SUMMARY

Pre-university science students in the Netherlands are novices in evaluating accuracy, reliability and validity (ARV) in inquiries in the school science subjects of biology, chemistry and physics. ARV can be considered as three interconnected concepts that together lead to consistency in the evidence in inquiries. Therefore, evaluating the ARV in inquiries is part of the procedural understanding about the construction of evidence in an inquiry. By learning to (self-)evaluate the ARV in inquiries, pre-university science students can increase their procedural understanding and learn to determine how 'strong' the evidence in an inquiry is. Up to now, there has been limited research on how the self-evaluation of ARV in inquiries, as part of improving procedural understanding, can be done by pre-university science students and how they can transfer this self-evaluation of ARV from one inquiry to another, especially from one school science subject to other subjects.

Chapter 1 contains a general introduction of the research in this thesis. In the theoretical background, describes which aspects of learning to inquire support the improvement of procedural understanding. The accuracy of an inquiry is defined as how the researcher performs the measurements and observations of an inquiry. Reliability in inquiries is defined as whether repeating the inquiry leads to similar results (repeatability) and whether the results can be or are reproduced by others (reproducibility). The validity of an inquiry is defined as whether a scientific researcher is measuring or observing what he wants to measure or observe. The transfer of evaluation of ARV from one inquiry to another can be done by means of mindful abstraction (high-road transfer). Four important factors to support this flexible application of the evaluation of ARV in a teaching-learning process are:

saskia_proefschrift_m_summary.indd 229 27-10-16 10:48

(1) relevant basic procedural understanding, (2) procedural understanding is taught in a variety of inquiries, (3) recognition of subject-specific patterns , and (4) an active role of students. An active role of students can be stimulated by self-evaluation of their performances. For this self-evaluation, an instrument could be developed that is feasible to evaluate ARV in inquiries in different school science subjects and is effective in improving this aspect of procedural understanding of pre-university science students. Therefore, the main research question of this thesis is: *What are the design characteristics of a feasible self-evaluation instrument and a supportive teaching-learning process (wherein the instrument is used) with which pre-university science students can effectively learn to self-evaluate the accuracy, reliability and validity of inquiries in different school science subjects?* To answer this question, a design-based approach has been used composed of an explorative study and two research cycles of designing, testing and evaluating.

In **Chapter 2** (explorative study) the knowledge of pre-university science students and science teachers about the ARV in inquiries is explored. This explorative study made use of the Concepts of Evidence (CoE) model in which concepts are described that together support evidence in an inquiry. A part of the CoE seemed to be feasible for the evaluation of the ARV in an inquiry in the school science subjects. By means of a think-aloud task, interviews and a questionnaire it was determined which of the selected CoE should be elaborated in a self-evaluation instrument for evaluating the ARV in inquiries by pre-university science students. Twenty-three of the CoE seemed to be feasible for elaborating in such a self-evaluation instrument.

Chapter 3 (first test cycle) describes the design of a set of rubrics as a self-evaluation instrument. For this design, four hypothetical design characteristics about the content, extent of complexity, extent of details and extent of general application were identified from the explorative study and from the literature. The content of the 19 designed rubrics was based on 23 CoE that were identified in the explorative study as necessary and sufficient for novices in evaluating the ARV in inquiries. The SOLO taxonomy was used to describe the extent of complexity in each rubric. Each rubric got an a step-by step analytical trait, to make it possible to self-evaluate the ARV in the different inquiry phases. To make the rubrics applicable in inquiries in different school science subjects the general descriptions for evaluating are accompanied by working examples from performed inquiries. The designed rubrics were tested in three inquiry units. In each inquiry unit, the students evaluated the ARV in their inquiries after writing an inquiry plan, after performing the experiment, after handling the data and after completing the inquiry. The reflection on the feasibility of the rubrics in the students' inquiries led to revisions in the four hypothetical design characteristics. The most important revision was to describe (only) 13 of the CoE in rubrics and to design two other tools as part of the self-evaluation instrument in which all the CoE could be elaborated.

saskia_proefschrift_m_summary.indd 230 27-10-16 10:48

Chapter 4, 5 and 6 report on three studies which were performed during the second research cycle. In these studies respectively the feasibility of the revised instrument, a supportive teaching-learning process for using the instrument in evaluating the ARV in inquiries and the effectiveness of the instrument in evaluating the ARV in successive inquiries were determined.

Chapter 4 focuses on the feasibility of the revised self-evaluation instrument for the evaluation of ARV in inquiries by pre-university science students. From this study on, the instrument is called the Evaluation of Quality of Inquiry (EQI) instrument. The design of the EQI instrument was based on the four design characteristics as yielded from the first research cycle. The EQI instrument consists of three parts: (1) an overview (ARV) card, (2) a checklist to check whether all important steps in an inquiry were done and (3) a set of twelve rubrics for self-evaluating the ARV in inquiries. The feasibility of the EQI instrument is tested in the same three inquiry units as used in the first research cycle. From the reflection on the results it can be concluded that the students used the EQI instrument as it was intended and can handle it in different inquiry phases and in different school science subjects. It is also proved to a sufficient level that handling the EQI instrument provides the teacher with information on the support that students need when evaluating the ARV in inquiries. Finally, use of the EQI instrument leads to an increase in adequate 'inquiry language' by students related to the evaluation of ARV in inquiries. Hence, it is plausible that the EQI instrument supports pre-university science students in evaluating the ARV in inquiries.

Chapter 5 describes the design characteristics of a teaching-learning process of different inquiries in which the students learn to evaluate the ARV in inquiries by using the EQI instrument. Eleven hypothetical design characteristics were identified from the literature and elaborated in three inquiry units in general science, biology and physics. In all inquiry units, the students had to evaluate the ARV in their own inquiries several times in different inquiry phases by means of the EQI instrument. The reflection focused on three intended formative feedback functions of the EQI instrument in the teaching-learning process: (1) self-evaluation of ARV by students, (2) support of evaluation by peers and teachers, and (3) facilitating of transfer from one inquiry to another. The designed teaching-learning process supported the fulfilment of functions 1 and 2. Function 3, regarding the facilitating of transfer, is fulfilled to some extent: students recognised analogies between the different inquiries and between scientific research and their own inquiries, but they did not sufficiently apply their knowledge on differences in evaluating ARV in different inquiries. The reflection has led to some revisions and additions of the hypothetical design characteristics for a (more) appropriate teaching-learning process.

saskia_proefschrift_m_summary.indd 231 27-10-16 10:48

Chapter 6 focuses on the student learning outcomes (effectiveness) of the previous tested teaching-learning process in which rubrics from the EQI instrument were used to evaluate ARV in inquiries in different school science subjects. In this study, pre-university science students performed the three previously mentioned inquiry units followed by two assessment tasks on chemical inquiries. The data analysis showed that students perceived the EQI instrument as relevant in evaluating the ARV in inquiries, gained sufficient knowledge about CoE and could flexibly apply, with the support of peers and teacher, these CoE in a new inquiry. Three inquiry units seemed to be insufficient to reach the flexible application of the evaluation of ARV in inquiries without guidance of peers and teacher. Further research is needed on the effect of the EQI instrument regarding transfer of the evaluation of ARV from one inquiry to another.

Chapter 7 contains a reflection on the outcomes of the five studies to answer the main research question. The designed EQI instrument, based on four design characteristics, can be used by pre-university science students in evaluating the ARV in inquiries in different school science subjects. The design characteristics about the teaching-learning process have led to a supportive process for learning to evaluate the ARV in the separate inquiries. For the EQI instrument and the teaching-learning process, some indications were found that both were supportive for the transfer of the evaluation of ARV from one inquiry to another but this could not be fully proven by the evidence from the empirical studies. From the students' learning outcomes it appeared that they were able to transfer the evaluation of ARV to a certain extent when supported by peers and teacher when using the EQI instrument. Improving the learning outcomes in evaluating the ARV in an inquiry with 'only' the support of the EQI instrument seemed to be difficult for some students. It can be expected that a teaching-learning process of three (large) inquiry units is too short for novices to learn to evaluate ARV in all inquiries without any support but the EQI instrument. Some limitations of the studies are the restricted number of participating students and the limited attention paid to the role of the teacher in the teaching-learning process. It also is not analysed whether the designed inquiry and assessment tasks had the same level of difficulty in applying the EQI instrument for evaluating the ARV in inquiries. The studies suggest several implications for research on pre-university science education, and for pre-university science education itself. It was learned how the CoE from the CoE model can be used as content for a self-evaluation instrument for evaluating the ARV in inquiries in different school science subjects. To fully support students in learning to evaluate the ARV, such a self-evaluation instrument for novices should be composed of tools with different functions: an overview tool, a checklist to check whether 'everything' has been done and rubrics for evaluating the ARV. The SOLO taxonomy proved to be a feasible taxonomy to describe the different levels of performance in the rubrics, especially to show to novices what their actual

saskia_proefschrift_m_summary.indd 232 27-10-16 10:48

level of performance is and where they can head to. To make the general descriptions in the rubrics feasible for as many students as possible, they should be accompanied by working examples from a previously performed students' inquiries. The teaching-learning process needs sufficient tasks to teach students as novices how to handle the self-evaluation instrument in different inquiries, by analysing the ARV in scientific research and two-by-two peer discussions and teacher's feedback about one's own inquiries. As a consequence of the new formal curriculum for school science subjects in the Netherlands, pre-university science students and their teachers should acquire procedural understanding about the meaning of ARV in inquiries and, more particularly, how to evaluate the ARV in inquiries in different school science subjects. The EQI instrument and teaching-learning process as developed in this research can be supportive to reach this aim. Although teaching the evaluation of ARV to science teachers was beyond the scope of this research, the studies provided insight into elements of a professional development programme in which science teachers can be trained on this aspect of learning to inquire.

saskia_proefschrift_m_summary.indd 233 27-10-16 10:48