Evaluation Criteria for Inquiry-based Biology Teaching

Kirsten Schlüter[1], Kathrin Bylebyl[2], Katharina Freund[2], Stefan Nessler[1]

[1] University of Cologne, Institute for Biology Education, Gronewaldstr. 2, D-50931 Köln, Germany

[2] University of Siegen, Institute for Biology and its Didactics, Adolf-Reichweinstr. 2, D-507068 Siegen, Germany

Abstract

In 2004, new national education standards for biology teaching were adopted in Germany. Since then, inquiry-based learning has got a higher priority in biology lessons. However, in Germany there are only a few teaching examples for this subject and there are even less suggestions how to evaluate students' achievements. In this article we want to refer to this deficiency and present an example how to assess students' contributions as well as general evaluation criteria for inquiry learning.

Key-words: inquiry teaching, evaluation, assessment, biology

Introduction

In 2004, the standing conference of the Ministers of Education and Cultural Affairs of the "Länder" in the Federal Republic of Germany (KMK), passed the national education standards which refer to the subject biology (KMK 2004; Elster 2009). In this resolution the KMK demands to promote four central competences. These competences can be divided in two subgroups. Group one covers content and concept-related competences. In other words, professional competence should be promoted. Students should acquire subject knowledge and be able to apply this. The second group refers to process-related competences which are (i) inquiry–based learning competence, (ii) communication competence and (iii) valuing and decision making competence. This article concentrates on the inquiry-based learning competence. This competence consists of three dimensions (Bayrhuber et al., 2007). The first one is a lab work. It relates to biological work methods such as to observe a phenomenon, to compare objects, and to use biological implements like a microscope. The second dimension comprises the phases of scientific inquiry and the third one regards the nature of science. The focus of this article is the second dimension: the phases of scientific inquiry. The different phases are: to formulate hypotheses, to make predictions, to plan experiments, to perform the experiments, to gather data, to present them and finally to interpret these data. Subsequent to this, the results have to be discussed. By this, communication competence is promoted, too. Communication competence is also promoted by presenting the results and by transforming the data into a graphic. Finally, the results have to be reflected. By this, valuing and decision making competence is fostered.

Aim

In this article, one teaching example will be outlined. There are various teaching units on inquiry learning in the English literature (e.g. Barell, 2007; Bass, Contant & Carin, 2009; Bentley, Ebert II, & Ebert, 2007; Hammerman, 2006; Koch, 2005, Goldsworthy, Sams, Smith, Watson & Wood-Robinson, 2000). In Germany this approach has been applied in chemistry and physics teaching (Schmidkunz &

Lindemann, 1976; Fries & Rosenberger, 1970). However, only very few examples exist for biology lessons (Mayer & Krüger, 2006; Mayer & Ziemek, 2006). This is the reason why we have developed additional teaching examples (Bylebyl, Freund, Nessler & Schlüter, 2010). One of these is presented here. We want to describe potential students' activities in order to give teachers an impression what might happen during inquiry lessons. Another aspect taken into account is how to evaluate students' achievements. This latter aspect will be a focal point of the teaching example presented here.

Teaching unit: the water strider

The **task** given to the students is: "How has the water strider to be built in order to stand on the water? Develop a model for this and test experimentally if it is suitable." By this task the following **teaching aims** should be achieved. *Content-oriented teaching aims* refer to the characteristics of the water strider. Students should notice that it stands on the water, and that it uses the surface tension of the water for this purpose. Due to the fact that it is difficult for the students to work out the latter point on their own, this knowledge can be imparted by the teacher. In addition, the students should observe how the legs of the water strider are arranged, and figure out how the structure of legs and feet prevent the water strider from sinking. The *competence-oriented teaching aims* comprise that the students get to know the phases of the inquiry process. This is possible by reflecting the sequence of students' activities during the process of model-building and -testing. Moreover, students should develop criteria by which they can judge the quality of their models and the quality of the model-building process.


Fig. 1a/b: Water strider (Fam. Gerridae) (photographs by Sabine Nöbel)

As an **introduction** to the topic, the teacher can make an excursion to a pond where one can find water striders or the teacher brings the animals to class or shows pictures or films of the animals. The goal is that the students get to know the appearance and the movement of the water strider. The students should observe that the water strider has a slim, longish body and that it has six legs. Furthermore, they should notice that the second and third leg pair are attached at the sides of the body, that these legs have an x-shaped arrangement and that the feet have large contact areas with the water. It would be helpful for the students if the teacher presents additional information on not obvious mechanisms which improve the water strider's ability to stand on the water. One mechanism is that water striders have a hair felt at feet and legs. Between these hairs air is included so it seems as if the water strider walks on an air bed. Moreover, this hair felt is water-repellent because the water strider greases it.

Next, students get a box with several **materials** inside which they can use for building their models. These materials are cork, skewers, florist wire, modelling clay, flower moss, polystyrene, stones, balloons, drinking straws, pipe cleaners, ... Some of these materials are useful for model building, others are not. Thus, the students have to decide which materials to use and which not. In addition, scissors, different kinds of glues, tape, pliers to cut the wire, a knife and finally a water tub to test the water strider models are provided.

In an ideal situation, students would formulate **hypotheses** and deduce **predictions** before they start with the model-building. A *first category of hypotheses refers to the choice of material*. One hypothesis says that "one should use floating material for the feet because the feet carry the water strider and enable its body to be above the water surface". From this hypothesis the following prediction can be deduced: "If one uses floating feet material, the model stands on the water. If the material for the feet is not floating, the model will sink." Another hypothesis concerns the body: "The body should be of light material in order to reduce the weight of the water strider". The corresponding prediction is: "The lighter the body weight, the better the model stands on the water". The *next group of hypotheses concerns* the *structure of the body or parts of the body*. Regarding the feet it is assumed that sufficient large contact areas lead to a better weight distribution. The prediction is: "The bigger the feet, the longer the model stands on the water". Concerning the legs, the students can suppose that if the legs are attached at the side of the body, the weight of the model will be distributed in a better way. The deduced prediction is: "If the legs are laterally attached, the model stands longer on the water surface. In contrast, if the legs are under the body, the model will sink faster." Another hypothesis might be that the second and third leg pair should have an x-shaped arrangement for a better weight distribution. It can be predicted that an x-shaped leg arrangement will lead to a model which stands longer on the water surface. Straight leg arrangement, however, will make the model sink faster.

The next phases of the inquiry process relate to **planning** different **models**, **building** and **testing** them as well as **documenting** the **results**. These phases will not be described in detail at this point because they will be revisited in the following paragraphs referring to the evaluation of the modelling process. In terms of documenting the models this can be done either by photographs, by drawings, or by a description of the chosen materials and their arrangement. When testing the models, the standing times on the water should be listed in a table.

Analysing the quality of the phases of the inquiry process

In order to analyse the quality of the inquiry process the students passed through, one can refer to problems which might occur during the different phases of the modelling process.

Concerning the **hypotheses,** four different types can be distinguished. First, there are *good and practicable hypotheses*. Good means that the hypotheses are justified and reasonable. Practical means that the hypotheses can be tested in the classroom. The second category relates to *good but not practical hypotheses*. This means they cannot be proved in biology lessons because special equipment is needed. The third category covers *unscientific or shortened hypotheses*. *Shortened* means that an explanation for the hypotheses is missing. An example would be the statement that there is a relation between the size of the water strider and its ability to stand on the water. Why such a relation is reasonable (or not) is not explicated. An *unscientific* hypothesis is one which cannot be proved because it is not testable at all or cannot be tested repeatedly by different persons. An example for this is the statement that the water strider stays on the water surface due to its inner will. In addition, unscientific hypotheses will not allow to deduce relations between the variables. Thus, there is no "if-then"- logic and no "the more-the…"- relation. In other words, no predictions concerning the experimental outcomes can be done. The last type of *hypotheses* has *no direct relation to the question*. An example would be the statement: "A water strider must have wings because it is an insect". This characteristic of the water strider will not affect its ability to stand on the water surface. Hence, it is of no relevance.

Referring to "**planning of experiments**" the following problems could occur: the students are *not sure about the issue which has to be tested*, therefore they are uncertain concerning the dependent and independent variables, and cannot make clear predictions. In the case of the water strider, it might be that the students build a model that looks most similar to the real living organism. Nevertheless, they do not test the model because they do not care about its functionality (dependent variable). Thus, they do not take into account the density of the used materials (independent variable). Another problem might be that the *students do not consider control experiments*. This happens when students construct and compare models, which differ in more than one variable, e.g. two models consist of different bodies as well as different legs. If one model stands better on the water surface than the other one, the reason for this result (whether it is caused by the material of the body, of the legs or because of both) cannot be concluded. Even if two models exist whose bodies consist of different substances while the legs and feet are made of equal material, a comparison might be difficult. The reason might be that the position of the legs is not the same. In the first case the legs are under the body and in the second one the legs of the water strider go more sideways. So – here again – two variables differ. A further problem could be that students are *uncertain what to measure* when testing the models. It might be that students do not exactly know what "standing on the water" means. So a floating model is correct for them.

When **performing the experiments**, that means when testing the models, different problems could occur. One problem could be *measuring errors*. They will occur if the models are not put on the water surface in the same manner. For instance, one model might be put from above on the surface, the second one might be put at an angle on the surface, and the third model is put with pressure on the water surface so that it immerses. Another measuring error could occur if there are residues of soap in some water tubs so that the surface tension is reduced. A further issue might be that students choose an *inappropriate or wrong experimental set up*. This would be the case if students put their models on the water surface and immediately state whether the models stand on the water or not. It would be more appropriate to watch the models for a longer time and to determine the time span the models can stand on the water. Another problem might be that the *experimental conditions are not constant*. For instance, when testing the models, a student might hit against the water tub. By this, waves will occur. This wave movement will increase the probability of the model to sink.

During the **documentation and** the **presentation of** the **data**, the following problems could occur: Maybe the students do *not distinguish between important and unimportant data*. Instead of measuring one meaningful time span per model, students list three time values: the time span the body of the water strider stays above the water surface; the time span the body floats on the water; and the time span until the body sinks entirely. Only the first one of these three values is of interest. Another problem could be that the *data are not clearly presented*. Maybe the students summarise their data as a text. However, it would be better to draw a chart or to produce a graphic. Another issue is that there could be *errors or missing information in the data presentation*. When preparing a graphic it might be that the axes of the coordinate system are not labelled or students use the wrong units of measurement. Maybe the number of repetitions is not given or the number of examination objects is not listed. It is also possible that the students choose the wrong kind of graph. For instance, if students document the standing times of three different models, and connect these measured points to a curve, they have chosen a wrong presentation.

Referring to the **data interpretation**, it might be that *students only focus on selected data*, like those ones which are especially striking or which they received at last. Or they only concentrate on those data they hoped for and they predicted. Another problem might be that the *interpretation of the data is too generalized*. This would be the case if students transfer the characteristics of the body of the water strider to other animals which live near the water surface, e.g. to a common backswimmer/water boatman. It might also be that *confounding variables are not taken into consideration*. This could happen if students do not document if a colleague hits against the water tub, which leads to a faster sinking of the model. Finally, it might be that *data are interpreted by preconceived opinions*. If a group of students has three water strider models of the same design and one of these models stands on the water surface and the other two are floating, it would be wrong to say the model stands on the water.

The analysis of students' research process could be very helpful to figure out its optimal structure and to draw attention to questionable phases. The problem, however, is that a teacher usually cannot observe students' work in detail. An alternative to analyse students' work would be to evaluate the models with regard to certain criteria and to evaluate the experimental protocol.

Analysing the quality of a model

When analysing the water strider-models two main aspects should be focussed on: the **appearance of the model** and **its functionality**. Both main aspects can be subdivided in several criteria. Each criterion is marked by a plus, which stands for a positive evaluation; a minus symbolising a negative one; a wave meaning it is in between and the question mark illustrates that this aspect cannot be evaluated, because there is an uncertainty. At the end, there is an overall evaluation concerning the appearance and the function of the model.

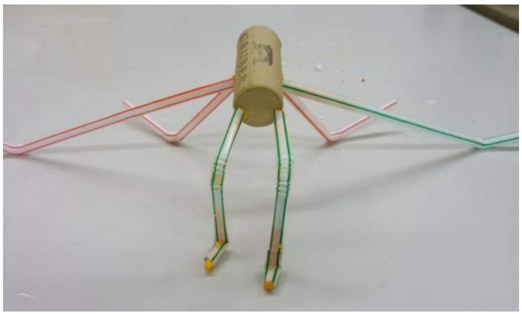

Fig. 2: Model of the water strider

Regarding the model in figure 2, it can be stated with respect to its *appearance* that the body does not touch the "ground" (+), its shape is slim and longish (+), the model has six legs (+), the second and the third leg pair are longer than the first one (+), they are in an x-shaped arrangement (+), and they are attached to the sides of the body (+). The anterior leg pair looks different (+). The feet are not oversized (+) and the relation between the body, the legs, and the feet is consistent (+). Referring to the *functionality* of the materials of the model, it is obvious that the body (made of cork) consists of light, not water-absorbent material (+). The legs are made of drinking straws, which contain wire and modelling-clay to bring them in the right position and to prevent water from entering the drinking straws. However, by these modifications the legs are not very light anymore (~). The feet do not possess large contact areas (-) and they do not have an air-cushion (-). If the model stands on the water, is not clear because there are no hints in the students' protocol whether they tested the model or not (?). The result of the *overall evaluation* is that the appearance of the model is really good (+) but not its functionality (-).

Analysing the quality of an experimental protocol/research report

In order to evaluate the experimental process, students should write a research report. Often such a report is not very detailed and does not correspond to the different phases of the inquiry process. Nevertheless, the different models which have been constructed successively in a group work are described in such a report. With the help of this report, the progress of the modelling process can be evaluated. Figure 3 shows a research report of a student group.

Research report

[Model Ia] We consider which materials we want to use. First, we took a cork and put 6 toothpicks into it. For the feet we would like to take polystyrene because it has a small density and floats on the water. We try to find materials which are light and can float on water. The first model sank because it was too heavy. We thought that polystyrene would float on the water but the cork, which we used for the body, was too heavy.

[Model Ib] Then we took bigger pieces of polystyrene. And the model could float.

[Model Ic] Besides, we had the idea to put some glue on the bottom of the polystyrene because we thought that it might float better. But there was no difference.

[Model II] Then we built a new model, whose body was a cork and whose legs consist of wire. But this model did not float.

[Model III] Another idea was to form the body of the water strider with modelling clay and to build the legs with drinking straws. But we realized that the modelling clay is too heavy and does not float on water.

[Model IV] Then we took a sponge and provided it with pipe cleaner as legs. This was too heavy and sank as well.

[Model V] Another idea was to use a balloon as the body and the pipe cleaners as legs. This model sank.

[Model VI] Someone put drinking straws into a cork. This model looked very similar to the water strider. Then she put some modelling clay into the drinking straws to preserve the shape. Then, wire was added for stabilization. At first, some of us were sceptical because wire usually is too heavy.

[Model VII] Next, we built a body with modelling clay and the legs with toothpicks. On the legs we put some pieces of polystyrene. We used modelling clay because the body of the water strider is heavy. This model floats on water. Actually, the water strider does not float on the water but he walks. Consequently, it does not have permanent contact with the water.

Fig. 3: Research report of a student group *(translation by the authors). The different models are numbered (by the authors) to facilitate the subsequent discussion of the modelling process.*

Fig. 4: Model I

Concerning **model Ia** it can be stated that the chosen materials were listed (cork for the body, toothpicks for the legs, and polystyrene for the feet) and the choice was explained. The students wanted to use materials which have a small density and float on the water. Students tested the model and explained the test result. They thought that the body was too heavy and therefore their model sank. Thus, they modified it and got to a second one, **Ib**. The students described the modification: they used bigger pieces of polystyrene as feet. The reason for this modification is given with model Ia. Again the model was tested. As a result students noted down that the model "floats".

However, the standing times were not recorded in the protocol. So the test results were not documented very precisely. Afterwards students modified the model again (**Ic**) by putting some glue on the bottom of the polystyrene. The explanation they gave is not clear because students just wrote that they expected the model to float better. They did not indicate what the effect of the glue might be. Probably it should close little holes in the polystyrene so that no water could penetrate into the material. Students tested the model but could not find any difference between the models Ic and Ib.

Referring to **model II** it can be concluded that it is a half new model. The body is the same as before. It is made of cork but the legs are different, they consist of wire. An explanation for this modification is missing. The students did not reason why they chose wire for the legs. We **interpret the lack of an explanation as an absence of hypotheses**. Due to the fact that students are not used to formulate hypotheses, we cannot expect them to do so. However, students should reason their activities. In our opinion these explanations can be equated with the presence of non-verbalised hypotheses. The testing of the model revealed that it sank. The result was not explained by the students.

Model III is completely new and there is no connection to the models before. The body consists of modelling clay and the legs are made of drinking straws. An explanation for the choice of materials is missing. The test had a negative result: the model sank. Positive is that the students explained this result. They supposed that the clay is too heavy.

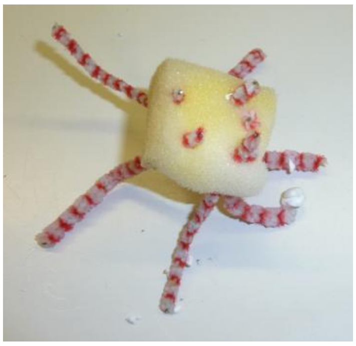

Fig. 5: Model IV

Regarding model IV, it is a new one again. A sponge is used for the body and pipe cleaners for the legs. Again the students did not explain why they chose these materials. A comparison with previous models is not possible because two variables have been changed. A test was done but the explanation is unsatisfying. Students said the model sank because it was too heavy. But that is only half the story: the model got heavier during testing because the materials absorbed water.

Model V is a half new one. In comparison to model IV the body is different (a balloon instead of a sponge) but the legs are the same (pipe cleaners). The choice of material was not explained. The test was negative and no conclusion was given, although one could easily conclude that the reason are the inadequate legs.

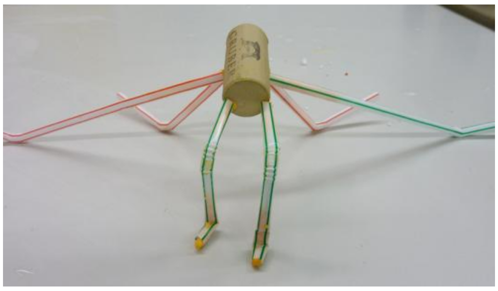

Fig. 6: Model VI

Concerning **model VI**, it can be stated that it is a combined model. The body is the same (cork) as in model Ia-c. The legs are a combination of those in model II (wire) and III (drinking straws). However, this connection was not mentioned by the students, so no pairwise comparison of these models was intended. The choice of materials was explained: the reason was to obtain visual similarity with the real living object, but the functionality or density of the materials was hardly considered. Nevertheless, there was a reflection on the choice of materials because the students mentioned that wire, which was put into the legs (made of drinking straws), would probably be too heavy. Whether a test was done or not was not mentioned in the protocol.

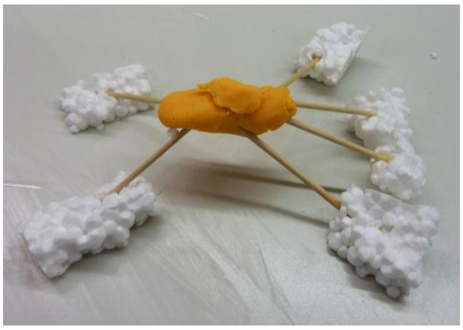

Fig. 7: Model VII

The last **model**, number **VII**, is a combination of others: it resembles the models Ia-c, as it possesses corresponding legs (toothpicks) and feet (polystyrene pieces), and the model III because of its body (modelling clay). The reference to previous models was not mentioned by the students and a direct comparison between these models was not realised. The choice of material for the body was explained: students expected the body to be heavy and therefore used modelling clay. The model was tested and met the expectations. Positive is that the students finally noticed the difference between a floating model and a model standing on the water surface while the body does not touch the water. The latter is true for model VII.

After this detailed analysis of the modelling process we want to **summarise** the **valuations**. For this we consider the following *criteria*: (i) *the construction process of the models* (choice and arrangement of materials) *is reasoned*, (ii) *the models refer to each other* so that a pairwise comparison is possible, (iii) *the results are explained*, and (iv) *the results are used for planning* further models so that a stepwise approach/improvement becomes visible.

Overall we can detect three different phases in the research protocol of the student group.

Phase	Model	Reasoning of construction process / modifications	Reference to other models	Reasoning of results	Use of result for further planning
I	Ia	+		+	+
	Ib	+	Ia	(+)[3]	+
	Ic	(+)[1]	Ib	–	–
II	II	–	(I)[2] (body)	–	–
	III	–	–	+	–
	IV	–	–	(+)[4]	–
	V	–	(IV)[2] (legs)	–	–
III	VI	+	(I, III, II)[2] (body, legs, feet)	–	–
	VII	+	(I, III)[2] (legs, body)	– / +[5]	–

Fig. 4: Evaluation of the inquiry process by analysing students' research report*. + = Positive evaluation. - = Negative evaluation. Roman numeral (no.) = Model number. [1] (+) = The modifications are explained, but not in detail. [2] (no.) = There is a reference to previous models, but students do not mention it. [3] (+) = The explanation was given before (referring to model Ia). [4] (+) = The results are explained, but not in detail. [5] -/+ = The results are not explained (-), but students noticed the difference between the terms "floating" and "standing on the water surface".*

The **first phase** comprises the models Ia, b and c. All models refer to each other. Stepwise modifications have been done in order to improve the models systematically. The modifications are reasoned predominantly. The test results are explained (as far as possible). Altogether it is a productive, well-conceived approach.

The **second phase** includes four models: number II-V. The modifications of the models are unexplained. In two cases there are no references to other models at all. If there is a reference, it is not explicitly mentioned by the students and the test results are not compared with each other. Only in one case there is a sound explanation for the test result. All in all, the second phase can be interpreted as an unproductive one. Students seem to investigate unsystematically.

In the **third phase** the models VI and VII are built. The construction plans respectively the choice of materials are explained. References to other models are not mentioned, even if they exist. In one case it is uncertain whether the model has been tested – which is a rather negative point. Concerning the interpretation of the test result of the VIIth model, it has to be positively stated that students finally noticed the difference between floating and standing on the water. Altogether, this third phase results in interesting models and indicates a structured, reasoned approach by the students – despite some points of criticism.

Conclusion

The analysis of students' experimental protocols can visualise important aspects of the research process. In the ideal situation, students develop different criteria for assessing the protocols by themselves. By comparing different research reports, students will notice the kind of information which makes the research process meaningful and comprehensible. Working out these criteria gives students an idea how research should proceed.

References

Barell, J. (2007). *Problem-based learning. An inquiry approach*. California: Corwin Press.

Bass, J.E., Contant T.L., & Carin A.A. (2009). *Teaching Science as Inquiry*. 11th edition. Boston: Pearson.

Bayrhuber, H., Bögeholz, S., Eggert, S., Elster, D., Grube, C., Hössle, C., … (2007). Biologie im Kontext - Erste Forschungsergebnisse. Mathematischer und Naturwissenschaftlicher Unterricht, 60(5), 304–313.

Bentley, M.L., Ebert II, E.S., & Ebert, C. (2007). Teaching constructivist science K-8. Nurturing Natural Investigators in the standard-based classroom. California: Corwin Press.

Bylebyl, K., Freund, K., Nessler, S., & Schlüter, K. (2010). *Selbständiges wissenschaftliches Arbeiten im Biologieunterricht*. Hohengehren: Schneider Verlag.

Elster, D. (2009). Biology in Context: teacher's professional development in learning communities. *Journal of Biological Education*, *43*(2), 53–61.

Fries, Eberhard & Rosenberger, Rudi (1970). Forschender Unterricht. Ein Beitrag zur Didaktik und Methodik des mathematischen und naturwissenschaftlichen Unterrichts in allgemeinbildenden Schulen, mit besonderer Berücksichtigung der Sekundarstufe. 2. Aufl. Frankfurt a.M.: Verlag Moritz Diesterweg.

Goldsworthy, A., Sams, J., Smith, C., Watson, R., & Wood-Robinson, V. (2000). *Developing understanding in scientific enquiry*. London: King's College.

Hammerman, E. (2006) *Essentials of inquiry-based science, K-8*. Thousand Oaks, California: Corwin Press.

KMK (Kultusministerkonferenz). (2004a). *Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss: Beschluss vom 16.12.2004*. München, Neuwied: Luchterhand. Retrieved from http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Bildungsstandards-Biologie.pdf

Koch, Janice (2005). *Science stories. Science methods for elementary and middle school teachers*. 3rd edition. Boston: Houghton Mifflin Company.

Mayer, J. & Krüger, D. (2006). Forscherheft: Biologisches Forschen planen und durchführen. *Kompakt Unterricht Biologie*, 318.

Mayer, J. & Ziemek, H.-P. (Ed.). (2006). Offenes Experimentieren. *Unterricht Biologie*, 317.

Schmidkunz, Heinz & Lindemann, Helmut (1976). Das Forschend-entwickelnde Unterrichtsverfahren. Problemlösen im naturwissenschaftlichen Unterricht. München: Paul List Verlag KG.