Students use EQI instrument at intended points | 100 | 100 | 100 |
| **1c** | Students change ARV of inquiries | | | |
| | i. Inquiry plan | 50 | 83 | 83 |
| | ii. Conduct, inferences, reflection on completed inquiry | 0 | 0 | n.a. |
| **1d** | Self-evaluation before peer and teacher feedback | 83 | 100 | 100 |
| **2a** | Peers give feedback before teacher | 100 | 100 | 100 |
| **2b** | Peers suggest changes of ARV | 50 | 42 | 83 |
| **2c** | Teacher gives written feedback with EQI instrument | 100 | 100 | 100 |
| **2d** | Teacher gives oral feedback followed by discussion on inquiry plans[b] | 100 | 100 | 100 |
| **3a** | Students recognise analogies in different inquiries[a] | n.a. | 41 | 78 |
| **3b** | Students recognise analogies with scientific research | 8 | n.a. | 75 |
| **3c** | Students apply knowledge on differences in evaluating ARV | n.a. | 75 | 42 |

*n.a. means not applicable.*

*a Percentage of individual students*

*b Written feedback on the completed inquiries was discussed in a whole-class discussion instead of individually with the twelve student groups (see 5.6.2).*

***Summary findings function 1: Self-evaluation of ARV***

Regarding the first function of the EQI instrument, self-evaluation of ARV by students, the findings of the biology and physics units show that each of the four indicators was present in the data sources of more than 80% of the students. It should be noted that indicator 1c, changes of the ARV, was only found in more than 80% of the students' inquiry plans and not in the completed inquiries, although all student groups made general comments on changing the ARV of their completed inquiries. None of the student groups repeated separate aspects of their inquiries to try to improve the ARV. Table 5.5 shows an overview of the findings regarding the indicators for the three intended functions of the EQI instrument in the teaching-learning process during the general science unit, the biology unit and the physics unit.

### 5.6.2 Function 2: Support by peers and teacher

***Indicator 2a: Peer feedback before teacher***

In all three inquiry units, all twelve student groups (100%) discussed the ARV of each other's inquiry with a peer group before receiving feedback from the teacher, as was shown in the field reports, the teacher's reflection reports and the audio and video recordings. In the biology unit, nine student groups (75%) discussed the ARV in each other's inquiries with three or more peer groups, as was intended in the educational design. The three other groups (25%) discussed each other's inquiries with one or two other groups.

Moreover, in the teacher's reflection report of the biology unit it was found that the teacher wrote on the peer discussion:

- *Some student groups tried to talk to three other groups as soon as possible. I listened to a peer discussion, or in fact a conversation between two students. One student asked, 'What would you yourself change on the ARV in your inquiry?' The other student mentioned enlargement of the sample and she answered: 'Yes, I would do the same.' And she went with her group to the next group.*

Analysis of the questionnaire data showed that one student wrote after completing the biology unit:

- *'The discussion with other groups was very chaotic, more order was needed.'*

***Indicator 2b: Peers' suggestions on the ARV***

During the general science unit, each student group explained to their peers what they actually inquired into, after which their peers gave suggestions on how to change the ARV in the inquiry, as was shown in the audio recordings. In three discussions between two peer groups (50% of the student groups), the students tried to convince their peers to change the ARV with adequate arguments related to items from the EQI instrument (score category 1). For example, *'You should repeat the inquiry*

*with a thermometer that was surely calibrated on zero degrees* [Celsius]. *Now you do not know whether it was measuring the right temperature'*. In the other three peer group discussions (50%), the peer groups exchanged information to get clarity on ARV in the inquiries, but did not make suggestions to change the ARV (score category 2). Analysis of the audio recordings and the students' workbooks showed that during the biology unit five student groups (42%) made suggestions for ARV changes in peer groups' inquiries (score category 1). The other seven groups (58%) made general comments on the inquiries of peers or repeated what their peers had mentioned. These seven groups wrote, for example, *'Hypothesis'* and *'Change subjects of experiment'* (score category 2). During the physics unit, in five discussions between two peer groups (83%) items were discussed from the EQI instrument to change the ARV (score category 1). In one peer group discussion (17% of student groups), general suggestions about the quality of an inquiry were made, such as: *'The table* [with measurements] *could be presented more clearly'* and *'You could have written more in your workbook so that it was clearer to us'* (score category 2).

***Indicator 2c: Written feedback from teacher***

In all three inquiry units, the teacher gave written feedback on all twelve student groups' (100%) inquiry plans with the checklist and rubrics from the EQI instrument. She also filled out rubrics from the EQI instrument on the completed inquiries for all twelve student groups (100%). Analysis of the teacher's reflection reports showed that the teacher wrote some extra comments below the checklist or a rubric when she thought the EQI instrument was not supportive enough and perhaps needed revision, for example *'How many times will you repeat the measurements?'* in the cooling coffee inquiry. Next to this, in the teacher's reflection report there was found a comment about the written feedback:

- *Due to a lack of time I could not fill out a checklist for all student groups while they were conducting their inquiries. When necessary, I gave all groups oral feedback on the items from the checklists, after which the students could make changes during conducting the inquiry.*

***Indicator 2d: Oral feedback from the teacher***

The data in the field report, as well as in the audio and video recordings, showed that the teacher gave twelve student groups (100%) oral feedback on the conduct of their inquiries in the three inquiry units with the checklist at hand. After returning the filled-out checklists and rubrics from the EQI instrument to the students, the teacher went around the twelve groups to ask whether the students understood the written feedback. Furthermore, she gave oral feedback to all twelve student groups, and to six student groups (50%) she gave information on the interpretation of and the descriptions in the EQI instrument.

Regarding the content of the discussions between the student groups and the teacher during the three inquiry units, the analysis of the field reports and audio and video recordings revealed that students did not ask any questions concerning the teacher's written feedback. The observer wrote in the field reports: *'All student groups, except one, studied the teacher's written feedback; three student groups changed their inquiry plan without consultation of the teacher.'* In all inquiry units, the teacher took the initiative to discuss one or more changes in the ARV of their inquiry plan with the students, for example in the general science unit: *'Do you have suggestions how the validity of the inquiry question can be improved?'* One student from the group answered: *'Well, perhaps more details, or [be] more precise about the measurements.'* Then the teacher asked them to look at various levels of the rubric 'Inquiry question', after which one of the students remarked: *'Oh, we should say something about the cooling rate of coffee as is described in the theoretical framework in the workbook.'* Similar examples of discussion between teacher and students were found in the biology and physics units.

The teacher's filled-out rubrics about the completed inquiries were not discussed by the teacher and students in any of the inquiry units. As the teacher wrote in her field report of the biology unit:

- *At the beginning of this afternoon session, I handed out my filled-out rubrics from the cooling coffee inquiry* [general science unit]*, but only asked students to look at aspects they could improve in future inquiries. Perhaps it would have been useful for the students to have a retrospective discussion on the completed inquiry, but I did not know how to manage this in class, discussing with one student group while other groups are waiting. So I decided to make some general comments about the ARV of the completed inquiries.*

The video recordings showed that the teacher thereafter discussed the influence of calibrating of measuring instrument on the accuracy of measurements. She also asked students about the control of other influencing variables, after which one student answered: *'Next time we should keep the window closed to prevent the cooling of some of the cups with hot water [being] influenced.'* At the beginning of the sixth afternoon session, just before starting the physics unit, a similar whole-class discussion took place.

***Summary findings function 2: Support by peers and teacher***

The findings regarding the teaching-learning process' fulfilment of the second function of the EQI instrument, support by peers and teacher, show that in all inquiry units peer feedback was given before teacher's feedback by all student groups (indicator 2a). The norm for indicator 2b is met in the physics unit: more than 80% of the student groups discussed changing the ARV of their inquiries. In the preceding units, this set norm for indicator 2b was not met. Next, the findings showed that the

norm for indicator 2c was met because the teacher gave written and oral feedback in all units at all intended points to all student groups, although she chose to give oral feedback with a checklist during the conduct of the inquiry instead of written feedback. The norm for indicator 2d is partly met: the teacher discussed her written feedback on the inquiry plans with all student groups, but clarified her written feedback on the completed inquiries in a whole-class discussion (see Table 5.5).

### 5.6.3 Function 3: Facilitating transfer

#### *Indicator 3a: Analogies in different subjects*

The data in the questionnaires that were filled out after the biology unit sessions showed that eleven students (41%) recognised analogies in the learning materials and educational design. In the questionnaire they wrote comments such as: *'It was useful in this inquiry to use the same tables* [the EQI instrument] *to change your inquiry, that's helpful in increasing your understanding of inquiries.'* After the physics unit, 21 students (78%) made these types of comments. One of these students wrote on the use of analogies, *'You had to do the same thing every time* [during the inquiry units]. *This would be useful when the inquiries were not one immediately after the other. Now you know it by heart and do not need it* [the EQI instrument] *again.'* The teacher noticed that students were better in using the EQI instrument in the physics unit than they were in the preceding units. She wrote in her reflection report after a physics unit session:

- *The students are now* [in the physics unit] *faster in using the EQI instrument than before* [in the general science and biology units]. *They seem to understand better and better how to use it in their inquiry. I am not very sure whether they read the descriptions well enough or think 'it runs something like this'.*

#### *Indicator 3b: Analogies with scientific research*

In the general science unit, only one student group (8%) explicitly applied knowledge about the ARV from the exemplary scientific research study on the use of probiotics when they worked on their own inquiry plan. They made use of the ARV card and checklist from the EQI instrument during the task about probiotics, as well as while writing their own inquiry plan. The above was analysed from the field reports, teacher's reflection reports, audio and video recordings and the student workbooks. The audio recordings of this student group revealed their discussion:

- *Yes, just as those researchers* [in the probiotic research study] *we did not think of all control variables. For example, whether the door was closed during our pilot experiment. It was not reliable anymore when someone opened the door and a gust of wind occurred.*

In the physics unit, nine student groups (75%) used the information from the task on the accuracy of measuring instruments to change their inquiry plan and, more

particularly, their inquiry method. The field reports and video recordings showed that these groups all made use of the same parts of the EQI instrument during the transfer tasks (see Table 5.2, p. 111) as well as while changing their inquiry plan. Regarding this situation, the teacher's reflection reports about the physics unit showed:

- *I saw that some groups used the task on the accuracy of the measuring instrument when working on their own inquiry plan. One group, for example, said that a spring rule is the most accurate measuring instrument to set a 9.2m trajectory.*

The field reports of the physics unit revealed:

- *In one group they discussed the accuracy of the measuring instrument and conclude[d] that for their inquiry a spring rule really is more accurate than a slide rule, whereas the task concerns the accuracy of measuring instruments in general. Therefore, they apply their general knowledge about measuring instruments before it is asked for* [in their workbooks].

***Indicator 3c: Knowledge of differences in evaluating ARV***

After the student groups wrote their inquiry plans in the biology unit, the teacher gave, as intended, an explanation of the difference between evaluating reliability in an inquiry with living test objects and one with non-living objects, which was shown by the video recordings and field reports. During the explanation, she highlighted the relevant items in the EQI instrument to the students. After the explanation, nine student groups (75%) wrote correctly in their workbooks the situations in which a representative sample is important, as well as when it is important to repeat an inquiry. The teacher's reflection report revealed:

- *During the whole-class explanation it seemed to take some time before the students understood why a representative sample contributes more to the reliability of an inquiry with living test objects than the repetition of measurements* [as in the cooling coffee inquiry]. *Only when I asked about the differences regarding internal differences between living test objects and non-living objects, a student answered: '*[...] *their DNA differs and you do not know what influence this has on the outcome of inquiry'.*

Six student groups (50%) wrote in their workbooks, concerning the evaluation of the reliability of their inquiry in the biology unit, that their inquiry would have been better if they had taken a representative sample. Furthermore, two other student groups (17%) wrote that a strong point regarding the reliability in their inquiries was that they repeated the measurement with their living test objects. Four student groups (33%) did not write about repeating measurements and taking samples in their evaluations of the reliability.

Further analysis of the physics unit revealed that five student groups (42%) were able to explain to the teacher correctly, during their own inquiry, at which points it was appropriate to repeat measurements and in which situations to take a sample. With the other seven student groups (58%) the teacher individually explained the appropriateness of repeating measurements and of taking a sample. Moreover, four student groups (33%) correctly mentioned in their inquiry plans in the physics unit which measurements they had to repeat and whether they had to take a representative sample in their inquiry.

***Summary findings function 3: Facilitation of transfer***

Regarding the extent to which the teaching-learning process facilitated transfer of self-evaluation of ARV from one inquiry context to another, the findings showed that in the physics unit, indicators 3a and 3b were present. The set norms of 80% of the students were almost met in this inquiry unit. In the biology unit (indicator 3a) and the general science unit (indicator 3b), the students seldom recognised the analogies between different inquiries and scarcely applied external knowledge about ARV. In the biology and physics units, indicator 3c about knowledge of differences in evaluating ARV in different inquiries was found, but the set norm was not reached (see Table 5.5).

## 5.7 DISCUSSION

The aim of this study was to determine whether three intended functions of the EQI instrument were fulfilled in a designed teaching-learning process wherein pre-university science students use the EQI instrument in evaluating an inquiry's ARV in different school science subjects. Regarding *function 1, self-evaluation of ARV by students,* it can be concluded that all students had sufficient understanding of the content of the EQI instrument to use it for self-evaluation of the ARV of an inquiry in the biology and physics units. During the teaching-learning process, the students had a central and active role in the evaluation of their inquiries' ARV. The students changed, as part of their self-evaluation, aspects to improve the ARV of their inquiry plans, but they did not change the ARV of their completed inquiries. For students, a completed inquiry seemed to be an 'end point', while we expected it to be a reflection point to gain procedural understanding on how the ARV could be improved in their future inquiries. In all inquiry units, the students did a self-evaluation with the EQI instrument before any peer or teacher feedback was provided. Regarding *function 2, support of self-evaluation by peers and teacher,* it can be concluded that students received peer feedback before the teacher gave feedback. During the general science and physics units, most of the students discussed with their peers the ARV of each other's inquiries and suggested changes of the ARV. In all units, the teacher gave written and oral feedback to the students at all intended points and discussed changes in the ARV of the inquiries with the students. Concerning *function 3, facilita-*

*tion of transfer,* it is concluded that students sometimes identified the analogies with the evaluation of ARV in the preceding inquiry units in the last part of the teaching-learning process, but this could be done much better by the students. Most students also applied knowledge about the evaluation of ARV that they acquired from evaluating the ARV of a scientific research study or from general tasks about the ARV of an inquiry. Only a few students expressed knowledge on differences in evaluating the ARV in different school science subjects before the teacher explained these differences.

In the next sections, the outcomes of this study will be discussed, followed by a reflection on the hypothetical design characteristics (hDCs) of the teaching-learning process to determine a set of revised design characteristics.

### 5.7.1 Function 1: Self-evaluation of ARV by students

As stated above, the students have a central and active role in the self-evaluation of the ARV in their inquiries. Most student groups' self-evaluation occurred before they received feedback from peers, and always before the teacher gave feedback (see Table 5.5, indicator 1d). As such, during the teaching-learning process as described in this study, all students had the chance to self-regulate their learning process, and passive reception of feedback by the students was avoided (Nicol & Macfarlane, 2006).

Although all students in all inquiry units used the EQI instrument at the intended points (see Table 5.5, indicator 1b), it can be concluded that in the first unit, general science, the teacher needed to be more active in indicating how the students should use the instrument than was expected by the designers. In this study, the instructions in the student workbook insufficiently triggered the students to use the EQI instrument of their own accord. The teacher therefore had to remind the students to use the EQI instrument for performing a self-evaluation. The role of the teacher in the process of modelling and guiding novices' self-evaluations is important and should be highlighted in a teacher's guide on introducing the EQI instrument and in any teacher training as an important step. The teacher is the only one who can adequately explain the relevance, content and use of the instrument on its first usage in class (Sadler & Good, 2006). In the next inquiry units, it was found that the students used the EQI instrument largely without (extra) guidance from the teacher. The teacher should guide individual students at particular moments to make sure that students use the EQI instrument with the understanding that the instrument is a means for self-evaluation of an inquiry and contains helpful information for changing the ARV of their inquiries. Doing this can probably prevent a situation occurring in which students, after a while, start to consider the use of the instrument as an inevitable and routine task (Andrade & Du, 2005; Tan, 2004). Another possibility to maintain students' attention at an appropriate level is to incorporate tasks on how accurate, reliable and valid

scientific research contributes to knowledge development in science. For most students it is not clear why inaccurate scientific research cannot be used to construct new technologies (Schwartz, Lederman, & Crawford, 2004). When students learn more about this aspect of knowledge construction, it can become clearer to them what the function of the use of the EQI instrument is in their own inquiries.

This study showed that in the biology and physics unit, most of the students made correct changes regarding the ARV of their inquiry plans after using the EQI instrument (Table 5.5, indicator 1c–i), although it is known that before this study, the students were not familiar with the CoE items about ARV of an inquiry. A tentative conclusion is that using the EQI instrument for self-evaluation improved the students' understanding on how to change the ARV of their inquiry plans. Therefore, it is reasonable to posit that the students, during the teaching-learning process, increasingly understood how the content of the EQI instrument could support improvement of evaluating the ARV of their inquiries. However, it is also found that students did not change the ARV of their completed inquiry after self-evaluation with the EQI instrument (Table 5.5, indicator 1c–ii). Usually, in the system of regular education, students are not often asked to reflect on their completed inquiries, let alone to change aspects of their inquiries or repeat their inquiries after completing them. For pre-university science students, the role played by reflection on inquiries and repeating an inquiry in scientific knowledge development is not often made explicit (Schwartz, Lederman, & Crawford, 2004).

Another reason for not changing the ARV of their completed inquiry may be that students had already looked at the EQI instrument during their conduct of an inquiry (see *Findings about indicator 1b*) and already confronted their own thinking with the content in the EQI instrument, instead of changing their written notes afterwards. This process of changing thoughts 'along the way' could lead to a better restructuring of knowledge than is provided by only reflecting after performing a task (Vosniadou & Brewer, 1987). Therefore, to fulfil the self-evaluation function of the EQI instrument completely, in a redesigned teaching-learning process the evaluation of ARV of completed inquiries should be accompanied by intervening ARV evaluations during the conduct of the experiments, drawing conclusions and reflecting on inquiries.

### 5.7.2 Function 2: Support by peers and teacher

Regarding function 2 of the EQI instrument, support by peer and teacher feedback, it could be concluded that, during the general science and physics units, sufficient students received peer support for their self-evaluation of inquiries' ARV. As was intended in the design of the teaching-learning process, peer discussions took place after self-evaluation and before the teacher gave feedback. Most students suggested changes to the ARV of their peers' inquiries (Table 5.5, indicators 2a and 2b). However,

during the biology unit a different peer feedback procedure was used. In this procedure most of the students only exchanged information, instead of having a real discussion on the ARV of a peer group's inquiry and without suggesting changes to the ARV. The students and the teacher experienced the peer feedback discussions during this unit as chaotic, and therefore the feedback as less effective than in the other two inquiry units with paired small-group feedback. Probably, peer feedback among novices progresses less chaotically when two students or two small groups of students are paired, instead of a procedure wherein students give feedback to several other (groups of) students. Two-by-two peer discussions create more focus on the content of the feedback (Nicol & Boyle, 2003). Additionally, students were not used to constructing valid, high-level arguments during group discussions about inquiries. This should be guided, or scaffolded, by a teacher. This scaffolding is easier in two-by-two discussions than in a discussion between more than two students or student groups (Kind, Kind, Hofstein, & Wilson, 2011). Consequently, in the teaching-learning process of this study, which was designed for novices, only two-by-two peer discussions should take place. When students gain more expertise in evaluating the ARV of an inquiry, the procedure of the peer discussion can be varied during the teaching-learning process to keep students motivated to have peer discussions (Ledford & Sleeman, 2000). Further research should address the question whether more focus in peer discussions of the biology unit leads to more and better suggestions for improvement of the ARV students' biology inquiries (Nicol & Macfarlane, 2006).

Regarding the teacher feedback (Table 5.5, indicators 2c, 2d), it can be concluded that the enactment of the teacher's written and oral feedback was planned at appropriate points during the teaching-learning process, except for the provision of written feedback during the conduct of inquiries by the students. It should be noted that the teacher, as one of the designers of the EQI instrument, fully understood the content of the instrument and was more or less expert in evaluating the ARV in students' inquiries due to the preceding design process of the EQI instrument. Because pre-university science teachers are mostly novices in evaluating ARV in an inquiry (see Chapter 2), it could be expected that the teacher in our study was probably better equipped to fill out the EQI instrument and to give appropriate oral feedback on the ARV of an inquiry than other teachers are. An in-service training trajectory can help teachers learn how to scaffold and support students in evaluating ARV in inquiries, although not much is yet known about how this scaffolding by the teachers should be designed (Akerson, Nargurd-Joshi, Weiland, Pongsanon, & Avsar, 2013; Bransford, 2000).

As noted in the theoretical framework, it is not clear whether and when written or oral feedback is most effective (Nicol & Macfarlane, 2006). In this study, both written and oral feedback from the teacher seemed to support the self-evaluation of the ARV, but

it is more practical and perhaps more effective to give students oral feedback while the inquiry is being conducted, instead of the intended written feedback by filling out a checklist. As mentioned before in the discussion about function 1, Vosniadou and Brewer (1987) concluded that changing thoughts 'along the way' could lead to a better restructuring of knowledge than simple reflection after performing a task, and this could also concern the support from the teacher's feedback. Further research is needed to gain more insight into the effectiveness of both types of teachers' feedback to support the students' self-evaluation of ARV.

### 5.7.3 Function 3: Facilitating transfer

The third intended function of the EQI instrument is to facilitate the transfer of ARV evaluation from one inquiry to another. Students identified analogies in the evaluation of ARV in the different inquiry units, culminating in the physics unit (see Table 5.5, indicator 3a). Hypothetically, this transfer effect might be enhanced when students, during their secondary education, learn to use the same EQI instrument and get the same questions, instructions and lay-out in the workbooks of relevant inquiry units during their pre-university school career. This implies the need for more inquiry units in which the students use the EQI instrument for evaluating the ARV in inquiries, as well as spreading them over several years of schooling. In addition, inquiry units could be developed in which various aspects of evaluating ARV in inquiries are further explored, such as the formulation of a valid inquiry question or the reliability of an inquiry method (Neber & Anton, 2008; Allie, Buffler, Campbell, & Lubben, 1998; Davidowitz, Lubben, & Rollnick, 2001).

This study found an increase in students' ability to recognise analogies between ARV in scientific research and in their own inquiries (see Table 5.5, indicator 3b). During the physics unit, most student groups applied general knowledge on the accuracy of measuring instruments to the accuracy of measuring instruments for their own inquiry. This is in strong contrast with the results in the general science unit, in which the students did not transfer knowledge about the ARV of the scientific probiotics research study to their own inquiry plan. An explanation could be that, due to the use of different CoE items (Salomon & Perkins, 1989), the gap between the evaluation of ARV in the probiotic research task and in the students' own inquiry about the cooling rate of coffee was too great. The difference between these two tasks could also be caused by the students' unfamiliarity with ARV evaluation in an inquiry during the transfer task about the probiotics research study in the first, general science unit. The abstract knowledge on ARV had probably not yet been internalised by the students and therefore could not yet be used in external instantiation tasks (Bransford, 2000; Chen, Yanowitz, & Daehler, 1995). It could also be that the students were overloaded with what was, for them, new knowledge in the probiotic research task, without the development of coherent mental maps of transferable knowledge (Gilbert, Bulte, &

Pilot, 2011) about the evaluation of ARV of an inquiry. It is necessary to analyse in more depth whether the students can connect information from the probiotic task to the content of the ARV card of the EQI instrument. Coherent mental maps of transferable knowledge perhaps develop better when novices first learn the concepts of ARV in relation to their own inquiries in successive, different inquiries (Van Rens, Pilot, & Van der Schee, 2010). For the teaching-learning process as designed in this study, the above could imply that transfer tasks aimed at applying knowledge about ARV from scientific research studies probably have to be introduced later on in a teaching-learning sequence of successive inquiry units.

Lastly, it can be concluded that explanations of differences in evaluating ARV in inquiries of different school science subjects insufficiently increase the students' ability to evaluate ARV in their own inquiries in different inquiries (see Table 5.5, indicator 3c). After three inquiry units, the students understood the meaning of the CoE items in the EQI instrument (see function 1), but were unable to transfer their understanding of repeating measurement and taking samples from one inquiry to another. The students as novices probably needed more practice in evaluating ARV in inquiries before they could flexibly handle the items in the EQI instrument in various inquiries. Furthermore, explanations of differences are only meaningful, or high-road transfer can only occur, when students have the mental abilities to concretise the explained, abstract knowledge in the inquiry in which they are working (Salomon & Perkins, 1989). Moreover, three inquiry units are limited in terms of the number of items in which differences regarding ARV evaluation in different school science subjects can be made explicit. As mentioned above, learning to evaluate the ARV of an inquiry should start in earlier school years. This implies the need for a teaching-learning trajectory with several successive units of inquiries in different school science subjects (Méheut & Psillos, 2004) and for teachers, or other experts, who can scaffold the transfer of the evaluation of ARV to different inquiries (Van Oers, 2001).

## 5.8 IMPLICATIONS FOR THE DESIGN CHARACTERISTICS

In this section, the hypothetical design characteristics (hDCs) are considered and, with respect to the findings of this study, the design characteristics are unchanged (DC), revised (rDC) or newly formulated (nDC). An overview of these design characteristics can be found in Table 5.6.

Regarding function 1 of the EQI instrument, self-evaluation of ARV by the students, it can be concluded that the designed teaching-learning process fulfilled this function to a sufficient extent. The hypothetical design characteristics (hDCs) about self-evaluation led to an educational design in which students could learn the language from the EQI instrument (DC1) and had a central role in evaluating the ARV of their inquiries (DC4). hDC2, about explanation of the relevance of self-evaluations, could

Table 5.6
*Design characteristics for fulfilling the three functions of the EQI instrument: unchanged design characteristics (DC), partly changed design characteristics (rDC) or newly added design characteristics (nDC). The rDC and nDC are written in italics*

| | |
|---|---|
| **Function 1: Self-evaluation of ARV with the EQI instrument** | |
| **DC1** | Learning tasks to learn concepts of EQI instrument. |
| **rDC2** | Explanation about relevance of self-evaluations and how to conduct these evaluations: *first guided then supported by teacher.* |
| **rDC3** | Supportive in changing ARV performance to higher levels during preparing, *performing, drawing conclusions* and reflecting on inquiry. |
| **DC4** | Central activity of the students. |
| **nDC12** | *Learning tasks about effective strategies in evaluating the ARV of inquiries.* |
| **Function 2: Support by peers and teacher** | |
| **DC5** | Peer feedback before teacher feedback. |
| **nDC13** | *Novices in peer feedback: only two-by-two peer feedback discussions.* |
| **rDC6** | First exchange of information between peers, then discussion. *Teacher: role in scaffolding during first peer discussions.* |
| **rDC7** | Teacher gives written feedback *on written students' products* linked to the criteria to which students can head. |
| **nDC14** | *Teacher gives oral feedback to students 'on the way'.* |
| **DC8** | Discussion between teacher and students about improving inquiry. |
| **Function 3: Facilitating of transfer** | |
| **rDC9** | Analogies in learning materials of different inquiries. *Analogies should be derived from earlier performed inquiry tasks or inquiry units.* |
| **rDC10** | Scientific research examples to explain abstract terms on ARV. *Novices: scientific research close to students' inquiries.* |
| **rDC11** | Explanation of differences in evaluating ARV in specific subjects *needs more than three inquiry units.* |
| **nDC15** | *Teaching-learning process needs smaller inquiry units to demonstrate one aspect of evaluating ARV in different school science subjects.* |

be worked out in more detail to lead to a more appropriate educational design to fulfil function 1. In the first inquiry units in which the EQI instrument was used, the teacher's role was to actively explain to the students how and why the EQI instrument should be used (rDC2). As the students gradually began to use the instrument by their own initiative, the teacher's role could be to motivate students to keep using the EQI instrument as a self-evaluation instrument to improve the ARV of an inquiry. In the designed teaching-learning process, hDC3 seemed to lead sufficiently to changes in the ARV of an inquiry plan, but changes in the ARV of the other phases of an

inquiry could probably be improved by introducing more ARV evaluations during those phases of performing, concluding and reflecting on an inquiry (rDC3). To fulfil function 1 more extensively, a new design characteristic (nDC12) about self-evaluation could be added: the teaching-learning process needs more tasks in which evaluating the ARV of inquiries could be learned by the students by teaching them effective strategies in evaluating. To prevent 'overloading' the students, it is worthwhile to study first at which points in the teaching-learning process these tasks can be introduced to be most effective for the students.

Function 2, support of the self-evaluation by peers and teacher, has also been fulfilled to a sufficient extent in the designed teaching-learning process. The hypothetical design characteristics (hDC5) about the support of self-evaluation by peers have led to a teaching-learning process wherein peer feedback was given before the teacher's feedback (DC5) and wherein peers mostly first exchanged information before they started a discussion about the ARV of an inquiry. The hDC6 can be supplemented with a more extensive role for the teacher (rDC6) during the peer discussions: the teacher should ask more questions to scaffold students in constructing valid, high-level arguments (Van de Pol et al., 2012). As a new design characteristic, for novices in peer feedback, discussions should take place between pairs of small student groups until students have attained a good understanding of all items on ARV in the EQI instrument (nDC13). About hDC7 and hDC8 it can be stated that both oral and written feedback has to be shown to be helpful in supporting the students. Written feedback can be given most practicably on students' written documents (rDC7). During performing the inquiry, it seems most practicable to give oral feedback (nDC14). Hypothetical design characteristic 8 also led to sufficient time in the educational design for discussion between teacher and student groups about improving the ARV of an inquiry (DC8).

Function 3, facilitating transfer, has been fulfilled less than functions 1 and 2 in the designed teaching-learning process, but first indications have been found that the use of the same EQI instrument in inquiries in different school science subjects gives students a better understanding of the analogies between these inquiries (hDC9). This hDC should be completed with 'these analogies should also be derived from inquiry tasks or inquiry units that were performed in previous school years' (rDC9). Regarding hDC10, it can be stated that some tasks with concrete, scientific examples actually illustrated the abstract terminology. For novices, it seems to be necessary for the scientific research topics used to be closely related to inquiries in pre-university science subjects. When the students have more procedural understanding of evaluating the ARV of an inquiry, the gap between the inquiries can be made larger (rDC10). To explain the differences in evaluating ARV in inquiries in different school science subjects, it seems to be necessary to have a series of more than three inquiry units

in which these differences for pre-university students can be taught (rDC11). Finally, to prevent students from experiencing cognitive overload during the inquiry units, it could be necessary to design smaller inquiry units (e.g. for a one-lesson period) to demonstrate one aspect of evaluating ARV in different inquiries (nDC15). However, a teaching-learning process aiming to perform function 3 appropriately needs more than three successive inquiry units to make it possible for students to apply their knowledge flexibly regarding the items in the EQI instrument in inquiries in different school science subjects.

Although some design characteristics are supplemented and a few new design characteristics identified in this study, it can be concluded that the eleven hDCs have led to a teaching-learning process wherein two of the three functions of the EQI instrument are sufficiently fulfilled. Hence, the students' performances in the teaching-learning process could be used to determine the effectiveness of the EQI instrument (see Chapter 6). It is expected that, by performing this teaching-learning process, students can improve their procedural understanding of evaluation of the ARV of an inquiry.

**REFERENCES**

Akerson, V., Nargund-Joshi, V., Weiland, I., Pongsanon, K., & Avsar, B. (2013). What third-grade students of differing ability levels learn about nature of science after a year of instruction. International Journal of Science Education. doi:10.1800/09500693.2012.761365.

Allie, S., Buffler, A., Campbell, B., & Lubben, F. (1998). First-year physics students' perceptions of the quality of experimental measurements. International Journal of Science Education, 20(4), 447–459.

Andrade, H., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. Practical Assessment, Research & Evaluation, 10(3).

Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. Theory into Practice, 48(1), 12–19.

Biggs, J., & Tang, C. (Eds.). (2007). Teaching for quality learning at university (3rd ed.). Buckingham, UK: Open University Press.

Bransford, J. D. (2000). How people learn: Brain, mind, experience, and school (Expanded ed.). Washington, DC: National Academy Press.

Chen, Z., Yanowitz, K. L., & Daehler, M. W. (1995). Constraints on accessing abstract source information: Instantiation of principles facilitates children's analogical transfer. Journal of Educational Psychology, 87(3), 445–454. doi:10.1037/0022-0663.87.3.445

College voor Examens (2013). Nieuwe examenprogramma's biologie, scheikunde en natuurkunde – centraal examen 2016 [New formal curricula Biology, Chemistry and Physics – final exams 2016]. Retrieved from http://www.examenblad.nl/9336000/1/j9vvhinitagymgn_m7mviwvlaj75kw6_n11vg41h1h4i9qe/vg41h1h7m8vx?topparent=vg41h1h4i9qe

Davidowitz, B., Lubben, F., & Rollnick, M. (2001). Undergraduate science and engineering students' understanding of the reliability of chemical data. Journal of Chemical Education, 78(2), 247–252.

Denscombe, M. (2007). The good research guide for small-scale social research projects (3rd ed.). Maidenhead, UK: Open University Press.

Gagné, R. M. (1985). The conditions of learning and theory of instruction. New York: Holt, Rinehart and Winston.

Gagné, E. D., Yekovich, C. W., & Yekovich, F. R. (1993). The cognitive psychology of school learning (2nd ed.). New York: HarperCollins.

Gibbs, G., & Simpson, C. (2004). Conditions under which assessment supports students' learning? Learning and Teaching in Higher Education, 1, 3–31.

Gilbert, J. K., Bulte, A. M. W., & Pilot, A. (2011). Concept development and transfer in context-based science education. International Journal of Science Education, 33(6), 817–837.

Gott, R., & Duggan, S. (1996). Practical work: Its role in the understanding of evidence in science. International Journal of Science Education, 18(7), 791-806.

Gott, R., Duggan, S., Roberts, R., & Hussain, A. (n.d.). Research into understanding scientific evidence Retrieved from http://www.dur.ac.uk/rosalyn.roberts/Evidence/cofev.htm

Gravemeijer, K., & Cobb, P. (2006). Design research from a learning design perspective. In J. Van den Akker, K. Gravemeijer, S. McKenney & N. Nieveen (Eds.), Educational Design Research (pp. 17–51). London and New York: Routledge.

Hodges, L. C., & Harvey, L. C. (2003). Evaluation of student learning in organic chemistry using the SOLO taxonomy. Journal of Chemical Education, 80, 785–787.

Hodson, D. (1999). Building a case for a sociocultural and inquiry-oriented view of science education. Journal of Science Education and Technology, 8(3), 241–249.

Janesick, V. J. (2000). The choreography of qualitative research design. In H. K. Denzin & Y. S. Lincoln (Eds.), Handbook of Qualitative Research (pp. 379–399). Thousand Oaks, CA: SAGE Publications.

Juran, J. M., Gryna, F. M., & Bingham, R. S. (Eds.). (1974). Quality control handbook (3rd ed.). New York: McGraw-Hill.

Kind, P.M., Kind, V., Hofstein, A., & Wilson, J. (2011). Peer argumentation in the school science laboratory. Exploring effects of task features. International Journal of Science Education, 33(18), 2527–2558.

Ledford, B. R., & Sleeman, P. J. (2000). Instructional design: A primer. Greenwich, CT: Information Age Publishing.

Levins, L., & Pegg, J. (1993). Students' understanding of concepts related to plant growth. Research in Science Education, 23, 165–173.

Lynch, R., Mannix McNamara, P., & Seery, N. (2012). Promoting deep learning in a teacher education programme through self- and peer-assessment and feedback. European Journal of Teacher Education, 35(2), 179–197.

Mayr, E. (1996). The autonomy of biology: The position of biology among the sciences. The Quarterly Review of Biology, 71(1), 97–106.

Méheut, M., & Psillos, D. (2004). Teaching-learning sequences: Aims and tools for science education research. International Journal of Science Education, 26(5), 515–535.

Minogue, J., & Jones, G. (2009). Measuring the impact of haptic feedback using the SOLO taxonomy. International Journal of Science Education, 31(10), 1359–1378.

Minstrell, J., & Stimpson, V. (1996). A classroom environment for learning: Guiding students' reconstruction of understanding and reasoning. In L. Schauble & R. Glaser (Eds.) Innovations in learning. New environments for education (pp. 175-202). Mahwah, NJ: Lawrence Erlbaum.

Neber, H., & Anton, M. (2008). Promoting pre-experimental activities in high-school chemistry: Focusing on the role of students' epistemic questions. International Journal of Science Education, 30(13), 1801–1821.

Nicol, D. J., & Boyle, J. T. (2003). Peer instruction versus class-wide instruction in large classes: A comparison of two interaction methods in a wired classroom. Studies in Higher Education, 28(4), 457–473.

Nicol, D. J., & Macfarlane, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. Studies in Higher Education, 31(2), 199–218.

Polkinghorne, D. E. (2005). Language and meaning: Data collection in qualitative research. Journal of Counseling Psychology, 52(2), 137–145.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.) Handbook of Self-Regulation (pp. 451–502). San Diego, CA: Academic Press.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. Instructional Science, 18(2), 119–144.

Sadler, P. M., & Good, E. (2006). The impact of self- and peer-grading on student learning. Educational Assessment, 11(1), 1–31.

Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: Rethinking mechanisms of a neglected phenomenon. Educational Psychologist, 24(2), 113–142.

Schalk, H. H., Van der Schee, J., & Boersma, K. Th. (2013). The development of understanding of evidence in pre-university biology education in the Netherlands. Research in Science Education, 43(2), 551–578.

Schwartz, R. S., Lederman, N. G., & Crawford, B. A. (2004). Developing views of nature of science in an authentic context: An explicit approach to bridging the gap between nature of science and scientific inquiry. Science Education, 88(4), 610–645.

Tan, K. H. K. (2004). Does student self-assessment empower or discipline students? Assessment & Evaluation in Higher Education, 29(6), 651–662.

Van de Pol, J., Volman, M., & Beishuizen, J. (2012). Promoting teacher scaffolding in small-group work: A contingency perspective. Teaching and Teacher Education, 28, 193–205.

Van Oers, B. (1998a). From context to contextualizing. Learning and Instruction, 8(6), 473–488.

Van Oers, B. (1998b). The fallacy of decontextualization. Mind, Culture, and Activity, 5(2), 135–142.

Van Oers, B. (2001). Contextualisation for abstraction. Cognitive Science Quarterly, 1(3-4), 279–306.

Van Rens, L., Pilot, A., & Van de Schee, J. (2010). A framework for teaching scientific inquiry in upper secondary school chemistry. Journal of Research in Science Teaching, 47(7), 788–806.

Vosniadou, S., & Brewer, W. F. (1987). Theories of knowledge restructuring in development. Review of Educational Research, 57(1), 51–67.

Yin, R. K. (2003). Case study research: Design and methods (3rd ed.). Thousand Oaks, CA: SAGE.