Keywords: physical education, knowledge, human biology, research, empirical study

An Empirical Study Investigating Interdisciplinary Teaching of Biology and Physical Education

Abstract
This paper deals with an empirical study examining the effectiveness of interdisciplinary teaching in biology and physical education (PE) regarding the students’ growth in knowledge.

The study was conducted with 141 German sixth form students. In groups, they were taught three hours a week for a period of six weeks. In order to compare the effectiveness of regular lessons with the project, the study uses a test and control group design. This means a test group takes part in interdisciplinary teaching (IdT) lessons while a control group of peers receives subject-related lessons. After being exposed to the different kinds of lessons for a certain time, we then compared the effects on the students’ skills in a pre-/post-test. The IdT lessons are designed to complement each other so that the biology lesson supplies explanations for processes occurring in PE, whereas the PE lesson provides a visualisation of the theoretical biology content.

In this paper, first the current German curricula requirements as they relate to interdisciplinary teaching will be outlined. After this, the state of research concerning interdisciplinary teaching, methodical-didactic expectations and learning targets will be provided. Following on from this, the data investigated in the course of the study will be presented, such as the study design and measuring instruments. The results of the pre-/post-test will be described and then discussed in the conclusion. They indicate a significant growth in knowledge with the test group (IdT) when compared with the control group (F(1, 113.24)=126, p<0.01, η²_p=0.43).

Introduction
For years, interdisciplinary teaching (IdT) has been regarded as a form of education imparting cross-linked and applicable knowledge to students beyond the bounds of individual disciplines (MSWWF, 1999a/b). Because of the vague and variable definitions of IdT that are found throughout literature, this article utilizes the term IdT to represent all methods of teaching related to more than one subject (Häning, 2009; Labudde, 2003). This paper deals with the combined teaching of biology and physical education (PE).

IdT is recognised as an important teaching method and can be found in numerous current syllabi of German federal states, for instance in the curricula of Hesse (2010a/b), Lower Saxony (2007a/b; 2010), North Rhine-Westphalia (2014a/b; 2008), Saxony (2011a/b) and Baden-Württemberg (2004). Baden-Württemberg’s curriculum (2004) considers IdT an indispensable prerequisite for scientific propaedeutical issues and thus relevant for upper school teaching. Saxony (2011a/b) further strengthens IdT’s relevance by demanding that each student must engage in cross-disciplinary learning for at least two weeks in each school year, irrespective of school age.

During the last ten years, both German and international studies concerning IdT have been published. They examine different aspects of this method, such as the general availability of it in schools and how it is realized there already (e.g., Gerdes, 2001; Hodgson, Keck, Patterson & Maki, 2005; Lambert, 2005; Schwartz-Bloom, Halpin & Reiter, 2011; Stübig, Ludwig, Bosse, Gessner & Lorberg, 2006). However, these studies are unable to reveal any explicit statement about the concept’s effectiveness, especially regarding the students’ improvement of application skills. This is also confirmed by Klos (2007), who sees a lack of substantiated empirical data to support IdT. Therefore, the study presented in this paper examines IdT’s effectiveness in respect to the students’ growth in knowledge in biology and PE. It must be noted that this area has not yet been researched extensively.

Expectations Regarding Interdisciplinary Teaching
As mentioned above, different German federal states already demand IdT in their curricula and regard it as a method for helping students to acquire knowledge that is of great personal relevance to them and, consequently, easily applied to everyday situations (Müller, 2006; Schecker & Winter, 2000). Dealing with problems from everyday life situations and following a problem-oriented teaching approach results in a deepened learning experience for students. This is based on the premise that students learn more readily by means of practical experience and relevance to their environment (Stübig, Bosse & Ludwig, 2002). The learning objectives anticipated with IdT are related to three aspects, namely developing cross-linked and scientific knowledge, learning how to act and learning to transfer knowledge to other situations. Consequently,
Within the concept of IdT, the approach of situated learning has to be mentioned. This concept includes learning content that is meaningful and of personal relevance to students and their environments (Müller, 2006; Schecker & Winter, 2000). Hence, it describes learning conditions that consider various application situations within the learning process and thus try to prevent passive knowledge. Passive knowledge arises because of differences in the situation of learning and the situation of application where new content should be used (Hartinger & Mörtl-Hafizovic, 2009). Following the strictly constructivist theory, knowledge has to be created by an individual actively and is bound to the situation in which it is acquired. As there is no such knowledge to be acquired in one situation that can easily be applied to a different situation, transferring knowledge to different contexts is possible to a limited extent only (Klauer, 2006). This, however, raises the question of how learning situations should be designed in order to enable a certain degree of behavioural continuity across different situations to support the students’ transfer and application skills. Therefore, the features of situated learning will be outlined in the following paragraph.

A central characteristic of situated learning conditions is the attempt to establish authenticity in the learning situation. Since knowledge is bound to a situation and cannot be transferred to a different setting easily, the situation in which knowledge acquisition takes place has to be designed analogous to other situations. However, providing students with authentic learning situations is not enough to prevent them from developing inert knowledge. Teachers will not be able to integrate every single potential application situation into their teaching design (Hartinger & Mörtl-Hafizovic, 2009). Nonetheless, students should be given the chance to look at learning content from different perspectives in order to learn how to apply knowledge practically. Further, it should be aimed at articulation and reflection phases. During the learning process conversations provide the opportunity to correlate learning processes and results. Moreover, practical knowledge can be supported by making students aware of other relevant perspectives and situations. Not mentioning these possible applications during learning can lead to inert knowledge and the students’ lack of transfer skills. Additionally, social contexts are to be created which originate from cooperative and interactive learning forms. Both negotiation processes and discussions occurring between learners in social contexts have proven to exert a positive effect on the learning process since students engage more intensively in lessons. Further benefits of social contexts are the promotion of generic and social competences such as communication skills and the development of a higher frustration tolerance. It should be noted that in order to support their students’ learning processes best and prevent inert knowledge, teachers should design their lessons so that they promote active and self-regulated learning, which means processes that can be influenced by the learners themselves (Kasztner, 2009). Generally, teachers should consider potential situations of application in advance when designing lessons. If possible, learning and application situations should correspond with each other (Renkl, 2010).

The study presented in this paper examines whether IdT increases the students’ growth in knowledge and application skills. It is based on a teaching project employing IdT in biology and PE lessons. The IdT lessons are designed to complement each other so that the biology lesson supplies explanations for processes occurring in PE, whereas the PE lesson provides a visualisation of the theoretical biology content. In combination the IdT lessons offer students a physical experience and a critical reflection on what has been learned theoretically and practically.

**Study Design**

The study at hand was conducted with six courses and 141 students. In groups, they were taught three hours a week for a period of six weeks. In order to compare the effectiveness of regular lessons with the project, the study uses a test and control group design. This means a test group takes part in IdT lessons while a control group of peers receives subject-related lessons. After being exposed to the different kinds of lessons for a certain time, we compared the lessons’ effects on the students’ skills in a pre-/post-test. This creates a quasi-experimental design since the students were not assigned to the groups by chance. A strictly experimental design would have caused the disadvantage of lowered external validity because of the test situation’s artificiality, aside from pragmatic problems like organizing different timetables.

Seventy-five of a total of 141 participating students belonged to the test group, whereas 66 students were assigned to the control group. The test and control groups contained three courses of students each. The students were allocated to the groups by means of their PE classes since different PE classes of two Bielefeld grammar schools took part in the study. Attention was paid to the fact that the students were of a similar background and the parents’ education. University students who were already experienced in teaching supervised the study. In order to prevent as many confounding variables (e.g., teaching style, experience, etc.) as possible, the following measures were taken beforehand. Prior to this study, university students were provided with an extensive introduction to the project during a block seminar. Furthermore, they were given detailed lesson plans and materials, which allowed for a precise imagination of the intended teaching process. This ensured the lessons’ contents were discussed and planned in detail and the expected teacher behavior was transparent. An additional clarification phase before every lesson helped the university students to focus on what to do during the lessons. A teaching university
student as well as an observing university student and the doctoral candidate conducting the study were present during the lessons. The observing person made sure the lesson was realized as planned and, if necessary, supported the teaching student. Every detail of the lesson was recorded by the doctoral candidate so that deviations could be reflected on retrospect.

With regard to the test’s validity check (Shadish, Cook & Campbell, 2002; Cook & Campbell, 1979), several hazards can be considered irrelevant for this study. Confounding impacts occurring randomly could have influenced the dependent variable, meaning the test score. Those impacts might be personal interest, previous knowledge or a decrease in performance due to personal reasons that individually or in sum could have altered the results. Since significance tests were carried out and proved the opposite, those concerns can be discarded.

A great number of significance tests make for a meaningful interpretation of random results. This is not the case in this study as a pilot study was conducted to check the tests, the content and the course of action and yielded results of the same significance as could be observed in the main study. Also, a follow-up study produced correspondingly significant results (still to be published).

Regarding the internal validity sensu stricto, interim events have to be taken into consideration. They could possibly have led to a change in the independent variable (test and control group) that cannot be attributed to the treatment itself, but rather to a structural incidence such as a project day on the study’s content concerning all the participants. Due to the great number of students, this event seems very unlikely, even more so because it would have needed to have happened across various schools and also, nothing comparable was announced to either the school or the conductors of the study, which is why the internal validity can still be presumed.

In the context of external validity and specifically in terms of construct validity, the treatment was only measured on the basis of one operationalization. Knowledge was investigated with the help of a knowledge test on a quantitative level and further enriched by consultations between teacher and student on a qualitative level, which revealed the students’ deep understanding of the content. Also, filming took place throughout the whole study, recording any student remarks. The comments confirm the scores of the knowledge test and the findings of the evaluation talks that were held at the end of the endurance study.

Instead of using a variety of methods, this study addresses one method only, namely the quantitative element of the knowledge test. The qualitative interviews with the students that were also filmed could not be evaluated quantitatively due to financial and time reasons. The study at hand has to be rated as a complex treatment, meaning that there might be interdependencies between several factors, such as student-teacher relation, the teacher personality, the students’ motivation and condition. Nonetheless, wherever possible, measures were taken to eliminate distorting impacts.

Regarding the external validity sensu stricto, the interdependency of environment and treatment has to be surveyed. The question is whether the results found in a certain environment can be transferred to another, thus addressing the issue of sample selection. Due to selecting students from different schools to participate in the study, this argument can be invalidated. It is true that the students stem from a similar social background, but the follow-up study mentioned before took this into consideration and sampled students from a different city, providing comparable results. Again, inter-happenings between the pre- and the post-test could have influenced the findings; however, the repetition of the study proves that this aspect’s relevance can be excluded.

The teaching units’ content of both control and test group was based on the North Rhine-Westphalia syllabi for PE and biology of fifth and sixth form classes. The thematic emphasis was put on topics of human biology, such as muscle build-up and contraction, the cardiovascular system, nutrition, training guidelines plus strength and endurance training (MSWWF 1999a/b). The following table (see Table 1) provides an overview of the biological and sportive contents the teaching units dealt with and explains how these contents were combined. The exact procedure is depicted in detail afterwards.

Whereas the test group was taught in cross-disciplinary lessons that intensified the transfer between the two subjects, the control group dealt with the topics separately in PE and biology lessons. The test group received interdisciplinary block-classes (three lessons a week) so that biological and sportive contents could be connected better. The lessons combined theory and practice with each other and took place in the fitness studio. The teaching units were divided into a short theoretical sequence functioning as an introduction followed by a practical phase to put the theories into practice. During the latter, the scientific input was applied to physical perceptions, observed and reflected upon. In this way the students turned the observation object into their own and discussed their observations in order to improve the training units they were to follow. The practical units further comprised different running forms such as interval training and endurance runs as examples of the duration method. Special attention was paid to the endurance run which was carried out in groups of four to five students. After the teacher provided the students with theoretical input about the project, the students organized their endurance training on their own. Within six weeks the students were to plan and implement twelve training units in running groups of four both during and after school. With the help of a training log, the students created individual training plans and set goals for themselves on which they reflected after each training unit. Moreover, the test group’s students were required to run cooperatively which means the group had to make arrangements in order not to over- or under-challenge its group members since the groups were composed heterogeneously. Hence, one of the project’s main goals was to support the students’ sense of cooperative skills via collective training.
<table>
<thead>
<tr>
<th>theoretical component</th>
<th>practical component</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>What happens inside a muscle during training?</em></td>
<td><em>The first guided training</em></td>
</tr>
<tr>
<td>Rules for training</td>
<td><em>The first autonomous training</em></td>
</tr>
<tr>
<td>Setting the training focus (training plan)</td>
<td>Introduction of the bicycle</td>
</tr>
<tr>
<td>Recording the performance in a training log (reflection)</td>
<td>Individual training time in groups with a bicycle — adjustment of running pace and usage of the bicycle</td>
</tr>
<tr>
<td><em>How does a muscle contract?</em></td>
<td><em>How does the training have to look in order to be successful?</em></td>
</tr>
<tr>
<td>The role of ATP in muscle contraction</td>
<td>Training principles and the setting of training stimuli</td>
</tr>
<tr>
<td>Recording the progress in the training log (reflection)</td>
<td>Recording one’s aim and the respective training principles (number of sets and repeats) in the training log (reflection)</td>
</tr>
<tr>
<td><em>How can the energy supply for the muscles be ensured?</em></td>
<td><em>Which exercises are helpful for training?</em></td>
</tr>
<tr>
<td>Purposeful setting of training stimuli as a necessary consequence of the retrieval of resources: types of energy supply, super compensation</td>
<td>Endurance and interval training — training methods</td>
</tr>
<tr>
<td>Consultation and support of the training by the teacher (at least one counselling session per group and student throughout the teaching unit)</td>
<td>Consultation and support of the training by the teacher</td>
</tr>
<tr>
<td>Recording the progress in the training log (reflection)</td>
<td>Establishing a relation between different training methods in one’s own training</td>
</tr>
<tr>
<td><em>Which exercises are helpful for training?</em></td>
<td>Recording the progress in the training log (reflection)</td>
</tr>
<tr>
<td>Endurance and interval training — training methods</td>
<td>Recording the progress in the training log (reflection)</td>
</tr>
<tr>
<td>Consultation and support of the training by the teacher</td>
<td><em>What happens when training too much? — sore muscles and its consequences</em></td>
</tr>
<tr>
<td>Establishing a relation between different training methods in one’s own training</td>
<td>Consultation and support of the training by the teacher</td>
</tr>
<tr>
<td>Recording the progress in the training log (reflection)</td>
<td>As appropriate, test on the acquired knowledge, feedback on teaching unit</td>
</tr>
<tr>
<td><em>What happens when training too much? — sore muscles and its consequences</em></td>
<td><em>What happens when training too much? — sore muscles and its consequences</em></td>
</tr>
<tr>
<td>Consultation and support of the training by the teacher</td>
<td>Consultation and support of the training by the teacher</td>
</tr>
<tr>
<td>As appropriate, test on the acquired knowledge, feedback on teaching unit</td>
<td>As appropriate, test on the acquired knowledge, feedback on teaching unit</td>
</tr>
</tbody>
</table>

They were to coordinate their training in order to manage running five and a half kilometres in the shortest time possible – as a group. In order to prevent physical overexertion and frustration for untrained students, every group had one bicycle (and one helmet for safety reasons) during the training units. The bicycle could be used by an exhausted student while the others continued running. During each running unit the bicycle was to be switched from student to student so that one always had the chance to gather new strength. The advantage of using the bike was that group members were not disrupted in their running flow\(^1\). In general, the students needed good stamina and had to communicate with each other in order to decide who was to use the bicycle to collectively run the route as quickly as possible.

The control group, on the other hand, was taught a double lesson of PE with the topic of endurance training as well as one single period of biology every week. Both treatment and control groups received lessons on endurance training so they dealt with the same sportive content. However, the control group was taught PE and biology lessons separately from each other in order not to link the subjects. To sum it up, both groups were exposed to the same problems and exercises during the PE lessons, but whereas the test group linked the PE content to biology directly, the control group did not. Their biology and PE lessons were completely separated from each other.

The school’s teachers taught the biology lessons after close consultations with the doctoral candidate to make sure the same content was taught to both groups.

---

1 For further information on the study’s biological and sportive concept study please read Wegner, Spintzyk & Gröben, 2011 and Wegner, Spintzyk & Gröben, 2013.
knowledge in the test and also transfer the newly learned content to novel tasks. The test has been performed in a preliminary study with 23 pupils and has been adjusted in terms of its reliability and selectivity. The knowledge test’s reliability amounted to Cronbach’s alpha of $\alpha = 0.710$ at the time of measurement $t_2$. The time of measurement $t_1$ was not considered in this test since it was to check the students’ pre-knowledge. Regarding the separation efficiency, the questions achieved results of $>0.2$, so a sufficient mutuality between the questions and the entire knowledge test can be anticipated.

Figure 1 provides some example questions of pre-post-test closed, semi-open and open questions:

Additionally, the students were interviewed repeatedly throughout the project in order to find out about their attitude towards the project on a qualitative level.

**Hypothesis**

In order to compare the test and control group with each other, these groups have to be representative with regards to specific characteristics. This means that all participating students have to be German grammar school students of that certain age and they should also be homogeneous concerning as many other significant variables as possible (Bortz & Döring, 2006). Thus, the study at hand uses the variables type of school, age and the factor performance level.

It is assumed that the test group’s students will have a greater increase in knowledge than the students in the control group, thus the groups’ development at the point of measurement $t_2$ has to differ from each other. This assumption is based on the explanatory statements listed in the introduction for a particular application-related growth in knowledge through IdT (KMK, 2004; MSWWF, 1999a/b; Müller, 2006, Schecker & Winter, 2000; Stübig et al., 2002). Based on this, the hypothesis is as follows:

**The test and control groups undergo different developments in their performance from $t_1$ to $t_2$. The test groups’ increase of performance will be higher than the control groups’ growth in knowledge.**

**Results**

The results of the study, driven by the proposed hypothesis, are presented in the following part. For the calculation of the results, the program SPSS 21 was used and it was tested on a 5% significance level. Functioning as the dependent variable was the total score of the knowledge test, whereas the independent variable was set by either the test or the control group. Further, the hypothesis was evaluated by means of an analysis of variance (ANOVA) with repeated measurements.

In order to check the preconditions for the conduction of the knowledge test, both the normal distribution and the homogeneity of variance was calculated for the test and control group. Normal distribution is given for the test group at both $t_1$ and $t_2$ (p>0.05), but not for the control group (p<0.05). Mauchly’s sphericity test was used to validate the ANOVA for checking the hypothesis. It could be shown that homogeneity of variance was not fulfilled ($F(1.126)=113.24$, p<0.05).

As Stevens (1999), among others, has pointed out, those irregularities are not to be assessed impedimental as the test is quite resilient to violations of that kind.

At the beginning, the test and control group differed only slightly in their descriptive values, with the test group $N=64$ achieving a score of $M = 6.67$, $SD = 2.28$ and the control group $N=64$ reaching a score of $M = 6$, $SD = 2.50$. At the time of measurement $t_2$, the test group $N=64$ achieved better average values ($M = 12.42$, $SD = 4.2$) than the control group $N=64$ ($M = 6.31$, $SD = 2.64$).
Whether the development of the test and control group differs significantly from each other was evaluated by means of the variance analysis with repeated measurements. Every one of the 64 students of the test and control group was included in this calculation. The evaluation showed a high significance for the time of measurement (\(F(1, 113.24)=126, p<0.01\)) and the treatment (\(F(1, 94.12)=126, p<0.01\)) in favor of the test group. The interaction effect of this significant developmental differential between the test and control group is high \(\eta^2_p=0.43\) (Bortz & Döring 2006, 606), so the treatment can be considered an important influence on the development of the students’ performance (see Figure 2).

**Discussion**

Not only do the findings reveal that the test group’s performance developed significantly from measurement point of time \(t_1\) to \(t_2\) but also that the students receiving IdT lessons achieved substantially higher values regarding their performance compared with the control group at \(t_2\) (\(\eta^2_p=0.43\)). Hence, the test group developed much better concerning their growth in knowledge than the control group.

Regarding the knowledge test’s overall results, the test group showed an enhanced ability to retrieve and apply the newly acquired knowledge at \(t_2\) when compared to the control group. Both groups dealt with the same learning content and their lessons only differed regarding the teaching methods to which the students were exposed.

Further, the interviews throughout the lessons have shown that the students of the test group were open-minded about the study and very motivated to organize their endurance training and reflect upon it. They displayed considerably higher interest in the topic of the project and were ready to engage with it theoretically and to carry it out in appropriate practice.

Hence, the test group’s high performance development and growth in knowledge is perhaps explained by the fact that IdT lessons are designed to support situated learning. According to the strictly constructivist theory, knowledge is bound to the situation in which it is acquired and only applicable in a limited number of similar situations. IdT lessons support the students’ practical knowledge by making them aware of other relevant perspectives and situations of application in class and providing them with authentic lessons that again influence the students’ commitment and interest in the matter.

Generally, the findings support the thesis that students taught in an interdisciplinary manner possess enhanced knowledge in comparison with schoolchildren taught in subject-related lessons (Fußangel & Gräsel, 2008; Stübig et al., 2002). Studies, such as those by Clary & Wandersee (2007) or Schwartz-Bloom et al. (2011), confirm the assumption of the study.

**Conclusion**

This study’s objective was to investigate IdT as regards its effectiveness in promoting applicable knowledge. The basis for this was established by a thorough literature review as well as current German curricula. An evaluation of the data collected during the project presented a positive picture of IdT’s effectiveness concerning the students’ performance levels and their growth in knowledge. Before the intervention’s commencement, the sampling was homogeneous in its specific characteristics. After the intervention then, the sampling displayed a significant heterogeneity as regards performance characteristics. Students in the test group attained a significant increase in knowledge growth whereas the control group did not. The reason for this can be found, amongst others, in the networking of different topics and subjects and the situated learning conditions that were lacking in subject-related teaching.

In accordance with the discussed test group’s performance, the control group’s weak development is to be explained by the following. Following the strictest form of the constructivist theory, knowledge is bound to the situation in which it is acquired, in this case biology lessons. Since subject-related lessons usually do not support the students’ practical knowledge and do not include possible application situations, they lead to the control group’s lack of transfer skills and hence inert knowledge. This is why...
the control group can only apply knowledge to a limited number of situations and is not able to transfer the isolated subject-knowledge from biology to the new situation of endurance training in PE lessons.

The test group’s better performance development can be explained by the situated learning conditions of IdT, for this teaching method has proven to be more effective regarding the students’ growth in knowledge than subject-related lessons. Just as assumed in theory, knowledge test reliability should be imitated situations in order to make it transferable and applicable.

It has to be mentioned that there are two aspects that should be taken into account in any follow-up study. First, the knowledge test reliability should be improved and second, the project should not be dependent on the weather, as this would simplify its planning and execution considerably. A follow-up study has already been planned and carried out, the results will follow shortly. This involves a unit in a fitness studio giving students the opportunity of greater self-determination, since they can make more individual decisions as regards their emphasis in training (strength or endurance). In conclusion, it can be said that the study’s objectives have been attained. Evidence was provided of the significant and quantifiably cognitive effectiveness of IdT, which is superior to subject-related teaching in the areas of applicable knowledge.

References
Niedersächsisches Kultusministerium (2007b). Kerncurriculum für die


Katharina Spintzyk, is a PhD student and a comprehensive school teacher for the subjects PE. She is writing her doctoral thesis in the Department for Didactics of Biology at Bielefeld University.

Friederike Strehlke, MA, is a trainee teacher for the subjects Philosophy and English at a secondary school and an academic assistant in the Department for Didactics of Biology at Bielefeld University. Correspondence concerning this article should be sent to: Friederike Strehlke, Jüngstraße 6, 33602 Bielefeld, Germany. Email: Friederike.Strehlke@Uni-Bielefeld.de.

Stephanie Ohlberger, is a Master student of the school subjects English and Biology. She is also a staff member in the project ‘Kolumbus-Kids’ in the Department for Didactics of Biology at Bielefeld University.

Prof. Dr. Bernd Gröben is a Professor and lecturer in the Department for Psychology and Sports Science at Bielefeld University.

Prof. Dr. Class Wegner is a senior teacher for biology and PE at a secondary school and lecturer at Bielefeld University in the Department for Didactics of Biology. He leads the project ‘Kolumbus-Kids’ and is an expert in the fields of giftedness.