Scientific Reasoning and Citizen Science: Enabling students and adults to become scientifically literate citizens of tomorrow's society

Dissertation
der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
M.Sc. Alena Rögele
aus Stuttgart

Tübingen
2021

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	10.06.2021
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter:	Prof. Dr. Christoph Randler
2. Berichterstatter:	Prof. Dr. Stephan Schwan

Table of contents

List of Figures

List of Tables

Acknowledgements

This dissertation project would not have been possible to accomplish without the ongoing support of numerous people and institutions. First, I would like to say a special thanks to my supervisors, Prof. Dr. Christoph Randler and Prof. Dr. Katharina Scheiter. You have made it possible for me do conduct my PhD in an interdisciplinary environment, during which I was able to improve my skills in the field of psychology as well as gain a lot of knowledge from interdisciplinary work with another field. Even though my PhD was marked by many setbacks and obstacles – many thanks to both of you for always working with me on a new way forward.

I would also like to take this opportunity to thank my second reviewer, Prof. Dr. Schwan. Thank you very much for stepping in as a second reviewer so spontaneously after the new doctoral regulations came into force last autumn.

Moreover, a big thank you goes to my great colleagues, Dr. Nadine Kalb, Dr. Corina Weidenauer and Naomi Staller for their support and scientific know-how.

Further thanks go to the LEAD Graduate School and Research Network and the Leibniz-WissenschaftsCampus Tübingen. Both programs offered me a unique opportunity to attend colloquiums or further education programs in order to foster my knowledge and my skills needed for conducting good and successful research. Moreover, I also came into contact with many incredibly nice other doctoral students and scientists through these programmes. It was a pleasure to get to know you all! Further, a special thanks to the Schools and Research department of LEAD for the manifold support I have received in carrying out my school studies.

Furthermore, I am very grateful for all participants who took part in the studies and thus supported my dissertation project. Many thanks to all headmasters, teachers, parents, students and others who made a successful implementation of the studies possible.

Finally, a big thank you to all my friends and family who always managed to motivate me, even when the project was at a difficult point again. Björn, thank you for your support and your patience with me, during the writing process of the dissertation.

And last but not least, a very special thank you to my sister Sarah. Thanks for helping me out with your amazing English skills, when I was in need of them, you are the best!

Zusammenfassung

Wissenschaftliches Schlussfolgern gehört zu den Kernkompetenzen, die benötigt werden, um erfolgreich in der Gesellschaft von heute und morgen teilhaben zu können (Osborne, 2013). Daher ist es essenziell, dass bereits Schülerinnen und Schüler auf diese Anforderungen vorbereitet werden und ihnen während ihrer Schulzeit die Fähigkeit zum wissenschaftlichen Denken vermittelt wird (National Research Council, 2012). Da wissenschaftliches Schlussfolgern jedoch ein gezieltes Training benötigt, stellt diese Dissertation mit Studie 1 eine neue Methode vor, bei der eine Teilnahme an einem Citizen Science Projekt in eine Unterrichtseinheit integriert wird, welche darüber hinaus theoretische Ansätze des forschungsbasierten Lernens, des Lernen innerhalb eines authentischen Kontexts und des außerschulischen Lernens kombiniert. Die Schülerinnen und Schüler übernehmen dabei authentisch die Rolle eines Wissenschaftlers. Die Ergebnisse zeigen, dass die Teilnahme an einem Citizen Science Projekts zu einer verstärkten Verbesserung des wissenschaftlichen Schlussfolgerns im Vergleich zu einer forschungsbasierten Lerneinheit ohne Teilnahme an einem Citizen Science Projekt führt. Hinsichtlich des Fachwissens und der Motivation zeigten sich keine Gruppenunterschiede.

Außerdem wird in der zweiten Studie, die in dieser Dissertation dargelegt wird, der Zusammenhang von wissenschaftlichem Schlussfolgern und der Teilnahme an einem Citizen Science Projekt bei Erwachsenen untersucht, um ein differenziertes Bild des Zusammenhangs erhalten zu können. Dabei erhielten Erwachsene im Rahmen ihrer Projektteilnahme entweder ein hohes Maß an Unterstützung oder nahmen ohne weitere Unterstützung am Projekt teil. Die Ergebnisse zeigen keine Überlegenheit der Gruppe, die viel Unterstützung erhalten hat, im wissenschaftlichen Schlussfolgern oder im Wissenstest.

Die Ergebnisse dieser beiden Studien weisen darauf hin, dass die Teilnahme an einem Citizen Science Projekt die Möglichkeit bietet, wissenschaftliches Schlussfolgern zu fördern. Um erfolgsversprechend zu sein, benötigt dies jedoch eine Integration in eine größere Unterrichtseinheit. Weitere Möglichkeiten, wie eine Projektteilnahme auch für Erwachsene vorteilhafter sein könnte, sowie die Praxistauglichkeit des Citizen Science Ansatzes werden diskutiert sowie die Möglichkeiten für zukünftige Forschung und Implikationen für die Praxis aufgezeigt.

Abstract

Scientific reasoning is an important core competency needed to successfully participate in today's and tomorrow's society (Osborne, 2013). Therefore, it is essential that students are prepared for these requirements and taught scientific reasoning skills during their school years (National Research Council, 2012).

However, because scientific reasoning requires specific training, in Study 1 of this dissertation project a new method is presented, in which project participation in a Citizen Science project is integrated into a teaching unit that also combines theoretical approaches of inquiry-based learning, situated cognition within an authentic learning context, and out-of-school learning. During this teaching unit, students authentically adopt the role of a scientist. Results show that participation in a Citizen Science project leads to increased improvements in scientific reasoning compared to an inquiry-based learning unit without participation in a Citizen Science project. No group differences were found regarding factual knowledge and motivation.

In addition, the second study presented in this dissertation project examines the relationship between scientific reasoning and participation in a Citizen Science project in adults in order to be able to obtain a differentiated picture of this connection. In this study, adults either received a high level of guidance as part of their participation in the Citizen Science project or participated in the Citizen Science project on their own, without further guidance. The results do not show any superiority of the group that received a lot of guidance concerning scientific reasoning or in a factual knowledge test.

The results of these two studies indicate that participation in a Citizen Science project offers the opportunity to promote scientific reasoning. However, to be promising, this requires integration of the Citizen Science project into a larger teaching unit. Other ways in which project participation could be more beneficial for adults, as well as the practicality of the Citizen Science approach, are discussed, along with opportunities for future research and implications for use in practice.

List of submitted manuscripts

Manuscript 1:
Rögele, A., Scheiter, K., & Randler, C. (submitted 2020). Linking out-of-school and in-classroom instruction using a Citizen Science Approach to Fostering Scientific Reasoning.

Manuscript 2:
Rögele, A., Scheiter, K., & Randler, C. (submitted 2021). Can Involvement induced by Guidance foster Scientific Reasoning and Knowledge of Participants of a Citizen Science Project?

1. Introduction

Scientific reasoning is regarded as a central competence required for successful partaking in a society of the 21st century (Osborne, 2013) and is one of the most important skills that students should acquire during their school years (National Research Council, 2012). However, education in school nowadays mainly focuses on cognitive thinking skills on a lower level, like comprehension and especially recall (Osborne, 2013). As a result, students struggle to develop scientific reasoning skills during their education at school, and thus, show low performances on assessment tests concerned with scientific reasoning (Pant et al., 2013).

Therefore, scientific reasoning skills seem to require a special training to develop, which might not be provided by traditional in-classroom instructions. The following sections outline possible reasons for the shortcoming of scientific reasoning skills in biology education at school and suggest possible solutions to overcome these problems.

1.1 Scientific Reasoning

Scientific reasoning is a complex, higher order cognitive skill, which includes "different epistemic activities (problem identification, questioning, hypothesis generation, construction of artefacts, evidence generation, evidence evaluation, drawing conclusions as well as communicating and scrutinising scientific reasoning and its results)" (Fischer et al., 2014, p. 39). It is regarded as a skill with major importance for the 21st century, as it is important for citizens to be able to evaluate and to make sense of an oversupply of information, with which they are confronted in today's society (Gilbert, 2005). Furthermore, employers of today's and tomorrow's society will be in need of workers, who have the ability to critically reflect new ideas and evidence (Hill, 2007) and to show a high problem-solving competence (National Research Council, 2008). In order to adequately prepare students for these demands of society, scientific reasoning is considered a key competency that students should acquire during their school years (National Research Council, 2012).

Yet, students seem to be more successful in gaining factual knowledge than in developing scientific reasoning skills during lessons at school. A nationwide study in Germany compared various schools in terms of the students' biology knowledge and relevant associated skills. It showed that the students in eight of 16 federal states only achieved the minimum standard in terms of scientific thinking and the understanding of patterns and hypotheses (knowledge production), instead of the recommended standard. However, concerning the recall of factual knowledge, students of all federal states achieved the recommended standard. If only German grammar schools are taken into account, the recommended standard concerning factual knowledge in biology is even exceeded by half of the federal states, whereas all federal states are only within the recommended standard in terms of knowledge production. In total, 15.5 percent of the students at grammar schools in Germany did not reach the recommended standard concerning knowledge production (Pant et al., 2013).

Given that the focus of school lessons is primarily on the recall of learned knowledge or its application (Osborne, 2013), and less on promoting critical thinking and evaluation (Ford, 2008; Ford, 2012) these results are not surprising. Hence, the development of scientific reasoning skills demands a special focus and a targeted training to overcome the obstacles that traditional in-classroom instruction may pose to fostering this ability.

1.2 Inquiry learning and situated cognition

Scientific reasoning is closely related to the steps of a scientific knowledge process. Thus, the inquiry-based learning concept, which asks students to construct knowledge by using methods like professional scientists (Keselman, 2003), can serve as a solution to the students' lack of scientific reasoning skills, since inquiry learning trains precisely this scientific way of working and the scientific knowledge gain.

Inquiry-based learning consists of several inquiry phases which form an inquiry cycle. Each of the inquiry phases represents a small unit of the complex scientific working process and is connected to its previous and following inquiry phases. The inquiry phases usually include Orientation, followed by Investigation and Conclusion and finally, Discussion (Pedaste et al., 2015). In the course of this inquiry cycle, students can generate own hypotheses and test these, for example by conducting experiments or by carrying out observations (Pedaste et al., 2012). This enables the students to be part in an authentic activity which, in this case, is a scientific discovery process (Pedaste et al., 2015). Besides improved learning success compared to traditional instruction (Furtak et al., 2012), inquiry learning also improves several inquiry skills, like formulating hypotheses, working on data and drawing conclusions

(Mäeots et al., 2008). Such skills are essential for valid scientific reasoning, as scientific reasoning can be seen as a transfer of scientific inquiry methods to tasks, in which reasoning is needed (Kuhn & Franklin, 2006).

As inquiry learning yields the largest effects if students receive an appropriate amount of guidance, teachers should help the students in the process, for example by providing a research question (Lazonder & Harmsen, 2016). In particular, for successful guidance students should be given the possibility to access relevant information concerning the inquiry topic. Moreover, teachers should help students to structure the inquiry learning process and restrict the complexity of the process (De Jong & Van Joolingen, 1998). Guidance is necessary for successful inquiry learning, as adolescents often have troubles with scientific methods and for example cannot clearly differentiate between hypotheses and predictions or only formulate very few hypotheses on their own at all (Gijlers & De Jong, 2005; Nijoo & De Jong, 1993). Therefore, students should be given guidance in order to improve their learning activities, performance success and learning outcomes (Lazonder & Harmsen, 2016).

Another positive aspect of inquiry learning is that it provides a possibility to increase student involvement, which is considered as a major factor for successful learning (Freeman et al., 2014). With adequate guidance, even students in elementary school can engage in an authentic scientific task and are able to learn and understand scientific methods (Lazoner & Harmsen, 2016).

In conclusion, inquiry learning has been shown to overcome some of the problems observed with students' scientific reasoning skills. Nevertheless, the mere use of an inquiry-based learning approach inside the classroom would still show too many differences to an authentic learning context outside of school and the students would not be able to transfer their knowledge and their skills into their daily lives, as the gap between inside and outside of school cannot be bridged without further facilitation.

Researchers in support of the situated cognition approach propose that the shortcoming of knowledge transfer and scientific reasoning might be due to the huge difference between learning at school and practical work outside the classroom (Resnick, 1987). For example, learning at school is mostly individual. Although being together as a class, the students do most of their work (e.g. homework) on their own and are judged individually in class tests as well. However, activities outside of school usually are embedded in social systems, in which cooperation is essential in order to gain achievements, as specific knowledge of single persons is often required. Moreover, schools dislike the usage of tools to improve performance and focus mainly on symbolic activities instead of working in meaningful, realistic contexts and try to teach general, universal skills, instead of situation-specific knowledge. Due to these differences between learning at school and practical work outside the classroom, students might become good learners in a school setting, but are not able to transfer and apply their knowledge to real-life situations (Resnick, 1987).

Thus, in order to improve learning success of students and to possibly also train scientific reasoning skills, researchers suggest learning to take place in an authentic context (Brown et al., 1989). Therefore, learning inside the classroom should be combined with an authentic activity outside the classroom, for example by having the students collect scientific data in the field, as a professional scientist would do. This could show similar advantages as an apprenticeship, as the cognitive demands during a field trip are the same for the students as for practitioners (Savery & Duffy, 1995), and might help the students to bridge the gap between theoretical learning at school and taking practical actions outside the classroom (Brown et al., 1989). Thereby, field trips help the students generate representations which could not simply be replaced by descriptions inside the classroom (Brown et al., 1989) and by doing so provide a context of application which enables situated cognition outside the classroom. Otherwise, students only have inert knowledge and are not able to develop useful practical knowledge (Whitehead, 1967). Thus, to improve students' scientific reasoning skills, students should work similarly to professional scientists by identifying a research question, formulating a hypothesis, collecting data in the field and analysing their results.

As mentioned before, inquiry-based learning serves as an authentic activity inside the classroom as well, as the students have the opportunity to act like professional scientists. Nevertheless, to help the students to bridge the gap between school and their everyday lives, it might be useful to combine inquiry learning with out-of-school learning.

1.3 Out-of-school learning

Looking at out-of-school learning, informal and non-formal learning are to be differentiated. Informal learning occurs spontaneously, often during the students' free time, and is not guided by a teacher, whereas non-formal learning is usually structured and pre-arranged, while often being led by a teaching person. Therefore, non-formal learning is taking place on field trips in nature (e.g., Randler et al., 2005), during a visit to the zoo (e.g., Seybold et al., 2014), a science centre (e.g., Itzek-Greulich et al., 2015) or a museum (e.g., Bamberger & Tal, 2007). Although non-formal learning shares certain characteristics with formal learning at school, such as taking place in a structured manner, it still has the ability to increase students' motivation which leads to learning that is typically intrinsically motivated. In contrast, formal learning is usually compulsory, evaluated and might even be repressive, typically resulting in extrinsic motivation (Eshach, 2007).

Motivation usually results from an individual's desire to achieve a specific goal. Therefore, a motivated action is always based on an intention (Deci & Ryan, 1985; Deci, 1992). The self-determination theory by Deci and Ryan (1985) assumes that the basic intention on which motivation is based is the satisfaction of basic, innate, psychological needs. They postulate three basic needs: competence, autonomy or self-determination, and social relatedness. For actions to be intrinsically motivated, especially competence and autonomy are crucial in strengthening self-determination (Deci, 1975). Intrinsic motivation occurs when an individual performs actions based on his or her own interest, i.e. without external impulses, such as a promised reward or a threat (Deci, 1975, 1992). Intrinsically motivated actions therefore represent a prototype of self-determined behaviour: A high degree of self-determination leads to a free choice of action, thus, the chosen action is in accordance with one's own opinion about oneself, which ultimately results in a high intrinsic motivation (Deci & Ryan, 1993). Hence, a mere high experience of competence is not sufficient for intrinsic motivation. The result of an action also has to be experienced as being caused personally, i.e. there has to be an experience of autonomy or self-determination (Ryan & Deci, 2002).

When taking a look at school learning, effective learning seems to require intrinsic motivation (Deci & Ryan, 1993). For example, Grolnick et al. (1991) have reported a positive correlation between motivation based on self-determination and learning performance. Moreover, students who received autonomy-supporting feedback in class, showed higher curiosity, more independence when solving problems, a more positive self-assessment (Deci et al., 1981) and higher knowledge gain, especially in a long-term assessment (Grolnick & Ryan, 1987).

Out-of-school learning, or especially taking students on field trips, provides the opportunity to increase students' intrinsic motivation based on self-determination. Since students can observe animals or plants in nature during a field trip, free-choice learning is partly possible. Free-choice learning is described as "the learning that individuals engage in throughout their lives when they have the opportunity to choose what, where, when and with whom, to learn" (Falk et al., 2007, p. 456). Free-choice learning leads to high self-determination, which results in a learning process driven by intrinsic motivation (Falk et al., 2007). During a field trip, students can choose more freely where and what they want to work on. This gives them the option to focus more on their personal interests than inside the classroom. Therefore, they can concentrate on parts of the topic which match their own interest and thus learn by free-choice and with a greater amount of intrinsic motivation. Falk et al. (2007) highlighted the importance of free-choice learning concerning interest in and understanding of science: Nearly half of the surveyed citizens stated that their scientific understanding was mainly achieved due to personal interest or curiosity, meaning due to free-choice of learning topics.

Furthermore, a field trip provides the possibility to make students realise the relevance of the learning topic for themselves. Since a field trip helps the students to better understand their personal environment, it leads to a higher subjective utility value of the learning topic. Eccles postulates in her expectancy-value theory that a higher subjective utility value leads to a higher level of involvement in the task, which in turn leads to an increase in interest and in knowledge gain of the students (Eccles, 1983). Especially for students with low expectations of success, higher personal utility values lead to higher motivation and better performance in science classes (Hulleman & Harackiewicz, 2009). Field trips should therefore be connected to prior knowledge and relevant topics or surroundings of the students to get them involved in the trip and the task (Eshach, 2007).

The impact of field trips on students' gain of knowledge has already been investigated by several previous studies. For example, Randler et al. (2005) compared a group of 3rd and 4th graders, who took part in an environmental conservation program, during which they were working with living amphibians to preserve them during migration, to a group which received in-classroom activity only. Results showed a greater increase in knowledge in the out-of-school group. Similar results were obtained by Seybold et al. (2014) for students of grade 5 and 6. In this study, two interventions were compared regarding their effectiveness in teaching students about primates. One intervention took place at a zoo, whereas the other intervention only contained in-classroom instruction. Students in the zoo group encountered living primates at the zoo and documented and discussed their observations of them, while the in-classroom group was working on the same topics in small groups at different workstations at school. Results showed higher gains of factual knowledge and a higher level of interest, which can be regarded as an indicator for intrinsic motivation, for the students of the zoo group, but a higher level of perceived choice, which represents the feeling of autonomy, for students of the in-classroom group. In contrast, in a different study by Wünschmann et al. (2017), a higher level of perceived choice was found for a zoo group in comparison to an in-classroom group. Equally, however, a greater increase in factual knowledge for the zoo group was also found in this study.

Especially when combined with formal learning methods, non-formal learning seems to be beneficial for learning performances. For students of grade 5 and 6, Randler et al. (2012) obtained results in favour of structured out-of-school learning. They compared four different learning methods when learning about vertebrates during a visit to the zoo, with one group receiving no instruction at all, the second group having a presentation by a teacher, the third group working in a learner-centred way with a summary by a teacher at the end, and the fourth group working learner-centred as well but with a peer-tutoring summary at the end. Results showed that students who received any kind of instruction outperformed students without any instruction concerning factual knowledge. Immediately after the teaching unit, students who received a presentation by a teacher showed highest knowledge scores, however, six weeks after the unit the learner-centred group, which received a summary by a teacher and the teacher-centred group both showed better factual knowledge compared to the peer-tutoring group. There were no significant differences between the teacher-centred group and the learner-centred group with the teacher's summary. Thus, in combination with formal instruction, a visit to the zoo is more beneficial for the students than an unstructured visit or a visit with peer-tutoring only, especially when looking at long-term effects (Randler et al., 2012).

Similar results were also obtained for university students by Pfeiffer et al. (2012) who investigated different interventions which combined an instruction part inside the classroom with a visit to an aquarium in order to improve identification of different fish species. During the preparation of the visit to the aquarium inside the classroom, the university students either worked with digital videos of the fish, with preserved specimens, or with both. Results showed that the digital-video group was better in identifying different species, yet showed lower motivational scores than the preserved specimens group. Best results were obtained, if both methods, learning with digital media and with preserved specimens, were used for the preparation of the visit to the aquarium. Hence, the implementation of digital media in a teaching unit that combines in- and out-of-classroom instructions might show even greater advantages.

However, there are also studies that do not show an advantage of non-formal learning in comparison to formal learning inside the classroom. For example, Itzek-Greulich et al. (2015) conducted a study which compared three different interventions on the topic of chemistry. One intervention took place at a science centre, the second intervention at school, and the third intervention combined learning at school and learning at a science centre. Results showed no differences regarding knowledge gain for the three intervention groups. Hence, when planning a field trip, one should keep in mind, that learning on field trips is influenced by several factors on different levels which should be taken into account to make non-formal learning effective. Reviewing literature concerning out-of-school learning, DeWitt and Storksdieck (2008) conclude, that out-of-school learning can improve learning outcomes, but only if it takes place under favourable circumstances. Eshach (2007) proposes a model for out-of-school learning that integrates personal, physical, social and instructional components, which all occur at both a cognitive and an affective level. For instance, the personal factor includes components as prior knowledge on a cognitive level and attitudes toward science or toward the environment on an affective level. Looking at the social factor, interpersonal interactions, in which knowledge is created, play a role on the cognitive level, whereas on an affective level, peers or teachers might influence the individual. Concerning the physical factor, the appearance of the surroundings during the field trip may influence students on an affective level, while the possibilities to influence or to interact with an environment play a role on a cognitive level. Finally, the instructional factor includes the highly important preparation for the field trip, on a cognitive as well as on an affective level, which is essential for the students to profit from the trip (Eshach, 2007). If these different factors are taken into account when planning a field trip, non-formal learning provides an opportunity for successful learning for students with individual strengths who can all play their own part in learning and solving tasks during the field trip (Eshach, 2007).

Yet, it often is problematic for students to connect the out-of-school learning to their everyday lives, as a connection between the field trips and the students' daily surroundings is missing in many cases. Therefore, simply taking students out on a field trip does not guarantee better learning and an improvement of students' scientific reasoning skills, as students often see the trips merely as an add-on activity and therefore do not really engage in the scientific process, neither inside the classroom, nor outside.

1.4 Citizen Science

To overcome this problem and to build up an authentic scientific process, in which all the single components are naturally connected to each other, it could be suitable to link the different learning locations by implementing the teaching unit in a Citizen Science project participation.

Citizen Science projects are designed to engage the public in scientific research. To do so, citizens who take part in the project take over the expert status of a scientist and gather data in their place in the field (Bhattacharjee, 2005). Participation in a Citizen Science project should be beneficial for both, the scientific institutions developing and coordinating the study, and the participants who volunteer to take part in the study and collect scientific data on their own. Advantages for scientific institutions include massive savings of professional scientists' working time and of money needed to conduct the study (Bonney et al., 2009). Bonney (1991) reports an example of a study concerning bird observations during which participants added over 200.000 hours of observation to the Citizen Science data base. Since the cost would have been in the millions if a scientist had collected this data, it would not have been possible to gather such a large amount of data without a Citizen Science project (Bonney, 1991). However, participating citizens can also benefit from a Citizen

Science project. Besides the possibility of contributing to scientific knowledge by means of their observations, they also gain insights into the nature and the manner of a scientific investigation and can expand their personal knowledge about the research object in a more structured and profound way than they might have done without professional guidance (Bonney et al., 2009; Trumbull et al., 2000). To ensure that partaking in a project is beneficial for participants as well, more and more Citizen Science projects are exploring not only the results of studies on the actual study topic, but also the effects that participation has on citizens. For example, the Cornell Lab of Ornithology (www.birds.cornell.edu) uses Citizen Science data to answer ornithological research questions on the one hand (Bonney et al., 2009), and on the other hand investigates outcomes of Citizen Science projects on participants (Bonney et al., 2016).

Participating in a Citizen Science project might have several positive effects on the development of scientific reasoning skills of students. Firstly, it supports the inquiry-based learning concept, as a Citizen Science project asks its participants to work like a scientist during the data collection. After finishing data collection, students could continue to work with and to analyse this data and thus continue the inquiry-based learning with real, scientific data. Secondly, participation in a Citizen Science project represents an authentic activity, as it is exactly the work of a scientist, which is delegated to laypersons (Bhattacharjee, 2005), So this participation can serve as some sort of apprenticeship and thus, enable situated cognition (Brown et al., 1989). Finally, participants of a Citizen Science project usually need to go outside to collect scientific data about the research object during their observations. This makes a Citizen Science project also suitable as a non-formal out-of-school learning activity. Furthermore, participation in a Citizen Science project offers the opportunity to link the in-school and out-of-school learning settings, as parts of the learning context, namely participation in the Citizen Science project, remain the same in both settings when the collected data is analysed in the classroom afterwards, like suggested before. Thus, transfer between in- and out-of-school knowledge should be facilitated, and the creation of inert knowledge avoided.

1.4.1 Citizen Science and digital media

Citizen Science projects and the internet are closely linked to each other (Bonney et al., 2016). Without the facilitation that the internet provides regarding sharing data and observations, Citizen Science projects would be tedious to implement. Since the collected data is usually shared on an online Citizen Science database, it is often useful for projects to be carried out directly with the usage of digital media. For the students, the use of mobile devices, like tablets, makes participation more fun, but additionally it also facilitates the transfer of knowledge between different learning contexts, as mobile devices support seamless learning. Chan et al. (2006, p. 6) define seamless learning as the possibility “that a student can learn whenever they are curious in a variety of scenarios and that they can switch from one scenario to another easily and quickly using the personal device as mediator”. Furthermore, tablets provide a portable learning environment (Chen et al., 2003) which enables the students to access learning content and other information during the field trip. Therefore, learning with mobile devices can also lead to higher intrinsic motivation, as the information can be provided on the tablet based on the learner’s own initiative. Knowledge transfer between formal learning inside the classroom and non-formal learning out on a field trip is hence simplified not only by taking part in the Citizen Science project, but also by using mobile devices.

So far, there is little research on combining out-of-school learning on animals with mobile devices (Thomas & Fellowes, 2016). Working with university students, Pfeiffer et al. (2009) reported higher knowledge gain when students were learning with digital materials how to classify fish in nature. Further research is needed to determine the chances and the obstacles of the use of mobile devices in the context of out-of-school learning and Citizen Science.

1.4.2 Effects of Citizen Science on Knowledge Gain and Scientific Reasoning

Different research groups have been able to show a gain in factual knowledge regarding the research topic in participants of Citizen Science projects. However, results concerning scientific literacy or scientific reasoning are scarce and inconsistent at the current time. For example, in a study of Crall et al. (2013) participants took part in a Citizen Science project concerned with invasive plant species. During their participation, they were asked to identify different plant species in the field, record their location with a global positioning system (GPS) and record their observations in a monitoring protocol. Results showed a modest knowledge gain concerning invasive plant species, however there were no changes in general scientific literacy of the participants. Another study about a Citizen Science project regarding invasive plants showed similar results (Jordan et al., 2011): Participants who collected data about invasive plants using a protocol and who discussed and analysed this data in a small group afterwards showed an increase in factual knowledge and in knowledge of scientific methods, yet did not improve their understanding of the way science is done. Furthermore, a Citizen Science project by Brossard et al. (2005), showed comparable results in the context of birds. In this project, participants were asked to observe nesting behaviour of birds by putting up a nest box in their backyard and afterwards documenting different information in connection with the birds' nesting behaviour, for example the size of the nest or the site on which the nest was build. There was a significant increase in factual knowledge regarding birds, yet no changes in participants' science literacy.

However, there are few studies that were able to show an increase in scientific thinking of participants: Trumbull et al. (2000) conducted a Citizen Science project with regard to food preferences of different bird species. Participants set up a feeder in their yard, which they filled in with different sorts of seeds. Afterwards, they were asked to record which bird species preferred which kinds of seeds. In a qualitative analysis of participants' feedback, Trumbull et al. (2000) reported that almost 80% of the participants engaged in scientific thinking processes. Similar results were obtained in a study concerned with birds' nesting behaviour, during which participants observed and reported nesting behaviour of different bird species close to their homes (Evans et al., 2005). Derived from the questions, which participants asked in connection with the project, Evans et al. (2005) conclude that participants, next to improving their factual knowledge, engaged in scientific thinking processes during participation.

Hence, the mere participation in a Citizen Science project might not be sufficient to foster learning and scientific reasoning skills. In schools, projects that are curriculum-based, and thus provide a higher level of structure and scaffolding, for example by being instructed by an adult, (Bonney et al., 2016) might provide the possibility to achieve a greater increase regarding factual knowledge and scientific reasoning. However, as science education of participants is one of the main goals of most of the Citizen Science projects (Bonney et al., 2016) and the ongoing effort of the National Research Council (1996) to provide possibilities for citizens to become more scientifically educated, opportunities for improvement in Citizen Science projects outside of school should not be neglected. An idea to improve participants' outcomes is discussed by Druschke and Seltzer (2012), who argue that a higher level of participants' engagement in the projects might lead to greater benefits. Possible ways to engage participants in the project might be either by a higher degree of participation, as Shirk et al. (2012) suggest, or by increasing the intensity of involvement (Wilmsen & Krishnaswamy, 2008). Both ways have been shown to have beneficial effects on projects outcomes (Shirk et al., 2012; Wilmsen & Krishnaswamy, 2008). This dissertation project investigates the possibility to stronger engage participants in the project by varying the personal involvement through different levels of guidance. Lazonder and Harmsen (2016) report positive effects of guidance for outcomes of the inquiry-based learning approach, even after a short period of time. As mentioned before, participation in a Citizen Science project can, in parts, be compared to learning in an inquiry-based manner, as the knowledge gain in Citizen Science projects results from the application of scientific methods. When an inquiry-based learning approach is pursued, an appropriate amount of guidance is essential for successful learning since inquiry without guidance is inferior to explicit instructions without inquiry learning. Hence, the success of an inquiry-based learning approach depends on the appropriate way and amount of guidance (Lazonder & Harmsen, 2016). In the context of Citizen Science, Evans et al. (2005) report that the availability and the collaboration of participants with the scientists, was essential for the success of the Citizen Science project. Having scientists available for queries and as contact persons if help is needed, can be regarded as a low level of guidance, which seemed to be important for participants' motivation in partaking in the project. Thus, it might be able to induce a higher level of motivation and engagement by providing a higher level of guidance to participants. Since these results could be transferred back to the schooling context as well, the investigation of the effects of different levels of guidance on scientific reasoning and factual knowledge of adults partaking in a Citizen Science project, might additionally point out implications for a successful training of scientific reasoning at school, which is essential to support students in becoming scientifically literate citizens in tomorrow’s society (Osborne, 2013).

2. Objective of the dissertation project

This dissertation project deals with the question how especially scientific reasoning, but also factual knowledge, can be improved in students and adults. Since scientific reasoning seems to require a special training to develop, a Citizen Science approach is pursued in this dissertation project in order to gain insight into the outcomes in the schooling context, but also on adults participating in a Citizen Science project outside of school. The following main questions are addressed in this dissertation project:

1. Is a Citizen Science approach suitable for training scientific reasoning at school?
2. Is a Citizen Science approach suitable for improving students’ factual knowledge?
3. Do students partaking in a Citizen Science project show higher levels of motivation, especially with regard to perceived choice as an indication for autonomy and thus, self-determination?
4. Does participation in a Citizen Science project improve adults’ scientific reasoning?
5. Does participation in a Citizen Science project improve adults’ factual knowledge?
6. Do different levels of guidance during a Citizen Science project lead to different outcomes regarding scientific reasoning and factual knowledge in adults?

Since scientific reasoning is an essential skill for successful partaking in today’s and tomorrow’s society (Osborne, 2013), which is, however, often still not sufficiently developed during science education at school (Pant et al., 2013), this dissertation project provides new knowledge regarding an important topic for schools and stakeholders of Citizen Science projects as well. Furthermore, as guidance in the context of Citizen Science projects has seemingly not yet been investigated, this dissertation project provides first insight into the opportunities which the combination of guidance, which is already known as a beneficial tool inside the school setting (Lazonder & Harmsen, 2016), with a Citizen Science project might possibly have for training school students' and adults' scientific reasoning.

3. Manuscript 1: Linking out-of-school and in-classroom instruction using a Citizen Science Approach to Fostering Scientific Reasoning

Author	**Author position**	**Scientific ideas %**	**Data generation %**	**Analysis & interpretation %**	**Paper writing %**
Alena Rögele	1	20	100	70	70
Katharina Scheiter	2	40	0	10	20
Christoph Randler	3	40	0	20	10
Title of paper:		Linking out-of-school and in-classroom instruction using a Citizen Science Approach to Fostering Scientific Reasoning			
Status in publication process:		Submitted in October 2020			

Please note: As this manuscript is the main part of the dissertation and contains the main findings obtained in the course of this dissertation project, parts of the following manuscript partly overlap in content and structure with the theory presented earlier (see point 1) and the discussion that follows (see point 5 below).

Abstract

Scientific reasoning is a central, but challenging competence to be achieved in science education. We aimed to improve scientific reasoning by implementing a teaching approach based on theoretical assumptions of situated cognition, inquiry learning, and out-of-school learning in a Citizen Science context. Fourteen Biology 5th and 6th grade classes (N = 345 students) first received an in-classroom lesson to the domain of study, biodiversity regarding waterfowl. Then, students in the experimental group observed waterfowl outside and recorded their observations in a Citizen Science database. The control group received further in-classroom instruction. Subsequently, both groups tested hypotheses inside the classroom using either data from the Citizen Science database (control group) or their own observations (experimental group). The experimental group showed greater improvements in scientific reasoning, even in a follow-up test; however, there were no group differences concerning factual knowledge gains or motivation. Concluding, the Citizen Science approach seems suited to overcome barriers in teaching scientific reasoning.

3.1 Introduction

Scientific reasoning, scientific argumentation and the critical evaluation of evidence are core competences students are supposed to acquire during their formal scientific education in school (National Research Council, 2012). However, according to national and international science assessments students across all age levels often fail to reason scientifically. The IQB National Assessment Study 2012 showed that students in Germany perform far worse on Biology tasks requiring scientific reasoning skills compared to tasks requiring only factual knowledge (Pant et al., 2013). Similarly, international large-scale assessments revealed that while in many countries students have good factual knowledge, they lack the ability to use this knowledge to solve scientific reasoning problems (Pant et al., 2013).

Researchers in education have argued that a lack of scientific reasoning skills goes back to at least two shortcomings in educational practice. First, students are hardly required to engage in scientific reasoning activities of their own. A stronger emphasis on inquiry learning has been proposed as a solution to this problem (Kuhn & Franklin, 2006). Second, students fail to recognize how they can use their knowledge to solve real-world challenges (Resnick, 1987). To counteract this problem, researchers have proposed to more strongly connect in-classroom instruction to real-world contexts. Recently, Citizen Science projects have been adopted in education as way to also teach citizens about science and improve their scientific reasoning skills (Cronje et al., 2011). In the present study, we adopted the Citizen Science philosophy to develop a teaching approach aimed at enhancing secondary students' scientific reasoning skills in the context of Biology education.

3.1.1 Scientific reasoning

Scientific reasoning is regarded as the skill to approach scientific problems in the way scientists do. This includes "asking scientifically oriented questions, giving priority to evidence in responding to questions, formulating explanations from evidence, connecting explanations to scientific knowledge, and communicating and justifying explanations" (National Research Council, 2000, p. 23).

During their scientific education at school, students are supposed to acquire scientific reasoning skills and learn scientific working methods (National Research Council, 2012). Yet, formal education appears to be more successful in conveying factual knowledge rather than reasoning skills. In a nationwide comparison of Germany schools on the subject of biology, students in half of the federal states did not achieve the required standard of knowledge production, which includes scientific thinking and understanding of patterns and hypotheses. In contrast, concerning factual knowledge, all federal states reached the required standard (Pant et al., 2013).

3.1.2 Teaching scientific reasoning through inquiry learning and out-of-school activities

Scientific reasoning comprises the formulation of a research question and hypothesis, including planning and implementation of the investigation, and ending with the evaluation and interpretation of the results (Mayer, 2007). Inquiry learning trains students in applying precisely this scientific way of working, which is why it is often seen as the golden route to teaching scientific reasoning skills. During inquiry learning, students are asked to test their hypotheses by means of experimentation or observation (Pedaste et al., 2012), thereby enabling them to partake in a scientific discovery process (Pedaste et al., 2015). Accordingly, inquiry learning contributes to the acquisition conceptual knowledge as well as to scientific reasoning skills (Furtak et al., 2012; Mäeots et al., 2008).

However, even when students engage in inquiry activities during classroom instruction, these activities are often detached from what they experience out of school and therefore have little personal relevance for them. To help the students to make an association to their daily lives, learning should be embedded in an authentic context (Brown et al., 1989) and furthermore, be taken into the students' natural habitat outside of school (Charney et al., 2007).

Researchers proposing a situated-cognition view on learning have argued that there are vast differences between learning in school and practical work outside the classroom. Due to these differences, students might become good learners in a school setting, but will be unable to transfer and apply their knowledge to real-life situations (Resnick, 1987).

To improve students' scientific reasoning and furthermore make these skills applicable to solve real-world problems, learning should therefore take place in an authentic context (Brown et al., 1989). Hence, in-classroom instructions should be combined with out-of-school learning to bridge the gap between knowledge and actions and to have students be part of activities that resemble those of professionals in the respective field, similar to an apprenticeship (Brown et al., 1989). This is assumed to prevent students from acquiring only inert knowledge that will remain unused when being confronted with real-world challenges (Whitehead, 1967).

Out-of-school learning can occur at several places outside the classroom, for example, during science centre visits (e.g. Itzek-Greulich et al., 2015), as well as during field trips in nature (e.g. Randler et al., 2005), or at zoos (e.g. Seybold et al., 2014). Field trips serves as an authentic activity, during which students can be given the opportunity to act like professional scientists. Therefore, the requirements that students have to accomplish during a field trip resemble those of practitioners, as they adequately represent the environment and the task of a scientist (Savery & Duffy, 1995)..

Furthermore, out-of-school learning enables students to recognize the relevance of a topic for themselves by highlighting the importance of understanding their personal environment; therefore, learning should have a higher subjective utility value for them. According to Eccles's expectancy-value theory, this higher personal value of the task should result in higher involvement, interest and learning (Eccles, 1983).

Moreover, motivation on a field trip might increase due to a higher level of self-determination of the students, as students have more opportunities to follow their own interests on a field trip than in the classroom. According to self-determination theory (Deci & Ryan, 1985), particularly competence and autonomy are important for intrinsically motivated, self-determined actions. Intrinsic motivation is a prerequisite to effective learning.

Previous studies have already shown positive effects of field trips on students' learning performance. For instance, Seybold et al. (2014), compared two teaching units concerning primates, with one taking place at the zoo and the other one at school. Students in the zoo-unit showed higher content achievement and higher interest or enjoyment as an indication of intrinsic motivation, but students at school reported higher values for perceived choice. In contrast, in another study by Wünschmann et al. (2017), perceived choice was higher for classes learning about reptilians at a zoo, compared to classes learning inside the classroom. Furthermore, the zoo-unit once again yielded higher contextual knowledge.

Out-of-school learning has shown positive effects on learning outcomes especially when combined with lessons inside the classroom. For instance, Pfeiffer, et al. (2012) combined a preparation inside the classroom with a visit to a public aquarium to train identification of European freshwater fish species. The university students either prepared with preserved specimens, digital videos, or a combination of both. Students who studied digital videos showed better identification performance, but less motivation than students, who had prepared with preserved specimens. Students who prepared using both methods showed good identification skills and high motivation scores, suggesting that the usage of digital media in combination with out-of-school learning might lead to even greater benefits.

However, not all studies show superior effects of out-of-school learning compared to in-classroom instruction only (e.g. Itzek-Greulich et al., 2015). Thus, effectiveness of out-of-school learning is influenced by several factors, like "the structure of the field trip, setting novelty, prior knowledge and interest of the students, the social context of the visit, teacher agendas, student experience during the field trip, and the presence or absence and quality of preparation and follow-up" (DeWitt and Storksdieck, 2008, p.181).

To further facilitate the connection between in-school and out-of-school learning and to enable the acquisition of scientific reasoning skills in an authentic context, students in the present study participated in a Citizen Science project. In Citizen Science projects, the expert status of scientists is transferred onto laypersons, who work like professionals while collecting scientific data in the field (Bhattacharjee, 2005). This leads to huge savings of time and money required for scientific research (Bonney et al., 2009), while also providing the possibility for participants to broaden their knowledge about science and to engage in scientific reasoning (Trumbull et al., 2000). As Citizen Science projects usually take place outside, they can serve as an out-of-school learning activity. Further, students collect and work with their own data in Citizen Science projects, which is likely to increase personal relevance and involvement in learning, thereby enhancing learning (e.g. Freeman et al., 2014).

Despite these promises, evaluations of Citizen Science projects have often revealed merely modest knowledge gains and no changes in attitudes towards science in adults (Brossard et al., 2005). A possible explanation for these findings is that most Citizen Science projects are not specifically designed to foster knowledge acquisition. As a consequence, Citizen Science projects that are curriculum-based, clearly structured, guided by an educator and potentially part of a larger teaching unit are expected to yield larger gains in scientific reasoning and knowledge (Bonney et al., 2016).

The possibility to easily collect data using digital devices and share information via internet has played a major role in the uplift of Citizen Science (Bonney et al., 2016). Next to making participation in a Citizen Science project easier and more fun for the students, mobile devices such as tablet computers also facilitate the connection of different learning environments by allowing to switch between scenarios, thereby enabling seamless learning (Chan et al., 2006). In line with this assumption, Pfeiffer et al. (2009) reported positive effects of mobile devices for linking different learning contexts. They asked university students to identify fish in the classroom with the help of digital learning materials on mobile devices. Afterwards, the fish were to be identified during a snorkelling field trip in nature. The mobile devices could be taken to the field trip location and could be used on site. Learning with the mobile devices enhanced students’ ability to correctly identify different fish species.

3.1.3 Current study and hypotheses

We developed a teaching approach that was grounded in theoretical assumptions of inquiry learning, situated cognition, and out-of-school learning and aimed to improve scientific reasoning, factual knowledge, and motivation. The teaching unit concerned water birds and bird migration and was, for the experimental group, integrated in a Citizen Science project and combined with a field trip; for the control group, the lessons were not part of a Citizen Science project and took place inside the classroom only. Furthermore, the experimental group collected data during the field trip and kept on working with this data later inside the classroom, which was expected to lead to a higher personal relevance for them and thus, a high level of involvement. In contrast, the control group worked with nationwide observation data from an online portal, which was expected to lead to less relevance and hence a low level of involvement. Students in the experimental group compared with the control group were expected to show a greater improvement in scientific reasoning (hypothesis 1), a greater increase in factual knowledge concerning birds (hypothesis 2), as well as to report more motivation (hypothesis 3) and more positive attitudes towards environmental issues (hypothesis 4) after the intervention.

3.2 Methods

3.2.1 Sample

The teaching unit on biodiversity of water birds was developed for students in grades 5 and 6 (secondary school) in line with the curriculum standards in Biology for the state of Baden-Württemberg in Germany. Approval for the study was granted by the state's Ministry of Education. For recruitment, information about the study were sent to 48 secondary schools. Fourteen classes agreed to participate. Students and their parents or guardians were informed about the study procedures and provided informed consent. 345 students (age: M= 11.21 years, SD= 0.55; 163 boys, 55 girls, 27 missing) took part in the study. For practical reasons, it was not possible to assign individual students to either the experimental or the control group; rather, group assignments were applied at the class level. This assignment procedure yielded 180 students participating in the experimental group and 165 students being part of the control group. Because there was no prior research which would have allowed to estimate effect sizes *a priori* as basis for a power analysis, we decided to run a sensitivity analysis instead using G*power (Cohen, 1988). This analysis ($\alpha = .05$; $1-\beta = .80$) suggested that with the aforementioned sample sizes the effect sizes would have to be at least $d = 0.27$ (Cohen's *d*) in order to be detected in a two-groups comparison with t-tests. Thus, the sample size would have allowed to detect even small effects. Because we actually did not run t-tests, but applied mixed linear models to also control for class membership, students' age and gender, the study sample would allow to detect even smaller effects.

Both age ($t(285.8) = -2.12$, $p = .036$) and gender ($U = 9510.00$, $Z = -4.35$, $p < .001$) differed between the two groups. Students in the control group (M = 11.28 years, SD = 0.58) were slightly older than students in the experimental group (M = 11.15 years, SD = .51) and more girls were in the control group (57 boys, 92 girls, 16 missing) compared with the experimental group (106 boys, 63 girls, 11 missing). Hence, these differences were controlled for in the analyses.

3.2.2 Study design

The study had two conditions (experimental vs. control group) and followed a BACI-design (before-after-control-impact, Randler & Bogner, 2008). Tests were presented three times to determine changes in the dependent variables. To assess a baseline, the first test was applied right before the beginning of the lessons (pre-test). To measure short-term effects, the second test (post-test 1) was administered immediately after the end of the lessons; finally, long-term effects were assessed by means of a follow-up test (post-test 2), six to eight weeks after the teaching unit. All tests mostly contained the same items, for exceptions see below.

3.2.3 Measures

Factual knowledge was assessed by using an adjusted version of a bird knowledge test developed by Randler and Bogner (2002). Students were asked to label pictures of birds ('Which bird species are shown in the photos?') and answer questions about the bird's ecology (e.g. 'What does the great crested grebe feed on?'). The 13 items of the adapted version only concerned water birds. Seven items were open-ended questions, six items were multiple choice questions. Open-ended questions were rated according to a rating system, which was developed with an ornithological expert and which was applied in the same manner to all the tests. Two items were excluded since students reported problems understanding them.

To assess scientific reasoning skills, we selected seven items of the *Trends in International Mathematics and Science Study* (TIMSS) concerning biology. Two items were taken from the 1994 study for seventh and eighth grade (IEA, 1994), two of the 2011 study for fourth graders (IEA, 2013b), and three of the 2011 study for grade 8 (IEA, 2013a).

An example for one of the tasks is:

Sandra has an idea that plats require minerals from the soil for healthy growth. She is placing a plant in the sun, giving it sand, minerals, and water. To check on her idea, she needs another plant. Which of the following should be used?

A. Dark cupboard, sand, minerals, and water
B. Dark cupboard, sand, and water
C. Sun, and sand
D. Sun, sand, and water
E. Sun, sand, and minerals

Three questions were open-ended, four questions multiple choice. Open-ended questions were rated according to the TIMSS guidelines.

Points for scientific reasoning and for factual knowledge concerning birds were converted into percentage correct.

Furthermore, we assessed environmental attitudes by using the 2-MEV Scales of Bogner (2007), which are two different scales. The first one, preservation, assesses students' attitudes toward preservation of the environment, by asking, for instance, whether the student would be likely to donate money for environmental protection. The second scale, utilization, focuses on attitudes toward using natural resources or changing nature for human needs, like cultivation of grain or constructing roads. The 2-MEV scales consist of 10 items each with a 5-point Likert-scale (from "totally wrong" to "totally right"). Higher values on the preservation scale show more positive attitudes toward the environment, while higher values on the utilization scale represent more negative environmental attitudes. Internal consistencies for the preservation scale (pre-test = .78; post-test 1 = .85; post-test 2 = .86) and for the utilization scale (pre-test = .75; post-test 1 = .80; post-test 2 = .83) were good. A test of the proposed factor structure with two related latent variables revealed that covariance between preservation and utilization was significant but small (*r* = -0.129, *p* = .002). Model fit based on root mean square error of approximation was good (RMSEA = 0.051) with confidence intervals between 0.040 and 0.062.

Lastly, we assessed situational motivation at post-test 1 with the KIM scale (Wilde et al., 2009; sample item: “The activity in the teaching unit was fun.”). The KIM scale is an adapted version of the Intrinsic Motivation Inventory (Deci & Ryan, 2003) consisting of four subscales with three items each (enjoyment, perceived competence, perceived choice and pressure), where students have to rate their level of agreement on a 5-point Likert-scale. Higher values on the subscales enjoyment, competence and choice represent higher motivation, whereas higher values on the pressure scale stand for higher pressure and thus, lower motivation. Internal consistency in the first post-test was .85 for enjoyment, .77 for competence, .41 for choice and .61 for pressure in our sample. The KIM’s factor structure has been confirmed in a large German school students sample (N = 1,861; Wüst-Ackermann et al., 2018).

3.2.4 Procedure

Students in both groups participated in three lessons that were held in consecutive weeks (Figure 1). The instructional time was the same for both of the groups with each lesson lasting about 90 minutes. All lessons were conducted by trained research assistants to ensure standardized procedures across all classes.

In lesson 1, all students first filled in the pre-test (no time limits) using paper and pencil. Afterwards, they received a short presentation with basic information about different bird species and learned about the functional ecology between beak forms and the connection with the preferred food of the birds. Then, they worked in pairs to apply their acquired knowledge from the presentation to six water bird species. They had to group water birds with the corresponding beak form and a tool that fulfils the same function as the bird’s beak (e.g., a pointed, narrow beak functions like tweezers). Afterwards, groups of four sat down together to discuss their results. Then, the class discussed the assignments together in plenary. Based on this discussion, each student was asked to formulate hypotheses for two of the six water bird species of his or her choice, how these species would behave in nature when feeding and which behaviour could therefore be observed when watching the birds during a field trip. Since the students were to check and possibly correct their hypotheses themselves during the next lesson, no feedback was provided at this stage.

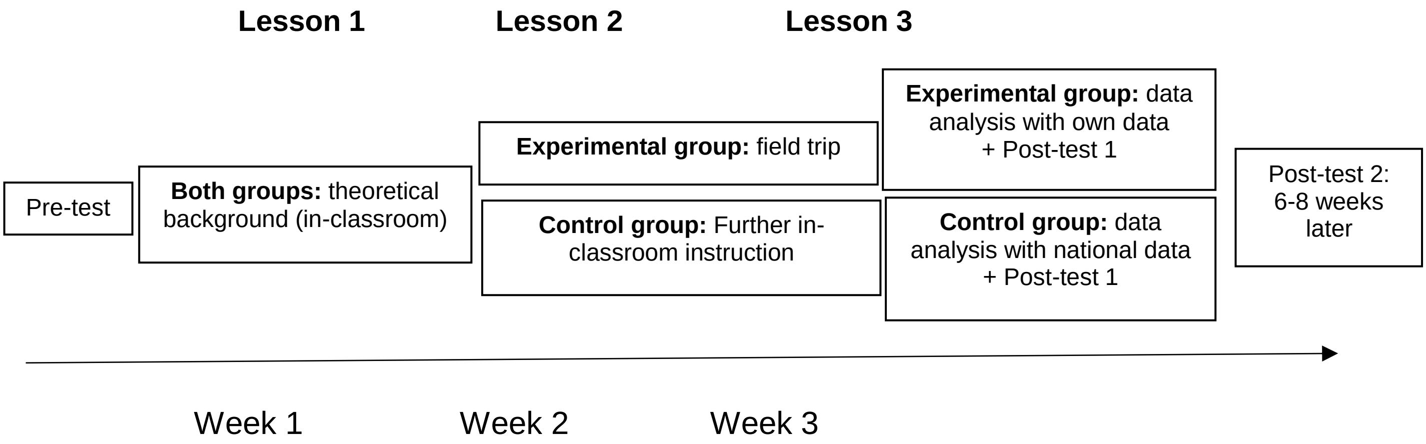

*Figure 1***.** Timeline of the study design. Lessons were implemented on regular school days, hence we divided the study in three parts and went into classes for three consecutive weeks. Each lesson lasted 90 minutes.

In lesson 2 the control group received another lesson inside the classroom, during which they learned more about water birds at different workstations. At work station 1, the students assigned different tools to the beak shapes of the waterfowl. A worksheet prompted them also to reason whether this beak shape is an adaptation to the birds' special diet. At stations 2 and 3, the students watched short film sequences showing the six different water bird species feeding on tablets. Based on these film sequences, the students could check and possibly correct their hypotheses of the last lesson. At workstation 4, the students were given multimedia instruction about how the waterfowl are adapted to their habitat (e.g. by webbed feet or special plumage), apart from their beak shape. These results were also written down on a worksheet. The four workstations could be visited in any order. The research assistants and the teacher were always available for questions and support. The results of all stations were discussed and corrected in plenary at the end of lesson 2.

In contrast, students in the experimental group went on a field trip to lakes or rivers close to their school, where waterfowl can be found. The students and their teacher first met with the research assistants in the classroom. Every student was given binoculars; furthermore, pairs of students were given a tablet and a clipboard; always four students were given a stopwatch. They received a short instruction about how to use the binoculars and afterwards went to the lake or river together with at least two research assistants. There the class was divided into at least three small groups. Each group went with a supervisor (either an assistant or a teacher) to a different spot close to the water, in order to make it easier to observe the birds without disturbing them. Students were given three tasks. First, they were asked to observe the behaviour of the waterfowl while feeding and use their observations to check and possibly correct their own hypotheses of the previous lesson. Second, they were asked to identify birds and submit their observations to the Citizen Science data base (ornitho.de, DDA, n.d.) using the tablets. The tablets contained the Citizen Science project's app for simplified bird registration (NaturaList), as well as a guide for easier bird identification (with pictures of different bird species, as well as information about colour and size, etc.). As the data of the birds were entered during the field trip, they were automatically registered for the location, where the students had seen them. Third, the students were asked to observe the general behaviour of the birds and to note on a worksheet how often a bird shows a certain behaviour within a minute (e.g. rests, flies away, eats grass, etc.). The research assistants and the teacher were always available for questions and support. At the end of the lesson, the class and the research assistants went back to school together.

In lesson 3, students of both groups were given the solution to the task from lesson 1 and the hypotheses were discussed and corrected with the help of the research assistants. In addition, students from the experimental and control group were assigned one of the six species of waterfowl for which they completed another worksheet. With the help of data from the Citizen Science database, they were asked to draw on a map where the birds of this species spend the winter in Germany. Further, they should consider why the species are present in certain areas during winter and not in others. The research assistants helped the students with these tasks if necessary. The experimental group was additionally asked to compare their observations with the data from ornitho.de. To make the amount of data more manageable for them, the data were pre-selected and handed out as prints. At the end of the lesson the results for each species were presented in the plenum and discussed and corrected.

At the end of the lesson, all students completed the post-test 1; post-test 2 had to be completed about six to eight weeks after lesson 3. There were no time limits for the tests.

3.2.5 Statistical analyses

All analyses were carried out with IBM SPSS and AMOS Statistics 26 (IBM, Somers, NY). To assess effects on scientific reasoning, factual knowledge, motivation, and environmental attitudes, mixed linear models with repeated measures were calculated.

The factors group (experimental group or control group), time (three measuring points), gender (male and female) and the interactions of group with time, group with gender, gender with time and group with time with gender were included as fixed effects in the mixed linear models with repeated measures. All calculations were corrected for class effects (as random factor) and age (as a fixed covariate), which had no significant effects on any of the dependent variables.

3.3 Results

3.3.1 Scientific Reasoning

Descriptive values can be found in table 1. Applying the linear mixed model to the data yielded a significant model (Rho = 0.54, $p < .001$). The model fit was best, when all parameters mentioned above were included (AIC = 3982.85, see supplementary materials (Appendix A) for details).

There were no effects of any of the control variables, namely, class (*Wald Z* = 1.09, $p = .277$), age ($F(1, 128.49) = 0.01$, $p = .941$) or gender ($F(1, 310.82) = 2.84$, $p = .093$), nor of group ($F(1, 14.37) = 1.36$, $p = .263$) or interactions of either experimental condition (group x gender: $F(1, 311.46) = 0.48$, $p = .489$) or the time factor with gender (time x gender: $F(2, 549.85) = 1.90$, $p= .150$), nor the triple interaction of group x time x gender ($F(2, 549.88) = .79$, $p= .453$). However, there was a significant effect of time ($F(2, 549.82) = 12.96$, $p < .001$), suggesting that across both groups students' scientific reasoning skills improved from the pre-test to the post-tests. Most importantly, there was a significant interaction of time and group ($F(2, 549.80) = 3.90$, $p = .021$) in that the improvement of scientific reasoning skills was more pronounced in the experimental group than in the control group (see Figure 2).

Table 1

Means and standard deviations in the dependent variables for the three measuring points by groups

		Experimental Group			Control Group		
		N	Mean	*SD*	*N*	Mean	*SD*
Pre-test	Scientific Reasoning	169	53.47	18.10	155	53.59	19.69
	Factual Knowledge	169	29.84	15.91	155	28.31	16.13
	Preservation	169	3.39	0.58	155	3.49	0.68
	Utilization	168	4.01	0.64	155	3.99	0.69
Post-test 1	Scientific Reasoning	160	60.63	17.77	154	56.49	18.79
	Factual Knowledge	160	68.22	15.69	154	70.50	15.61
	Enjoyment	156	4.13	0.67	147	4.07	0.91
	Competence	155	3.59	0.65	146	3.65	0.86
	Choice	155	3.09	0.89	143	3.35	1.82
	Pressure	156	1.95	0.82	147	1.93	0.92
	Preservation	159	3.40	0.65	154	3.47	0.81
	Utilization	159	4.17	0.64	154	4.19	0.71
Post-test 2	Scientific Reasoning	169	61.83	17.79	152	56.39	21.09
	Factual Knowledge	169	57.26	20.83	152	59.89	20.75
	Preservation	168	3.29	0.72	149	3.32	0.88
	Utilization	168	4.11	0.76	149	4.19	0.75

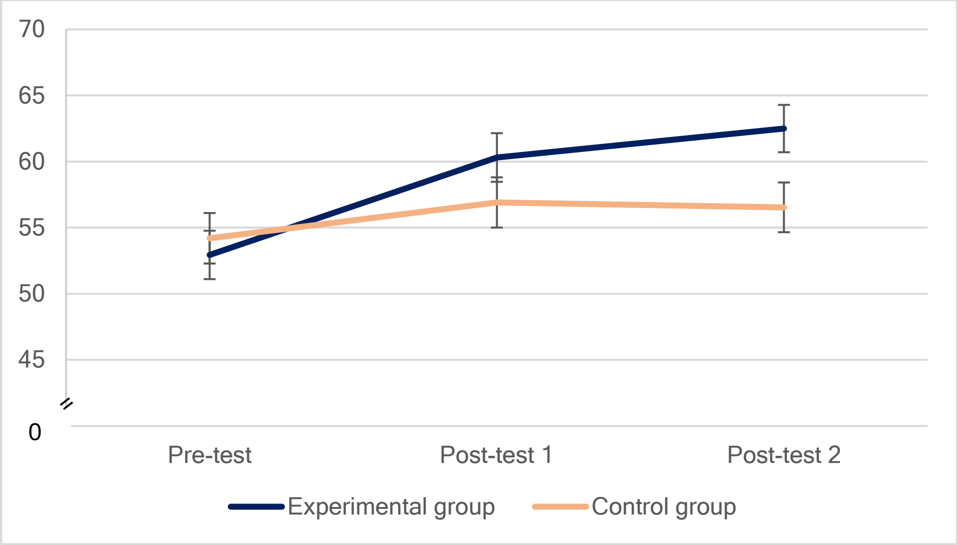

Figure 2. Changes in scientific reasoning (% correct) over time for experimental and control group. Bars indicate standard errors.

3.3.2 Factual Knowledge

Applying the linear mixed model to the data yielded a significant model (Rho = 0.39, *p* < .001). The model fit was best, when all parameters mentioned above were included (AIC = 3845.93, see Appendix A).

There were no effects of any of the control variables, namely, class (*Wald Z* = 1.49, *p* = .137), age (*F*(1, 176.74) = 0.01, *p* = .920) or gender (*F*(1, 288.55) = 0.12, *p* = .733), nor of group (*F*(1, 11.50) = 1.16, *p* = .304), and no interaction of gender with the time factor (time x gender: *F*(2, 526.39) = 1.34, *p*= .264). However, there was a significant effect of time (*F*(2, 526.33) = 577.66, *p* < .001), suggesting that across both groups students' factual knowledge improved from the pre-test to the post-tests (see Figure 3). Moreover, the model showed significant effects for both the interaction of group and gender (*F*(1, 286.51) = 4.70, *p* = .031) and for the triple interaction of group x time x gender (*F*(2, 526.48) = 3.57, *p* = .029). Contrary to our assumptions, there was no interaction of group and time, *F*(2, 526.29) = 2.16, *p* = .117, suggesting that the in-classroom instruction had been equally effective in teaching factual knowledge as the Citizen Science treatment.

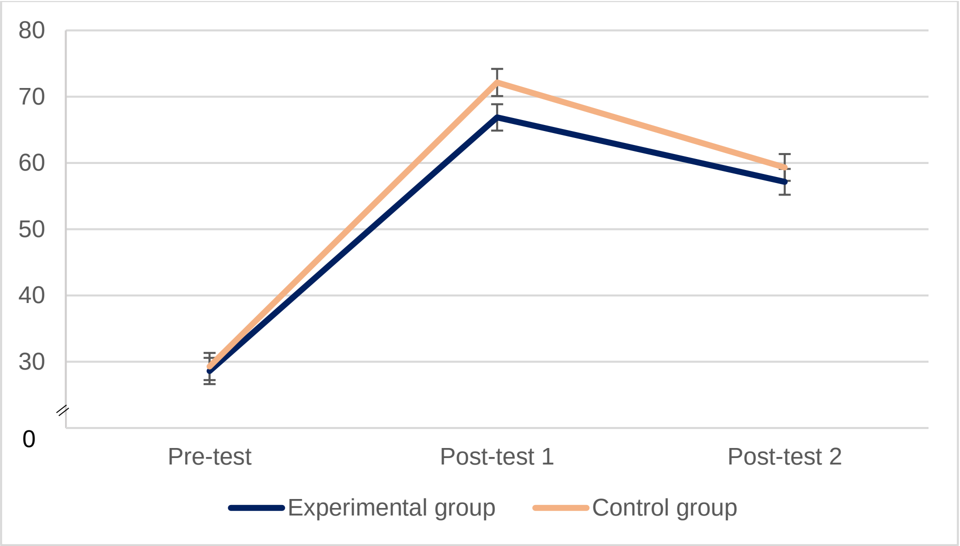

Figure 3. Changes in factual knowledge (% correct) over time for experimental and control group.

3.3.3 Motivation

Detailed results of the mixed models are summarized in Table 2. There was no difference between groups for any of the subscales, other than for the subscale pressure regarding the interaction of gender and group, which showed a significant effect, *F*(1, 266.99) = 5.08, *p* = .025. Calculations were based on the model with the best AIC fit, see supplementary materials (Appendix B) for details.

3.3.4 Environmental Attitudes

Analyses of the data on preservation showed that class did not converge in the mixed model. Hence, class was excluded for further calculations. Applying the linear mixed model with all remaining variables yielded a significant model (Rho = 0.76, *p* < .001). The AIC was 1353.38 (see Appendix C for details).

There was no significant effect for the control variable age (*F*(1, 313.96) = 2.75, *p* = .095), but for the control variable gender (*F*(1, 317.57) = 28.47, *p* < .001). Moreover, there was a significant interaction of time and gender (*F*(2, 560.05) = 3.72, *p* = .025) and group and gender (*F*(1, 317.22) = 3.96, *p* = .048) as well as of time (*F*(2, 560.04) = 8.78, *p* < .001). However, neither group (*F*(1, 318.32) = 0.29, *p* = .590), nor for the interaction of time and group (*F*(2, 560.06) = 0.44, *p* = .645) or the triple interaction of group x time x gender (*F*(2, 560.06) = 1.90, *p* = .151) showed significant effects.

Table 2

Results (p-values) of mixed model analyses regarding motivation.

	Variables included	Group	Age	Class	Gender	Gender* Group
Enjoyment	Group, age, class	.801	.481	.103	-	-
Competence	Group, age, class	.666	.624	.161	-	-
Choice	Group, age, gender, gender*group	.207	.308	-	.220	.065
Pressure	Group, age, class, gender, gender*group	.958	.468	.333	.085	.025*

Notes. Calculations were always based on the model with best AIC fit. Class was not further included in calculations regarding choice, as the model did not converge and thus, class did not have any effect in the model.

Based on these results and on the AIC model fits, group and all interactions including group were removed as variables from the model and the analysis was rerun, focusing on gender. AIC for this analysis was 1343.22 (see Appendix C for details). The model was highly significant (Rho = .76, p < .001). Girls showed significantly higher preservation scores than boys, *F*(1, 320.28) = 27.59, *p* < .001. Moreover, time (*F*(2, 564.67) = 9.65, *p* < .001) and the interaction of time and gender (*F*(2, 564.69) = 3.23, *p* = .040) still had significant effects in this corrected model. Age still showed no significant effect, *F*(1, 316.12) = 2.21, *p* = .138. Changes of preservation attitudes over time by gender are depicted in Figure 4.

Concerning utilization, AIC was 1423.3, when all variables were included (see Appendix C for details). Applying the linear mixed model to the data yielded a significant model (Rho = 0.73, *p* < .001).

This model did not show significant effects for group (*F*(1, 12.88) = 0.02, *p* = .881), nor for age (*F*(1, 173.79) = 0.68, *p* = .411), the interaction of either gender with time (*F*(2, 552.77) = 2.57, *p* = .078), or of group with time (*F*(2, 552.74) = .54, *p* = .583), or for class (*Wald Z* = 1.45, *p* = .147). Moreover, there was no triple interaction of group x time x gender (*F*(2, 552.80) = 0.12, *p* = .890). However, time (*F*(2, 552.73) = 14.14, *p* < .001), gender (*F*(1, 311.73) = 14.06, *p* < .001) and the interaction of group and gender (*F*(2, 309.69) = 4.13, *p* = .043) had significant effects on utilization attitudes. As the analysis of preservation attitudes had already revealed the stronger impact of gender than of the implemented teaching unit, group, group and all interactions including group were removed from the model. Model fit for this calculation improved, AIC was 1411.82 (see Appendix C for details). The model was highly significant, Rho = 0.74, *p* < .001.

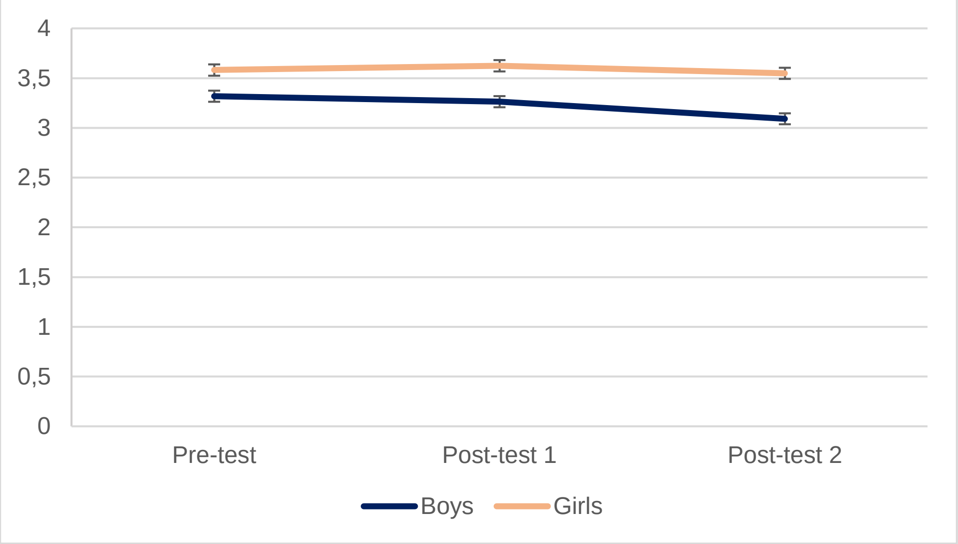

Figure 4. Changes in preservation attitudes over time for boys and girls.

The model showed significant effects for time, *F*(2, 556.49) = 14.44, *p* < .001, and for gender, *F*(1, 308.92) = 13.92, *p* < .001, with girls reporting higher utilization scores. The interaction of gender and time was significant as well (*F*(2, 556.56) = 3.10, *p* = .046) with girls showing increasingly higher utilization scores in the course of the unit, while the utilization scores of boys decreased. Age (*F*(1, 169.49) = 0.70, *p* = .405) and class (*Wald Z* = 1.49, *p* = .137) showed no significant effects. Changes of utilization attitudes over time by gender are depicted in Figure 5. Attitudes are depicted for boys and girls of both of the groups separately to make it possible to see the interaction between group and gender, which was significant in the first model.

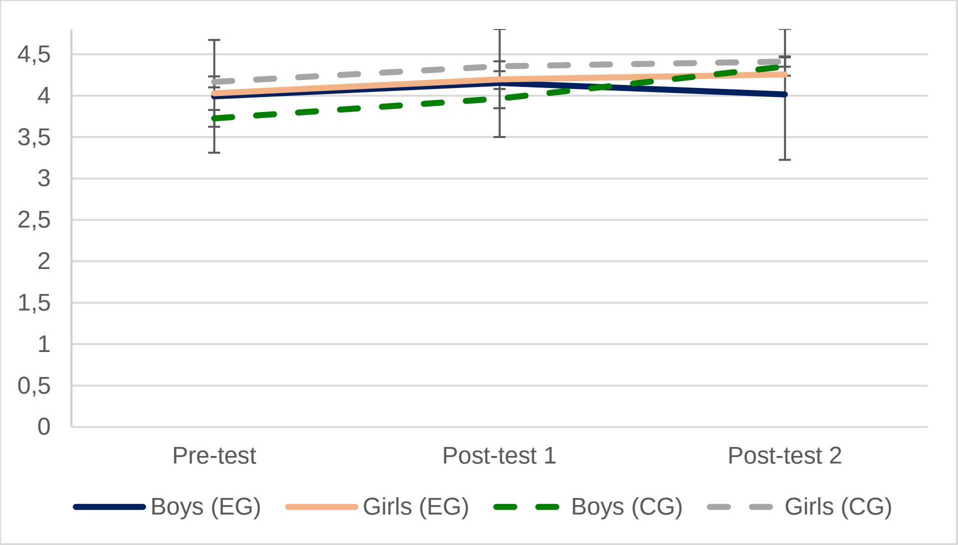

Figure 5. Changes in utilization attitudes over time for boys and girls in the different groups.

3.4 Discussion

In this study we investigated whether a teaching approach developed against the backdrop of inquiry learning, situated cognition, and out-of-school learning implemented as a Citizen Science project improves scientific reasoning, knowledge gain, motivation and environmental attitudes on students in 5th and 6th grade. Given the lack of research on this issue, our study provides first insights into possibilities and challenges when implementing such a teaching concept in schools. The teaching approach had a significant impact on improving students' scientific reasoning skills. Concerning knowledge, motivation, and environmental attitudes the Citizen Science treatment and the in-classroom instruction were equally suited to influence these variables.

3.4.1 Results regarding scientific reasoning, motivation, and factual knowledge

As out-of-school learning during a field trip provides an authentic activity for training and applying scientific reasoning in a real-life context, thus, explaining its benefit. Importantly, this benefit was still visible six to eight weeks after the end of the teaching unit. As Henry (1992) already demonstrated in the context of a museum visit, out-of-school learning yields lasting knowledge, which can often still be remembered after months and thus differs from knowledge acquired in school, which is often forgotten quite easily (Baumert et al., 1997). Identifying effective ways to teach scientific reasoning is highly important given the problems that students often face in this regard; it was hence the main aim of the study.

We could not find any differences in motivation between the two groups. Thus, the differences in scientific reasoning cannot be attributed to the fact that the teaching approach was more motivating, thereby possibly fostering student engagement. Instead, they are more likely due to enabling more direct cognitive benefits, where - due to engaging in professional-like observation and reasoning tasks situated in the students' real-world experiences – students find it easier to grasp what scientific reasoning is about (Brown et al., 1989). Moreover, we found quite high values in motivation also in the control group, which may have made it difficult to obtain an add-on benefit of the Citizen Science treatment.

There were also no group differences concerning factual knowledge. Thus, our findings are in contrast to several previous studies, for example by Seybold et al. (2014), who showed a positive effect of out-of-school learning on knowledge gains. On the other hand, there are also previous studies that have not found any effects of out-of-school learning either, for example, when students were learning in a science museum setting (Kubota & Olstad, 1991). Our results are probably due to the fact that the in-classroom instruction was sufficient to acquire this knowledge, which is why also students in the control group improved over time. This suggests that due to the in-classroom preparation all students possessed the necessary background knowledge to successfully engage in further in-classroom or in the field trip activities.

3.4.2 Unexpected effects on environmental attitudes

There were no effects on environmental attitudes concerning preservation or utilization. However, regardless of group membership, environmental attitudes developed negatively over time. At this point we do not have a good explanation for this finding. The educational program did not have an explicit environmental education component, but the changes in attitudes were assumed to occur in an implicit manner, because being outside, working on ecological questions are inevitably linked with environmental issues. Maybe secondary school students can perceive these issues only when addressed more explicitly, for instance, by adding corresponding tasks to the work sheets. Still, this interpretation does not explain why attitudes developed negatively. Further studies should attempt to replicate this finding to hopefully rule out that teaching approaches implemented as Citizen Science projects, but also in-classroom activities tailored towards enhancing scientific reasoning have such unintended side effects. At the same time, it has to be acknowledged that while being statistically significant the changes over time were rather small on an absolute level, thereby suggesting that they may not necessarily reflect a meaningful decline.

We found significant gender effects on both preservation and utilization scores. On both scales, girls showed higher preservation attitudes as well as higher utilizing attitudes. Regarding preservation, this finding is in line with previous research (e.g. Bogner & Wiseman, 2006). In contrast to what we found, however, girls usually show lower utilization scores than boys and hence more positive attitudes (e.g. Bogner & Wiseman, 2006).

3.4.3 Limitations

We used trained research assistants to conduct the teaching unit at the different schools to ensure a high standardization of the teaching unit and the study. Nevertheless, we cannot fully out that there were differences between the six research assistants in how they interacted with students. However, we controlled for potential variability by using class as random factor in the analyses. Furthermore, standardization is always difficult to maintain, when working with living animals in nature. Even though we checked the field trip sites beforehand with regard to the occurrence of water birds, it cannot be guaranteed that living animals can be found in the same places every day. Thus, some classes may have observed fewer individuals or fewer different species during their field trip than others. This leads both to an easier distraction during the field trip, as there are fewer task-relevant stimuli available, and to less learning opportunities, which might result in less knowledge gain, and no effective manipulation of involvement. As involvement was supposed to be induced by observing “own” birds and working with own data afterwards, the absence of birds might result in low involvement and frustration, instead of high involvement and motivation. As the role of involvement could not be finally clarified due to the manipulation possibly being too weak, further research focusing only on the manipulation of the involvement in Citizen Science projects should be carried out to clearly determine the effects of involvement.

3.4.4 Contributions

As the study was designed to fit into the regular schoolwork and was also conducted in schools and places close to the schools, the study is of high ecological validity. Hence, our results are valid for everyday school-life. Furthermore, transferring the teaching unit into a regular lesson plan is very easy. Therefore, teachers could implement the project in their classes without much effort.

The study highlights the effectiveness of a Citizen Science project for training scientific reasoning. In this paper, we presented a useful method for achieving the central goal of profound scientific reasoning skills in students (National Research Council, 2012) and thereby overcoming the deficits found in large-scale assessment studies in this area (IQB National Assessment Study 2012, Pant et al., 2013). Citizen Science projects have thus once again proven to be not only of use for the scientists, but also for the participants, who can, in this case, improve their scientific reasoning, making Citizen Science projects a win-win situation for both sides (Riesch & Potter, 2014).

4. Manuscript 2: Can Involvement induced by Guidance foster Scientific Reasoning and Knowledge of Participants of a Citizen Science Project?

Author	**Author position**	**Scientific ideas %**	**Data generation %**	**Analysis & interpretation %**	**Paper writing %**
Alena Rögele	1	35	60	85	70
Katharina Scheiter	2	35	5	5	5
Christoph Randler	3	30	35	10	25
Title of paper:		Can Involvement induced by Guidance foster Scientific Reasoning and Knowledge of Participants of a Citizen Science Project?			
Status in publication process:		Submitted in February 2021			

Please note: As this manuscript is building on the results of the first study presented before (see point 3), some topics and theoretical assumptions on which the second manuscript is build are the same. Thus, the following manuscript partly overlaps with the first manuscript (see point 3), the theory presented previously (see point 1) and the discussion that follows (see point 5) in content and structure.

Abstract

Citizen Science projects are continuously growing in popularity as they offer a unique possibility to conduct large scale research projects as well as allow citizens to broaden their knowledge about the research topic or the process of scientific investigations. However, the benefits for participants of a Citizen Science project vary, depending on the way and the amount of participants' personal involvement in the project. In this study we investigated whether additional guidance provided by a professional ornithologist would improve involvement and hence lead to greater knowledge gains, higher commitment and improved scientific reasoning of participants in a Citizen Science project on birdwatching. A group receiving guidance was compared to a group which only took part in the same project without receiving additional support. Results showed that both groups enhanced their knowledge about birds and their commitment to birding during the first three weeks of participation, while there were no significant changes in scientific reasoning. However, participants receiving additional guidance observed a higher number of different bird species and reported being able to identify significantly more species than participants without guidance. Accordingly, participating in a Citizen Science project either with or without further guidance can be seen as powerful in supporting birders at the beginning of their participation in a Citizen Science project. Thus, both versions should be taken into account when designing future Citizen Science projects.

4.1 Introduction

Citizen Science projects offer participants a unique possibility to contribute data to scientific research, as they can take over the part of a scientist (Bhattacharjee, 2005) and collect data in the field. For example, about 600,000 people worldwide use the online birding website eBird (www.ebird.org, The Cornell Lab of Ornithology, n.d.) and 19,000 its German equivalent Ornitho.de (www.ornitho.de, DDA, n.d.) to submit their bird observations to Citizen Science project. People using this portal stretch from novices to highly specialized birdwatchers with diverse knowledge, personal commitment and behaviour. The recruiting process is somewhat haphazard, and there are only few studies about people joining these networks, and their background in terms of knowledge, commitment and behaviour (Wood et al., 2011). Citizen Science projects aim both at gathering new scientific data as well as having the participants learn about the research topic and the process of scientific investigations (Bonney et al., 2009). To better understand the latter aspect, studies should focus on the citizens taking part in Citizen Science projects and investigate which benefits a participation holds for them and how project leaders could support participants in gaining knowledge and foster their scientific reasoning skills. Working with data obtained from eBird, Kelling et al. (2015) showed in a post hoc analysis that participation in the Citizen Science project over a longer period of time and the continuous reporting of data to the online database seemed to increase participants' knowledge about birds. Here, we focus on the process of acquiring knowledge, scientific reasoning skills, and commitment to birding by accompanying people during the first weeks of their Citizen Science initiation on ornitho.de, and by comparing a self-regulated, unguided online group with a group receiving additional guidance.

4.1.1 Citizen Science

Citizen Science projects are growing in popularity in scientific research as they can be used to study large-scale phenomena in nature (Bonney et al., 2009). During their participation in a Citizen Science project, the participants take over the expert status of the scientists (Bhattacharjee, 2005) and collect data in the field, which are then later used by scientists in their research projects. Citizen Science has become an important topic for both science and society. On the one hand, scientists have the possibility to gather a huge amount of scientific data, which would not be possible without the help of citizens. For example, in a project about birds, citizens contributed over 200.000 hours of data collection, which would be worth millions of dollars, if these data would have been collected by scientists (Bonney, 1991). On the other hand, Citizen Science participants have the opportunity to be part of a real scientific investigation and thus, might broaden their knowledge not only about the research topic, but also about scientific methods and the scientific inquiry process (Bonney et al., 2009).

Especially well-suited for Citizen Science projects is ornithology, as it is one of the fields in which even lay people can make important contributions concerning data collection (Trumbull et al., 2000; Bonney, 1991). Thus, many Citizen Science projects are based in ornithology, for example eBird, a website where participants can upload data concerning the presence, absence, or the exact number of a certain bird species at their location (Bonney, 2007) or its European counterpart ornitho, where some European countries use the same platform. For example, data from eBird were used to assess where citizens observe birds within a city, which can have implications for conservation efforts (Lopez et al., 2020). Further projects are concerned with monitoring nesting behaviour (Brossard et al., 2005; Evans et al., 2005) or food preference of birds (Trumbull et al., 2000). As birding is a popular recreational activity (Frątczak et al., 2020, Randler et al., 2020) and amateurs can contribute valuable data in this field, ornithology is the ideal scientific field to research benefits of participating in a Citizen Science project.

4.1.2 Birding

Birding is a nature-based recreational activity that can be classified alongside a gradient from novice to experienced birder (McFarlane, 1994). It represents a continuum between the generalists with low involvement and the specialists with a high involvement. Therefore, birding is an activity that can be explained in terms of recreation specialization theory (Bryan, 1977). Recreation specialization consists of the three dimensions: skill, commitment, and behaviour (Lee & Scott, 2004), with skill being related to the ability to identify birds, behaviour to the number of birding trips and money invested in equipment, while commitment is measured as a psychological variable, including personal and behavioural commitment. Concerning Citizen Science projects, such as eBird or Ornitho, birdwatchers can contribute a huge amount of data for science, often in a high quality, depending on their specialization gradient.

4.1.3 Learning about Bird Identification

Knowledge about bird species is essential to understand ecological relationships (Randler, 2008a). While there are some studies about bird species knowledge in students (Randler, 2008b, Gerl et al., 2021), only few analysed adult laypersons and professionals (Hooykaas et al., 2019). Learning to identify bird species in nature is far from easy. In controlled school and university settings, identification of only a handful of species can be learned within a given time frame (Randler & Bogner, 2006), so learning under real-life conditions is much more difficult. Nervertheless, participants of a Citizen Science project haven been shown to gain knowledge about bird biology during their participation in a project monitoring nesting behaviour (Brossard et al., 2005; Evans et al., 2005). Further, regarding Citizen Science projects that do not deal with birds, participants of a project about invasive plants were also able to increase their content knowledge by on average of 24 percent (Jordan et al., 2011). Moreover, participants of another project on invasive species, who were working with global positioning systems (GPS), showed a knowledge gain not only regarding invasive species, but also about GPS and concerning vegetation monitoring after an eight-hour training day (Crall et al., 2013).

4.1.4 Scientific Reasoning

Scientific reasoning is an umbrella term for various skills needed to be carried out scientific work or a scientific thinking process. Parts of scientific reasoning are ‘asking scientifically oriented questions, giving priority to evidence in responding to questions, formulating explanations from evidence, connecting explanations to scientific knowledge, and communicating and justifying explanations’ (National Research Council, 2000, p. 23). Scientific reasoning is not a clear, linear process that always follows the same pattern (Bauer, 1992), but a complex interplay of deduction, classification of different objects and the repeated search for connections and explanations (Duschl, & Grandy, 2008). Hence, scientific reasoning is a complex skill, which is often difficult to convey.

Improving public understanding of science and hence also learning how to reason scientifically is a major goal of many Citizen Science projects (Bonney et al., 2016). As Cohen (1997) argued, participants might acquire knowledge concerning the process of science or scientific inquiries, as they can be a part of a real scientific process. However, there are only few studies taking a look at learning outcomes or possibly increased understanding of science or scientific reasoning. Only a handful of studies reported changes in scientific understanding yet. For example, Trumbull et al. (2000) found that participants showed thinking processes that resembled those of scientists. Similarly, Golumbic et al. (2020) showed that data interpretation skills of participants improved by partaking in a Citizen Science project, which even reduced the influence of the scientific education level on this skill. Thus, participation in a Citizen Science project might trigger scientific thinking processes (Trumbull et al., 2000), but the assumption that knowledge gains and improvements in scientific reasoning simply occur due to exposure to a real scientific investigation does not seem to be true (Bonney et al., 2016). Rather, to help participants improve their knowledge and scientific reasoning, it may be necessary to have them engaged more strongly in the research process (Druschke & Seltzer, 2012). Engagement of participants can vary depending on the degree of their participation in a project. Shirk et al. (2012) propose a model that differentiates mainly between contributory, collaborative and co-created projects. Contributory projects are developed by scientists and participants are involved only in the data collection process. In collaborative projects, citizens can incorporate their experiences during participation in the project and can assist the scientists for example in refining the project or in data analyses. Finally, co-created projects are developed by scientists and citizens in cooperation and participants of the Citizen Science project are part of almost all steps of the scientific investigation. If projects are co-created and offer a high degree of participation, participants show increased understanding of science as well as higher content knowledge and scientific reasoning skills, whereas less positive outcomes for participants can be found in contributory projects. Although these participants show an increase in content knowledge as well, they do not develop a deeper understanding of scientific processes (Shirk et al., 2012). In line with these results, the National Academies of Science, Engineering, and Medicine (NASEM, 2018) state as well that scientific reasoning skills might not be fostered by simply participating in a Citizen Science project; rather, more scaffolding or commitment of participants may be required in order to acquire knowledge from them.

However, since not all projects can allow for a high degree of participation or even co-creation, this study deals with the extent to which citizens can be engaged in a project in other ways. Previous research already showed that for example the intensity of involvement can support beneficial outcomes as well (Wilmsen & Krishnaswamy, 2008). Based on these results, we varied the intensity of involvement for the participants by comparing two groups that received different levels of guidance. We expected additional guidance to result in higher commitment and a higher intensity of involvement in the project, thereby leading to more beneficial outcomes, which would be greater improvements in learning and scientific reasoning when participating in the Citizen Science project in a guided, compared to an unguided way. Thus, we investigate how positive outcomes for individual participants can be achieved even when only a small degree of participation in a contributory project might be possible by offering different amounts of guidance. To our knowledge, this is the first study to investigate the role of guidance in a Citizen Science project. However, moving beyond this specific context, guidance has been shown to improve conceptual understanding and scientific skills in more formal learning contexts, such as inquiry learning at school or university (Lazonder & Harmsen, 2016).

4.1.5 Hypotheses

Hypothesis 1: Both groups improve their content knowledge concerning birds, but the group receiving a higher level of guidance (guidance group) shows a greater increase than the group without guidance (online group).

Hypothesis 2: The guidance group shows a greater increase in scientific reasoning skills than the online group.

Hypothesis 3: As participants of the guidance group should have a higher feeling of involvement and commitment for the project, they report more observations and observe more different bird species during a three-week observation period than participants of the online group.

Hypothesis 4: Due to the higher feeling of involvement, the guidance group shows a greater increase in commitment to birding than the online group.

4.2 Materials and Methods

4.2.1 Participants and Data Collection

The study was approved by the Ethics Committee of the Leibniz-Institut fuer Wissensmedien (IWM) Tuebingen (LEK 2019/043). Recruitment e-mails were sent to employees and students of the Eberhard Karls University of Tuebingen containing information about the study procedure and the aim of the study. Respondents received more information and (in case of the online group) the link for the first online questionnaire. Participants gave informed consent before starting their participation.

The guidance group was recruited in winter 2019, the online group in autumn 2020. Due to the Covid-19 pandemic, adjustments of the study design were necessary (see more information below, section 4.2.4, Procedure).

We recruited 113 participants in total, 46 participants formed the guidance group (age: M = 34.44 years, SD = 15.95; 19 male, 27 female) and 67 participants were part of the online group (age: M = 34.82 years, SD = 13.96; 22 male, 44 female, 1 diverse).

There were no differences in age (*t*(110) = -.13, *p* = .895) and in gender ($\chi^2(2) = 1.44$, *p* = .488) between both groups.

4.2.2 Interventions

We compared two different interventions: A) an online group, B) a guidance group. Both groups received the same three-week bird observation period, flanked by a pre-test prior to the observation period and a post-test immediately afterwards. The groups differed in that the guidance group received personal instructions about birdwatching, the use of the Citizen Science portal Ornitho, as well as some birdwatching aids, such as how to identify bird species by one of the authors, whereas the online group was not given further aid before starting their observation period. This group received only a few general instructions about the study and the use of ornitho.

4.2.3 Measurement Instruments

Content Knowledge concerning Bird Species

Species knowledge was assessed by colour pictures. The selection of the 20 bird species was based on their abundance in the study area (see e.g., regional reports in Anthes et al., 2019; Anthes et al., 2020; Randler 2008b). The respondents had to identify the birds as precisely as possible. Scoring was based on previous partial credit models (Gerl et al., 2021; Randler, 2008b). Every correct answer was scored with 1, the correct identification of the genus with 0.5. For example, identifying a bird as a great tit (*Parus major*) received 1, identifying it as tit only received 0.5 points. Cronbach’s alpha was .90 in the pre-test and .85 in the post-test.

Additionally, we asked for the self-assessment of skill, i.e., how many bird species could be identified by sight or by sound without a field guide (Lee & Scott, 2004) with open-ended questions.

Scientific Reasoning

Scientific reasoning was measured using a selection of items of a scientific reasoning test (Hartmann et al., 2015; Krüger et al., 2020). We chose a selection of 12 items, which assessed interpretation of results, generation of hypotheses etc.. An example item is the following:

> ‘Squirrels hoard acorns in autumn as a food supply, which they dig up again in winter. Scientists hypothesise that the squirrels only find the acorns hoarded by themselves, but not the acorns of other squirrels.
>
> What experiment could be used to test this hypothesis?
>
> A) Investigating whether squirrels leave chemical marks on their own hoarded acorns.
> B) Investigating whether squirrels find acorns hoarded by humans less frequently than self-buried acorns.
> C) Investigating whether squirrels find hoarded acorns of conspecifics.
> D) Investigate whether squirrels find food in areas without self-hoarded acorns.’
>
> Cronbach’s alpha was .37 in the pre-test and .57 in the post-test.

Birding Specialization

We used four items to measure birding specialization: Two items covered the psychological commitment component (‘I would rather go birding than do most anything else’, ‘I find that a lot of my life is organized around birding.’, Scott & Lee, 2010, previously used in Lee & Scott 2004; Lee & Scott, 2006) and two items asked for support of family and friends (‘My family members are sympathetic about my going birdwatching.’, ‘I have supportive friends in birding.’, adapted from Scott & Lee, 2010). Participants could respond to these items on a 7-point Likert-scale ranging from 1 = ‘I totally disagree’ to 7 = ‘I totally agree’. Cronbach’s alpha was .73 and .69 for pre-test and post-test, respectively.

Birdwatching Activity

In the post-test, we further asked for the number of species and number of observations the participants had submitted to the ornitho database during the three-week observational period. To get an impression of bird abundance during the study periods, twice a week, one of the authors counted birds on a 5.5 km strip within the study area (eBird hotspot: Rottenburg-Weggental, Baden-Württemberg, DE). In addition, we compared the number of bird species observed during both three-week observation periods by checking the regional reports (Anthes et al., 2019; Anthes et al., 2020).

4.2.4 Procedure

Initially, the study was planned to compare a control group to two different intervention groups, who would have on-site appointments, but would receive different amounts of guidance during the three-week period between pre- and post-test. We started collecting data in the two intervention groups. Both intervention groups participated in the Citizen Science project during the three weeks between pre- and post-test, but received different amounts of guidance. Both groups received personal instructions and tips before starting their observations. The difference between the two experimental groups was that participants of the experimental group 1 had two additional small group meetings with an ornithological expert during their observation period, during which they could discuss their observations. Experimental group 2 did not have any meetings during the observation period, but heard a lecture about birds by an ornithological expert instead, before filling in the post-test. This kept instructional time similar. During data collection, due to Covid-19 face-to-face meetings were prohibited, and we had to stop data collection. Thus, we checked the already collected data and there were no statistical differences between both groups (data not shown). Therefore, we decided to merge them as the ‘guidance group’ and additionally recruited an online group, which did not receive any guidance, to assess the difference between guidance and no guidance.

Participants of the guidance group started their participation in the study with an on-site appointment, during which they were given personal instructions by an ornithological expert, as well as expert tips for birdwatching, before filling in the pre-test. The test contained the instruments described above (see 4.2.3) and was filled in using paper and pencil. There was no time limit. Afterwards, they received a bird identification book (Dierschke, 2017) and were instructed how to install and to use the app NaturaList, which is the app belonging to the online Citizen Science portal ornitho.de. After the end of this on-site appointment, participants started their three-week observation period. Participants were asked to report at least 30 observations during the three weeks.

The online group did not have an on-site-appointment and received information about the study and the Naturalist app via e-mail only. There was no instruction by an expert and no instruction how to use the app or the Citizen Science portal. However, the ornitho online portal provides enough detailed information on how to use it. The e-mail also contained an online link to the pre-test, which could be filled in on a computer or smartphone. The test contained the instruments described above (see 4.2.3). There was no time limit. Afterwards, participants received the same bird identification book as participants of the guidance group (Dierschke, 2017). After participants had filled in the pre-test, they started their three-week observation period. Participants were asked to report at least 30 observations during the three weeks.

During the observation period, participants of the guidance group received further support by an ornithological expert, either by meeting in small groups or by getting to know further information about birds by hearing a lecture (as explained above). Participants of the online group did not receive further guidance or information during the observation period.

After three weeks, participants of the guidance group had their last on-site appointment, during which they filled in the post-test. The test contained the instruments described above (see 4.2.3) and was filled in using paper and pencil. There was no time limit. Afterwards, participation was completed. Participants of the online group were sent an online link to the post-test, three weeks after filling in the pre-test. The post-test contained the instruments described above (see 4.2.3) and was filled in at a computer or smartphone. There was no time limit. Afterwards, participation was completed.

All participants, who had participated in all parts of the study (pre- and post-test, observations, possibly on-site appointments), took part in a raffle of eight wildlife camera traps after the study, which were sent to the winners by mail.

4.2.5 Statistical Analyses

We used SPSS 26 (IBM, Armonk, NY, US) to analyse the data. Data concerning number of species and number of observations reported have been log10 transformed prior to analyses. We ran a series of general linear models with repeated measures design (repeated measures analysis of variance) for variables that have been measured with a pre-/post-test design (content knowledge concerning birds, self-assessed bird species knowledge, scientific reasoning, specialization/commitment to birding). Number of observations and number of species observed during the observation period were assessed with univariate analyses of covariance where birding specialization and pre-test bird knowledge were entered as covariates. Partial eta-squared was used as a measure of effect sizes in all calculations.

4.3 Results

4.3.1 Content Knowledge

Descriptive values can be found in table 3.

The general linear model with repeated measures, which was calculated based on the results of the species knowledge test, revealed a significant main effect for time ($F(1,108) = 113.17$, $p < .001$, partial $\eta^2 = .51$), showing that both, the guidance group and the online group, achieved gains in content knowledge. There was neither a significant main effect for group ($F(1, 108) = 1.89$, $p = .172$, partial $\eta^2 = .02$), nor an interaction effect for group x time ($F(1,108) = 0.16$, $p = .693$, partial $\eta^2 = .001$). Changes in content knowledge concerning birds over time in the different groups are shown in figure 6.

Table 3

Means and standard deviations in content knowledge (bird species) and scientific reasoning

		Guidance Group			Online Group		
		N	Mean	*SD*	*N*	Mean	*SD*
Pre-test	Content Knowledge	46	0.61	0.24	67	0.59	0.22
	Scientific Reasoning	46	0.74	0.12	67	0.73	0.15
Post-test	Content Knowledge	43	0.76	0.19	67	0.70	0.18
	Scientific Reasoning	43	0.77	0.14	67	0.76	0.19

Notes. Possible values for content knowledge and scientific reasoning range from 0 to 1.

The achieved points in the knowledge test in the post-test showed a significant positive correlation with participant’s self-assessment about their species knowledge in the post-test, $r = .63$, $p < .001$.

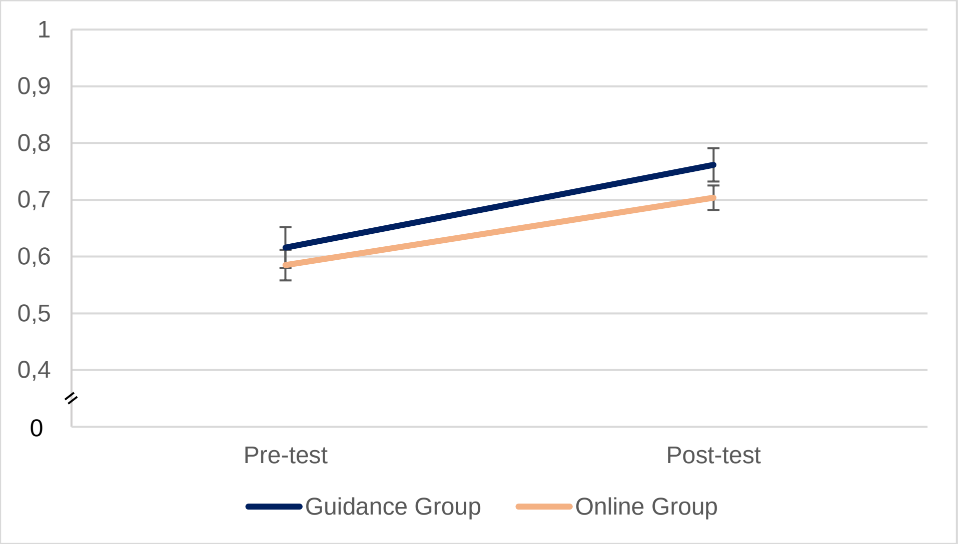

Figure 6. Changes in content knowledge (bird species) over time for guidance group and online group. Possible values ranging from 0 to 1. Bars indicate standard errors.

Concerning their self-assessment, participants of the guidance group reported they could identify on average 39.34 different bird species in the pre-test, which increased by almost 20 species in the course of the observation period, so that they reported being able to identify on average 59.14 species in the post-test. On the other hand, participants of the online group stated in the pre-test that they could identify on average 35.19 different species, which increased only by about 1, so that they stated they could identify only 36.51 different bird species on average in the post-test. Applying the general linear model with repeated measures to these data yielded a significant interaction effect for group x time ($F(1, 107) = 12.14$, $p = .001$, partial $\eta^2 = .10$), indicating that participants of the guidance group report a greater increase in self-assessed bird species knowledge than participants of the online group, who showed almost no increase at all. The model also showed a significant main effect for time ($F(1, 107) = 16.89$, $p < .001$, partial $\eta^2 = .14$), but not for group ($F(1, 107) = 2.42$, $p = .12$, partial $\eta^2 = .02$). Figure 7 illustrates the changes in self-assessed bird species knowledge for the two different groups. Although there was no group difference in the bird knowledge test, which contained 20 different bird species, the guidance group reported being able to identify a number of bird species in the self-assessment.

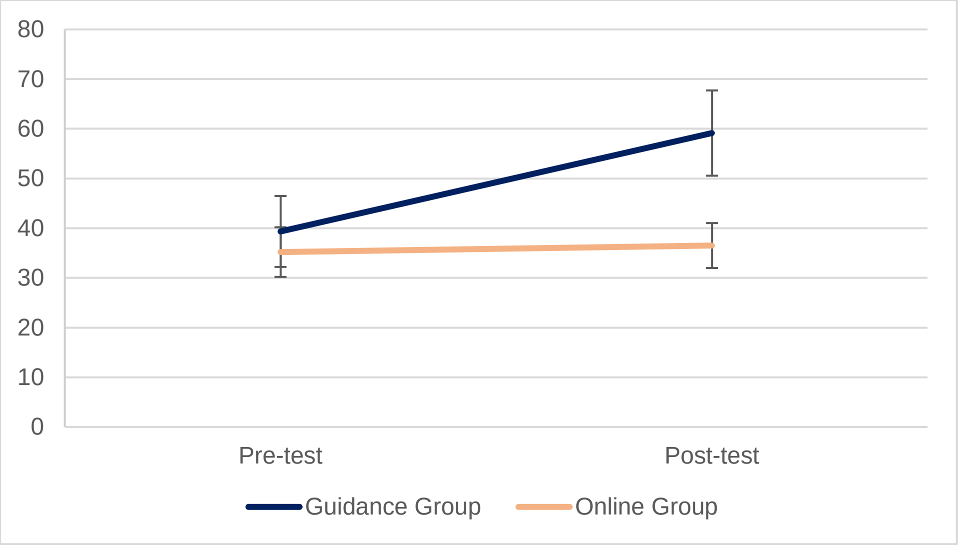

Figure 7. Changes in self-assessed bird species knowledge over time for guidance group and online group. Bars indicate standard errors.

4.3.2 Scientific Reasoning

Descriptive values are presented in table 1. The general linear model with repeated measures showed no significant effects. There was neither a significant main effect for time ($F(1, 108) = 3.71$, $p = .057$, partial $\eta^2 = .03$), nor for group ($F(1, 108) = 0.12$, $p = .726$, partial $\eta^2 = .001$), nor a significant interaction effect for time x group ($F(1, 108) = 0.04$, $p = .835$, partial $\eta^2 < .001$). However, we found a strong trend for time having an impact on scientific reasoning in both groups, as this main effect was on the verge of reaching significance.

4.3.3 Number of Bird and Species Observations

Descriptive data concerning the number of observations and the number of different species observed during the three-week observation period are presented in table 4. There was no significant difference between the two groups concerning the number of observations, $F(1, 107) = 0.74$, $p = .389$, partial $\eta^2 = .01$. Birding specialization, which was added as a covariate, did not have a significant effect either, $F(1, 107) = 1.07$, $p = .304$, partial $\eta^2 = .01$. However, pre-test bird knowledge (test score), which was added as a covariate as well, showed a significant effect on the number of observations, $F\ (1, 107) = 4.38$, $p = .039$, partial $\eta^2 = .04$). Thus, better skilled birdwatchers also reported more observations.

Table 4

Number of observations and number of different species observed

	Guidance Group			Online Group		
	N	Mean	*SD*	*N*	Mean	*SD*
Number of observations	46	54.41	51.87	67	50.97	70.74
Number of different bird species	46	21.57	12.96	67	16.24	12.17

Notes. Number of observations and number of different species observed were log-transformed before analysing them. For better illustration, data presented in the table above are before log-transformation.

Concerning the number of species observed during the three-week period, the univariate analysis of covariances yielded a significant difference between the two groups with participants of the guidance group reporting significantly more different bird species, $F(1, 107) = 6.97$, $p = .01$, partial $\eta^2 = .06$. This result fits with the findings reported above concerning participants' self-assessment, in which participants of the guidance group showed a greater increase in species knowledge as well (see 4.3.1, Content knowledge). The covariate birding specialization showed no significant effect ($F(1, 107) = 0.67$, $p = .415$, partial $\eta^2 = .01$), whereas the covariate pre-test bird knowledge (test score) had a significant effect on the number of different species reported ($F(1, 107) = 14.12$, $p < .001$, partial $\eta^2 = .12$).

During the two three-weeks observation periods, a total of 137 bird species were reported by more than 250 observers in the region. Out of these, 133 were reported from October 2019 and 105 from January 2020 (Anthes et al., 2019, 2020). This was significantly different (Sign Test: $Z=-4.50$, $p<0.001$). This is also reflected by the line transect counts in Weggental. During weekly counts/observations in the 'Weggental' a total of 34 species were counted in January 2020, while 46 were counted in October 2020 (Randler, 2020). Based on these data, it is highly likely that more bird species were present in the main study area in October than in January. Thus, the online group would have had the opportunity to observe even more different bird species, however participants of the online group have observed significantly fewer different species in October than the guidance group in January, suggesting that the guidance group invested more or acquired a higher knowledge.

4.3.4 Birding Specialization / Commitment to Birding

Descriptive values are presented in table 5.

Table 5

Means and standard deviations in birding specialization

	Guidance Group			Online Group		
	N	Mean	*SD*	*N*	Mean	*SD*
Pre-test	45	4.42	1.36	67	4.24	1.02
Post-test	43	4.80	1.12	67	4.44	1.00

Notes. Possible values for birding specialization range from 0 to 7.

Applying the general linear model with repeated measures to the data yielded a significant main effect for time ($F(1, 107) = 12.69$, $p = .001$, partial $\eta^2 = .12$), indicating that both groups increased in commitment or birding specialization during the participation in the Citizen Science project. There was no significant main effect for group ($F(1, 107) = 2.27$, $p = .135$, partial $\eta^2 = .02$), nor a significant interaction effect for group x time ($F(1, 107) = 0.19$, $p = .661$, partial $\eta^2 = .002$). Changes in birding specialization are illustrated in figure 8. Thus, both groups increased their commitment.

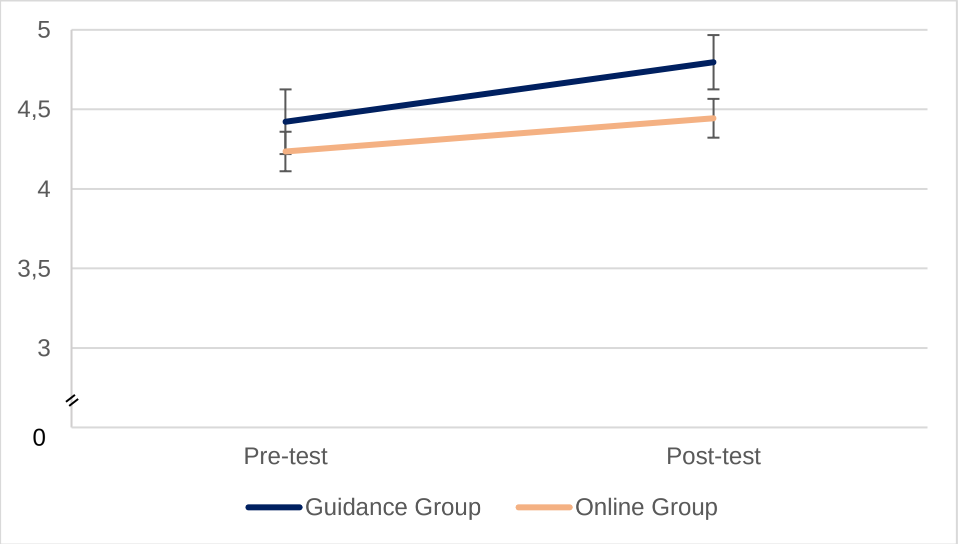

Figure 8. Changes in birding specialization over time for guidance group and online group. Possible values ranging from 0 to 7. Bars indicate standard errors.

4.4 Discussion

Both interventions improved species knowledge and birding specialization, while scientific reasoning could not be significantly improved. In terms of effect sizes, the highest effect size was found in species knowledge, followed by changes in birding specializations. This reflects the well-known fact that factual knowledge is easiest to improve compared with attitude.

4.4.1. Knowledge

Changes in the number of species being able to recognize showed that both interventions lead to a higher factual knowledge in bird species. Knowledge gain has been reported in some studies including a Citizen Science programme, both in students (Schneiderhan-Opel & Bogner, 2020), and adults (Jordan et al., 2011; Parrish et al., 2019). Similarly, science teachers assessed their bird species knowledge as higher after an intervention dealing with wild birds and ringing (Ortiz et al., 2020). Similarly, Brossard et al. (2005) used a self-assessment item about bird biology, and found an increase in self-assessment after a program concerned with observing nest boxes and birds, recording data according to the different protocols (Brossard et al., 2005). Thus, encountering birds in nature and participating in Citizen Science programs foster identification skills as measured by standardised tests as well as the self-assessment of ability.

4.4.2 Birding Experience

The guidance group was able to observe more species during their three-week observation period. This might be attributed to the guidance. Parrish et al. (2019) emphasized the importance of expert training for long-lasting Citizen Science programs. Probably as a result of the higher number of species observed, the guidance group also scored higher in the self-assessment of their bird species knowledge. These two aspects may be linked with each other.

Another reason for the differences in the numbers of observed species could be the different seasonal environments, which have an influence on the number of different bird species present in the observation area. The guidance group participated in the Citizen Science project during January, while the online group made their observation from mid-October onwards. By chance, one would have expected a higher number of species being observed by our online group in October, but the contrary was the case. Thus, in October, more species were present in the study area, but fewer were found and reported by the participants. Thus, the difference in activity cannot be linked to the species being present.

Another question is why the participants reported more species, assessed their own knowledge higher but did not perform better in the post-test. This may be owed to the number of species used in the test (N = 20), because we only asked for 20 species, and not for the possible 137 species that could have been encountered during the study period. Thus, bird species knowledge may have increased in species that had been observed but not asked for in the test. This shows that the self-assessment of one's own species knowledge is an important aspect. Both seem valuable because they were significantly correlated. This highlights on of the strengths of our study, which integrated the factual knowledge test, as well as a self-assessment in order to shed a light on two different facets of knowledge concerning different bird species.

4.4.3 Birding Specialization/Commitment

Birding expertise has been identified as an influential factor in the attrition to longer-lasting Citizen Science projects, with higher qualified birdwatchers staying longer in a specific program compared to the ones with lower specialization and commitment (Parrish et al., 2019).

To our knowledge, this is one of the few studies that longitudinally assessed changes in birding specialization, albeit over a short time period of only three weeks. Scott and Lee (2010) analysed birdwatchers over a five-year period. Their findings showed that some birdwatchers increased their specialization, but others remained on the same level or even declined their activity. The most influential factors on specialization were support from family members and retirement (Scott & Lee, 2010). Other studies in recreation research longitudinally assessed the specialization of boat owners (Kuentzel & Heberlein, 2006) comparing data from 1975, 1985 and 1997. These authors reported that specialization progression was the exception rather than the rule among boaters. However, short-term changes in specialization, as in our study, have not yet been addressed. The studies suggest that during birding initiation, an increase in specialization is the rule, while in the long run, e.g. over some years, progression, decline or stability can occur. Also, the increase in our study can be attributed to the program itself, but concerning longer durations, other factors seem to be more influential (see above, Scott & Lee, 2010).

4.4.4 Scientific Reasoning

Scientific reasoning seems more difficult to develop than species or factual knowledge. For example, Jordan et al. (2011) reported that knowledge (in this case of invasive plants) increased, but the participation in their three-day program was insufficient to increase understanding of how scientific research is conducted. Similarly, Brossard et al. (2005) reported an increase in knowledge about bird biology, but no increase in understanding of scientific methods. However, Fujitani et al. (2017) reported not only higher knowledge, but also a behavioural change when recreational anglers were involved in the experiments of scientists. Concerning scientific reasoning skills, there are few studies that report positive effects of a participation in Citizen Science projects, for example, Golumbic et al. (2020) found an improvement in data interpretation skills of participants. Further Trumbull et al. (2000) reported scientific thinking processes of participants, as well as Evans et al. (2005), who described basic scientific reasoning processes of participants, like asking questions, but also on a higher level, including drawing conclusions from their observations.

In line with most previous studies, we could not find a significant increase in scientific reasoning in our sample, which supports the assumption that improvements in scientific reasoning do not just emerge as citizens take part in a Citizen Science project, but need more targeted support (Bonney et al., 2016). Nevertheless, we found a strong trend for scientific reasoning improving in the course of the project participation for participants of both groups. Each group showed an improvement of three percentage points in scientific reasoning. Perhaps a longer observation period would have yielded significant changes. Another informal assessment for scientific reasoning (e.g., qualitative interviews) might have been useful to gather more information about whether the participants have actually improved in scientific reasoning or whether the trend we found might perhaps just be a retest effect.

In general, evidence concerning improvements in scientific reasoning after partaking in a Citizen Science project is still very rare. Only few Citizen Science projects assess scientific reasoning, and if they do, mostly indirect (Stylinski et al., 2020). Solely 13% of the articles being part of the review of Stylinski et al. (2020) used a formal assessment for scientific inquiry skills. Thus, our study adds new evidence about the relationship of Citizen Science projects and scientific reasoning to the little amount of studies currently available. Still, further research is needed to get a clearer picture of the extent to which participation in a Citizen Science project provides opportunities or obstacles to the development of scientific reasoning.

4.4.5 Limitations and Strengths

The guidance group consisted of two initially different groups. Study design could have been improved by having all participants of the guidance group receive the same sort of guidance. However, as there were no significant differences between the two initial groups, the amount of guidance does not seem to influence the studied variables. It is more likely that the mere presence or absence of guidance influences some outcomes for participants of a Citizen Science project, as for example the number of different species observed. Another limitation of the study are the different seasons, during which the participants of the different groups made their observations, as different seasons have a major impact on the number of different bird species present in the observation area. However, as mentioned above (see 4.3.3 Number of Bird Species and Observations), participants of the online group, who would have had the opportunity to report more different bird species, even reported fewer different species. This makes our finding concerning the number of different bird species observed even more remarkable.

Finally, the study could have been improved by having two groups in the same size. In our sample, the online group (N = 67) consisted of more participants than the guidance group (N = 46). This is due to the longitudinal design of the study, which can lead to dropouts that cannot be controlled for in advance. Moreover, having groups in exactly the same size is always difficult in field studies. Yet, the benefits of field studies outnumber the disadvantages, as field studies yield results of high ecological validity, which is particularly important in the context of Citizen Science research, because Citizen Science does not occur in lab situations. Another strength of the study is its longitudinal design, albeit over a short period of three weeks. Nevertheless, even though the participation period was quite short, we were able to show significant changes regarding bird identification skills and commitment to birding. As we used standardized tests to assess knowledge and scientific reasoning, we were able to add objective, quantitative results to the study base, which is especially with regard to scientific reasoning still quite sparse.

4.4.6 Conclusion

Both interventions increased factual species knowledge and changed the self-assessment of birding specialization. Thus, both interventions can be seen as powerful in accompanying birders starting to bird within a Citizen Science framework. This is an important result for Citizen Science project leaders, as it shows that participants even profit from participating in a project, if it is not possible to ensure a high degree of involvement or a high amount of guidance. Hence, our study results may help designing future Citizen Science projects in an effective way to achieve both the goal to gain scientific knowledge and the aim to achieve positive outcomes for participants.

5. Results and General Discussion

This dissertation project aimed at investigating whether scientific reasoning and factual knowledge gain can be improved by pursuing a Citizen Science approach to learning. To do so, two studies in different contexts were conducted.

In the study which was carried out in the school context with students of grade 5 and 6, the participation in the Citizen Science project was combined with methods based on theoretical assumptions of inquiry-based learning, situated cognition and out-of-school learning. Results revealed that students who participated in a Citizen Science project during the teaching unit showed a greater increase in scientific reasoning compared to students who received a teaching unit on the same subject, but did not go on a field trip and did not take part in a Citizen Science project. Concerning factual knowledge gain and motivation, there were no differences between the two groups.

In the second study, which was conducted with adults, participants were asked to collect data for a Citizen Science project during their free time over a period of three weeks. The participants received different amounts of guidance and were either guided by a professional or participated mostly on their own. Results did not show any differences in scientific reasoning or factual knowledge gain in the administered quantitative tests, however, participants receiving a higher level of guidance reported a higher level of factual knowledge in a self-report after the three-week observation period.

Implications of these results on the six main question posed at the beginning of this thesis are discussed in the following sections.

5.1 Citizen Science and Scientific Reasoning of Students

The first main question which this dissertation project wanted to address was whether a Citizen Science approach would be suitable to train scientific reasoning skills of students. As the results of the first study with students show, students who participated in the Citizen Science project did improve their scientific reasoning skills in a greater way than students who did not participate in a Citizen Science project.

This effect cannot only be attributed to the fact that the students learned according to an inquiry-based learning approach and were thus able to work authentically like professional scientists, since the group that did not participate in the Citizen Science project also worked accordingly to an inquiry-based learning approach inside the classroom. Rather, there seems to have been the additional positive effect that the students were able to collect their own data authentically in the field, which will ultimately be used for actual research findings within the framework of the Citizen Science project which was embedded in the teaching unit. Thus, due to being a part of the Citizen Science project both during the field trip and inside the classroom, the Citizen Science project can serve as a useful tool in bridging the gap between actions taken outside of school and knowledge acquired inside the classroom (Brown et al., 1989), thereby simplifying knowledge transfer and reducing the probability of the generation of inert knowledge (Whitehead, 1967).

Hence, a Citizen Science approach is indeed suitable to train scientific reasoning skills. These are major, novel findings with important implications for schooling practice, as the current approach to learning, which takes place predominantly within the classroom, has severe difficulties in fostering scientific reasoning of students (Pant et al., 2013). Since the demands of today’s and tomorrow’s society ask for citizens who have a high level of scientific reasoning and problem solving skills (Hill, 2007; National Research Council, 2008), it is a major goal of science education at school to have the students become scientifically literate citizens of tomorrow with strong scientific reasoning skills (National Research Council, 2012).

5.2 Citizen Science and Factual Knowledge Gain of Students

The next question this thesis sought to answer was whether a Citizen Science approach would be suitable to improve factual knowledge of students. The results in the first study, which was implemented in the school context, showed no differences regarding gains in factual knowledge between the students who took part in the Citizen Science project and those who did not. Therefore, there seem to be no additional benefits on factual knowledge if a Citizen Science project is included in the teaching unit.

A possible reason for this result might be that unlike the improvement of scientific reasoning, the acquisition of factual knowledge is not such a complicated process that it requires special and targeted training. Therefore, mere out-of-school learning that is not embedded in a larger teaching project is also sufficient to improve subject knowledge. As stated previously (see 1.3), examples of the effectiveness of out-of-school learning on factual knowledge gain of students can be found in the literature, for example Randler et al. (2005) or Seybold et al. (2014). Similarly, the use of the inquiry-based learning approach seems to be sufficient to achieve knowledge gains. The effectiveness of the inquiry approach for knowledge growth has been demonstrated several times (Furtak et al., 2012). In line with these results, our study, in which the control group that only learned inside the classroom also worked according to the inquiry-based learning approach, also showed a positive effect of inquiry-based learning on the students' factual knowledge growth. The additional participation in a Citizen Science project therefore does not seem to be necessary to generate gains in factual knowledge. Nevertheless, as both interventions, no matter if a Citizen Science project was included or not, were successful in achieving factual knowledge gains, a Citizen Science approach can be implemented at schools in order to foster scientific reasoning without harming gains in factual knowledge.

A reason why the Citizen Science approach was not superior to learning inside the classroom might be found by looking at non-formal learning from the perspective of social constructivism. According to social constructivism, knowledge is mainly generated through dialogue of different people (Vygotsky, 1978). Social constructivism serves well as a framework for construction of knowledge in a non-formal learning environment, as each individual has different prior experiences and a different understanding of the observed objects, hence, knowledge of each individual grows by interacting with others (Gilbert & Priest, 1997). During the field trip in the first study, students were working in small groups and were observing the birds together. Hence, the knowledge was generated by dialogue in small groups of students who received only a small amount of guidance by a teacher, as the aim was to have the students themselves think about their hypotheses and thus train their scientific reasoning skills. However, concerning the construction of knowledge, adults can often provide even better input in conversations to facilitate knowledge generation for students (Crowley & Callanan, 1998). Therefore, students might have needed a little more input or guidance by an adult to not only improve their scientific reasoning but also to increase their factual knowledge to a greater extent. Alfieri et al. (2011) already reported the huge importance of the right amount of guidance for the success of inquiry-based learning: if students are receiving the wrong amount of guidance, inquiry-based learning is less effective than other learning forms. Only with the correct amount of guidance, inquiry-based learning is superior to explicit instruction. However, effects of guidance only seem to influence performance success and not learning outcomes or domain knowledge (Lazonder & Harmsen, 2016).

Another factor, that might have influenced factual knowledge gain of the students who were partaking in the Citizen Science project, is the novelty phenomenon. Research has shown that the novelty of a surrounding highly influences the behaviour of individuals, especially in school groups (Falk, 1983; Falk & Balling, 1982). This leads to a high engagement with the new environment and a low engagement with the task (Balling, as cited in Eshach, 2007), resulting in a lot of off-task behaviour (Lucas, 2000). If the feeling of novelty is too high, children might even experience anxiety (Eshach, 2007). By choosing field trip sites close to the students’ schools in the first study, this problem was minimalised. Still, this cannot ensure entire familiarity of every student with the surroundings during the field trip. Moreover, as the students of the sample of the first study were still quite young (age M= 11.21), they might have been distracted quite easily during a field trip, even if the surrounding environment is familiar to them.

However, to sum this part about effects on factual knowledge up, it is important to note that the students showed significant knowledge gains, regardless of the group they were in. Therefore, the possibly influential factors discussed above are merely to be seen as reasons why the students who participated in the Citizen Science project did not enlarge their knowledge in a greater way than the control group.

5.3 Citizen Science and Motivation of Students

The third main question which was investigated with this dissertation project was whether a participation in a Citizen Science project would lead to a higher level of motivation in the students, which should show in higher levels of feelings of autonomy and thus, intrinsic motivation.

Results of the first study did not show any differences regarding motivation between the students who participated in the Citizen Science project and those who did not. As mentioned before (see 1.3), especially the feeling of autonomy is essential for developing intrinsic motivation (Ryan & Deci, 2002). Hence, the measure that was used for motivation in the first study, the KIM scale (Wilde et al., 2009), included a distinct subscale for autonomy, labelled perceived choice. In the sample surveyed for the first study, intrinsic consistency for this perceived choice scale was quite low, only being .41. This is a surprisingly low intrinsic consistency, since Wilde et al. (2009) report Cronbach’s Alpha being .75 or .79 for this scale. Perhaps there have been some problems with this subscale in the specific sample of the first study which resulted in the students’ feeling of autonomy not being correctly reflected by the perceived choice scale in this sample. Without having a reliable measure of the feeling of autonomy, evaluation of students’ motivation based on self-determination is thus difficult. Nevertheless, the KIM scale usually is a suitable instrument to assess motivation, since it has successfully been used in many different school settings before and a comparative factor analysis of 1.700 secondary school students showed good fit indices (Wüst-Ackermann et al., 2018).

Nevertheless, results of the first study revealed a high motivation, especially regarding interest for all of the students, no matter if they were partaking in a Citizen Science project or not, which showed in values being M= 4.13 and M= 4.07

respectively of a maximum of five points. Therefore, if the all the students had high levels of interest and thus motivation, detection of differences between the groups is hard.

Another possible reason for the missing differences in perceived choice as an indicator for motivation might be that the concept of self-determination might not be the theoretical concept, which best fits the emergence of motivation in the first study of this dissertation project. As stated previously (see 1.3), feelings of personal utility value as included in the expectancy-value theory might be important for the development of motivation (Eccles, 1983; Hulleman & Harackiewicz, 2009) during a participation in a Citizen Science project in an out-of-school learning context close to the students' respective schools as well. Thus, it might have been reasonable to include items that aim at assessing the personal utility value and the personal involvement in the task in the assessment of students' motivation. In addition to the enhanced insights into the motivational experience of the students in this specific study, it would also have provided a clue as to which factors are central to the development of motivation in the context of participation in a Citizen Science project as part of a larger teaching unit. Further research is needed in order to get a clearer view of relevant factors of students' motivation in this context.

However, similar to the results concerning factual knowledge, it is important to highlight that the students partaking in the Citizen Science project do not show a lower level of motivation, but that there is simply no difference between the groups. Thus, it should be noted once again that a Citizen Science approach can be implemented at school without harming students' motivation. Moreover, the lack of differences regarding motivation between the students who participated in a Citizen Science project and those who did not only further suggests that the improvements in scientific reasoning are due to the specific characteristics of the Citizen Science approach. Thus, as there were no group differences regarding motivation, motivation cannot be responsible for the greater improvement in scientific reasoning of students who were partaking in a Citizen Science project. Therefore, these improvements seem to have been specifically trained by the cognitive demands of the task, which authentically reflect the working methods of a professional scientist and thus, are suitable to foster scientific reasoning.

5.4 Citizen Science and Scientific Reasoning of Adults

The fourth major question which was addressed in this thesis was whether participation in a Citizen Science project would improve scientific reasoning of adults. This question was investigated during the second study, which worked with adult participants. Results of the second study showed no significant effect of the participation in a Citizen Science project on scientific reasoning, neither if participants were receiving a high amount of guidance, nor if they were receiving no guidance. However, there was a strong trend for an improvement of scientific reasoning for all participants, which might have reached significance, if the duration of the study would have been more than three weeks.

This result is in accordance with previous findings, which state that improvements in scientific thinking do not simply happen due to participants’ exposure to a real scientific study (Bonney et al., 2016). However, similar to scientific reasoning in the schooling context, there has to be a special emphasis on scientific reasoning or a targeted engagement of the adult participants in the Citizen Science project in order to achieving beneficial outcomes regarding scientific reasoning. In line with this statement, most of the studies dealing with scientific literacy in the context of Citizen Science projects could not find any improvements regarding scientific reasoning due to participation in the Citizen Science project (for example Jordan et al., 2011; Brossard et al., 2005). Moreover, the assumption that scientific reasoning can indeed be improved under certain circumstances during participation in a Citizen Science project, is, at least for the schooling context, supported by the findings of the first study of this dissertation project as well, as they showed that scientific reasoning skills improve, if a teaching unit puts an emphasis on training scientific reasoning within a Citizen Science approach. Thus, in principle, Citizen Science projects have the potential to be useful in improving scientific reasoning.

In accordance with the findings of the first study, some previous studies hint at a possible improvement of scientific reasoning skills during participation in a Citizen Science project for adults as well. For example, Trumbull et al. (2000) analysed letters of participants and reported scientific thinking processes for a majority of the participants. Similarly, Evans et al. (2005) derived scientific thinking patterns from respondents' questions about the Citizen Science project. However, quantitative evidence for these conclusions is still lacking. The second study that has been conducted for this dissertation project could not confirm the qualitative analyses regarding scientific reasoning of adults who partake in a Citizen Science project. However, since the results showed a strong trend towards an improvement of scientific reasoning, it would be premature to dismiss the possibility that a Citizen Science project might have a positive impact on scientific reasoning. The trend which was found in the second study might reflect the changes in scientific thinking processes reported in the qualitative analyses by Trumbull et al. (2000) and Evans et al. (2005). Yet, these possible changes in participants' thinking patterns were apparently not sufficient to achieve significant changes in the quantitative tests. Still, researchers should keep Citizen Science projects in mind as possible useful tools to train scientific reasoning of adults.

5.5 Citizen Science and Factual Knowledge Gain of Adults

The fifth main question, which this dissertation project wanted to address was whether participation in a Citizen Science project would lead to gains in factual knowledge for adults. The results of the second study revealed that the participants who took part in the Citizen Science project achieved significant knowledge gains, no matter whether they received a high amount of guidance or no guidance at all. This shows that the mere participation in a Citizen Science project is sufficient to improve factual knowledge and, in contrast to scientific reasoning, no targeted intervention is needed in order to support knowledge acquisition.

This is in line with the results of the first study implemented in the school context, in which the results did not show the need of a targeted training in order to heighten the gain of factual knowledge, either. As argued previously (see 5.2), acquisition of factual knowledge might not be as complex as improvement of scientific reasoning and thus, needs less targeted training and less support. This hypothesis is supported by several results of previously conducted studies with adults in the context of Citizen Science. For instance, Crall et al. (2013) and Jordan et al. (2011) both reported gains in factual knowledge for participants of a Citizen Science project concerning invasive plants. However, in both studies the results could not show changes regarding scientific literacy. Likewise, Brossard et al. (2005) reported the same results in the context of birding: factual knowledge concerning birds increased, yet there were no changes in scientific literacy of the participants.

Hence, Citizen Science projects are already known as being suitable and useful when aiming at factual knowledge gains of participants. The second study of this dissertation project was able to add further evidence to this data base and to confirm previous results that show an increase in factual knowledge, yet no improvement regarding scientific reasoning.

5.6 Citizen Science and Guidance

The final question, which this thesis investigated, was whether different amounts of guidance would influence outcomes of adult participants of a Citizen Science project regarding improvements in scientific reasoning and factual knowledge gains. The results of the second study of this dissertation project did not show different outcomes for the participants depending on the level of guidance they received, neither concerning scientific reasoning, nor with regard to factual knowledge gains as assessed with a quantitative factual knowledge test. However, participants who had received a higher level of guidance, reported a greater variety in observed bird species to the Citizen Science database and additionally they stated in the self-assessment that they knew a greater variety of bird species than the participants who had not received any guidance.

Since there are differences in the behavioural data and in the self-assessment depending on the level of guidance received, guidance might not have been sufficient to induce personal involvement or engagement in the Citizen Science project and thus, did not lead to different outcomes concerning scientific reasoning and factual knowledge. However, it might have influenced participants’ motivation. If participants with a higher amount of guidance had a higher level of motivation one would expect them to report more observations and a greater variety of species. Even though the number of reported observations did not differ between participants receiving guidance and those who did not, the higher number of different species, which was reported by the group of participants that received guidance, might represent the will to deliver data of higher quality to the Citizen Science database, which could result from a higher level of motivation. However, since the assessment did not include a measure for motivation in the second study, this connection can only be speculated on and thus should be analysed in more detail in future studies.

Since guidance does not seem to be sufficient to generate involvement in participants, further options should be considered on how to induce personal involvement in order to improve the results of participation in a Citizen Science project for the participants. For this dissertation project, it was originally planned to conduct a third study, again in a schooling context. One of the aims of this study would have been to gain further insight on the role of involvement for the results of Citizen Science projects for the participants. For this purpose, it was planned to establish involvement through a personal connection of the students to the Citizen Science project. Following theoretical assumptions of Eccles' expectancy-value theory (1983), a higher level of involvement was expected to result from this personal connection. Since engagement has already been discussed as a way to multiply the positive effects of a Citizen Science project for the participants (Druschke & Seltzer, 2012), it was subsequently expected that this increased involvement would have positive effects on improving scientific reasoning and increasing factual knowledge gain. However, due to the Covid-19 pandemic and the associated school closures in Germany, it was not possible to implement this school study within the time frame of the dissertation project. Therefore, these additional insights into the possibilities of involvement in the context of a Citizen Science project cannot be reported here.

Another possibility to induce engagement, which has already been better studied scientifically, is to vary the degree to which participation is possible in a Citizen Science project (Shirk et al., 2012). According to Shirk et al. (2012), there are mainly three different degrees of participation that indicate the extent to which participants are involved in the Citizen Science projects. In their model, projects which only allow for a low degree of participation are called contributory project. The characteristic of these projects is that participants are only part of a small part of the scientific investigation. The development before and the analysis of the data after the project are completely done by professional scientists. The citizens who participate in the project are only part of the data collection phase and add their data to the Citizen Science database. Projects which allow for a medium degree of participation are called collaborative projects. In these projects, participants can not only be part of the data collection process, but can also report back to the scientists with their experiences and problems, which in turn will be taken into account when refining the project or analysing the final data. Lastly, if scientist wanted the participants to have a maximum degree of participation in the Citizen Science project, they would aim for a co-created project. This type of Citizen Science project is characterised by the fact that scientists and citizens are cooperating in nearly every step during the scientific investigation, starting with a joint development of the study and ending with a joint discussion of the study results and the elaboration of new research questions. Co-created projects have already been shown to have a positive impact on scientific reasoning and participants' comprehension of scientific investigations (Shirk et al., 2012).

Since the Citizen Science project, in which the participants of both studies of this dissertation project took part, would be a contributory project following the model of Shirk et al. (2012), it might have been difficult to induce a higher level of involvement in participants, perhaps even if guidance could be a useful tool to do so in collaborative or in co-created projects. Thus, future research is needed to clarify if a higher amount of guidance could be a useful way in order to even further enhance participants' involvement in collaborative or co-created Citizen Science projects. Moreover, the amount of guidance received could still be an effective tool for enhancing the effectivity of Citizen Science projects on scientific reasoning, even if it is not useful to enhance involvement. As already known from research in the school context, guidance is an important way to achieve success in inquiry-based learning (Lazonder & Harmsen, 2016). Therefore, an adequate amount of guidance in Citizen Science projects in the school context might also be helpful to further promote outcomes in terms of scientific reasoning and factual knowledge gains, even without achieving an increased level of involvement. The results on guidance in an adult sample outside the school context are therefore not necessarily transferable to students. Therefore, it would be useful to conduct future research on the relationship between guidance and the outcomes of a Citizen Science project in the school context.

5.7 Differences between Citizen Science at School and for Adults

Next to a potentially different impact of guidance, there are some other factors that might have led to less beneficial outcomes for adult participants in the second study who did not show any changes in their scientific reasoning skills compared to the participating students in the first study whose scientific reasoning improved significantly.

An additional reason for these findings could be the study design and the way the Citizen Science project was integrated into the study process. Bonney et al. (2016) compared four categories of Citizen Science projects that differ in their design and thus in participants' engagement. They examined data collection projects, which involve participants in the process of collecting data for scientific research that will be analysed by professional scientists, data processing projects, which do not involve participants in the data collection phase, but rather in the analysis and examination of given data sets, curriculum-based projects, which are typically designed for students or youths, are clearly structured, accompanied by an adult and often incorporated in a national or regional school curriculum, and finally, community science projects, which place an emphasis on local issues and aim at having an impact on local decision-makers or stakeholders.

Even though the participants of both studies reported in this dissertation project collected data for the same Citizen Science project, the surrounding context, in which participation was based, differed in many ways. In the first study, participation was embedded in an entire teaching unit. Students received a lesson before and after collecting data for the Citizen Science project and were supported by research assistants and teachers at all times who they could ask for help or further guidance. Moreover, their collected data was not only used to add to the Citizen Science database and to be used for professional scientific investigations, but also for the students' own question and hypotheses, which were worked on after the data collection. Thus, this participation was part of a curriculum-based project. On the other hand, participants of the second study participated mostly on their own and received only a predetermined amount of guidance, if any at all. Furthermore, participants did not conduct any further analyses with their collected data and merely added their observations to the online Citizen Science database. Hence, the adults in the second study participated in a data collection project.

Depending on the category into which a project is put, Bonney et al. (2016) report different outcomes for participants. For participants of data collection projects, increases in content knowledge can be found (e.g., Jordan et al., 2011). However, only few studies report changes in the context of scientific reasoning (e.g., Trumbull et al., 2000). However, in curriculum-based projects, participants show increased content knowledge, as well as improvements in understanding the way in which scientific research is conducted as an indicator for improved scientific reasoning skills (Thompson, 2007). Therefore, the findings of the two studies of this dissertation project are in line with the results of the analyses of Bonney et al. (2016), as there was a significant increase in scientific reasoning in the curriculum-based project (study 1), but not in the data collection project (study 2).

Hence, participation in a data collection Citizen Science project, as in the second study of this dissertation project, might have similar effects on participants' outcomes as mere out-of-school learning has on students, which results in gains concerning factual knowledge (e.g., Seybold et al., 2014, see 1.3), but is often not sufficient to train scientific reasoning skills, as it might be regarded as an add-on activity, which does not lead to a deeper engagement with the scientific process for the students, or, in this case, the participants of the Citizen Science project.

Furthermore, another bias that might have influenced the different results of the two studies regarding changes in scientific reasoning is a possible self-selection of the adult participants in the second study. Although participation was of course voluntary in both studies, it can nevertheless be assumed that participants in den second study would show a higher interest in science and scientific investigations, as well as birding, compared to the students, who took part in the first study. Since the recruitment for participants in a school study begins with a request to the school's headmaster and afterwards with a request to the teachers, who would be involved in the study, the sample is already pre-selected when the actual participants, the students, are asked about their interest in participating. This means that students who would be very interested in participating in the study may not even have the opportunity to do so, because their headmaster or teacher has decided against their school or class taking part. Thus, in comparison to an open recruitment of an adult sample, a school study cannot be offered to all potentially interested students for participation. Moreover, this effect also occurs in reverse for the students in the participating classes. Since the project was selected by the headmaster or the teacher on behalf of the classes, it cannot be said that the students who ended up participating in the study would have taken part in the Citizen Science project of their own accord. As already mentioned, participation was of course voluntary for the students and no student took part without a written and signed declaration of consent from both him or her and a corresponding parent or guardian. However, other motivating factors might have played a role for taking part in the study, compared to a case in which participants come forward of their own accord in order to take part in a Citizen Science project in their free time. Hence, the adult sample in the second study may already have shown a greater interest in scientific questions and methods before the project than the participating students in the first study. Therefore, the adults in study two may already have had elevated scores in scientific reasoning at the beginning of the project compared to the overall population. This could explain the lack of significant improvement in scientific reasoning in the second study, as Citizen Science projects might be more suited to improve scientific reasoning among participants with little previous exposure to scientific enquiry or only an average interest in scientific research and scientific knowledge. However, since this dissertation project compares an adult sample to a student sample that differ in several ways other than a possible self-selection effect, this consideration can only be regarded as a hypothesis that needs further scientific testing, and not as a firm conclusion.

5.8 Strengths and Limitations of the Dissertation Project

This dissertation project provides new insights into scientific reasoning and its development in relation to participating in a Citizen Science project. The implementation of a Citizen Science project in a larger, structured teaching unit provided new insights into how scientific reasoning can be better promoted in the school context. These are essential findings with important implications for practice, as scientific reasoning is often not adequately trained in school contexts, which often only refer to the retrieval and application of knowledge (Osborne, 2013). However, scientific reasoning is one of the key competencies that students should not only develop during their school education (National Research Council, 2012), but also one of the competencies that is essential for successful participation in today’s and tomorrow’s society (Gilbert, 2005; Hill, 2007). The fact that this dissertation project was able to demonstrate a method through which scientific reasoning can be successfully trained without reducing students' gains regarding factual knowledge or students’ motivation is a central strength of this dissertation project with important results, both for research and practice.

Furthermore, as the potential of Citizen Science projects for improving scientific reasoning has been investigated from different perspectives in this dissertation project and therefore results from both the school context and with adults outside the school context can be reported, a differentiated picture of Citizen Science projects and their strengths and limitations in terms of improving scientific reasoning can be gained. Hence, implications can not only be derived for school practice, but also for future planning of Citizen Science projects outside of schools, so that maximum beneficial outcomes can be achieved for participants. In addition, for the first time, insights into the effects of guidance on factual knowledge gain and scientific reasoning in adult participants of a Citizen Science project could be gained.

However, this dissertation project has some limitations as well. Even though the new teaching concept which included a Citizen Science project in order to train scientific reasoning skills was successful in doing so, the implementation of this teaching unit into the daily school routine comes with an increased effort for the schools and teachers. For example, a suitable Citizen Science project has to be found and an out-of-school teaching unit, as well as the travel to and from this out-of-school learning location, have to be organised. Even though the results of the first study strongly suggest that this effort is worthwhile, its practicality may be limited. Therefore, as mentioned before, it had been planned to conduct a third study for this dissertation project, investigating participation in a Citizen Science project only within the school campus. As this study could not be conducted within the framework of this dissertation project due to the current Covid-19 pandemic, important findings are missing from the final conclusions in order to provide a potential opportunity for a more practical approach to training scientific reasoning by partaking in a Citizen Science project.

5.9 Outlook for Future Research

This dissertation project offers several points to build on in future research. Especially the role of guidance could not be clearly determined yet. Hence, future research could investigate the effectiveness of guidance in the school context to find out, whether the results obtained with adults in our second study are transferable to students. Moreover, the design of Citizen Science projects should be investigated further in future studies. Research concerned with this question could take a look at a possibility to design Citizen Science projects for adults similarly to curriculum-based projects at school in order to find out, whether this approach might lead to more beneficial outcomes for adults as well. Finally, future research could build on the initially planned third study for this dissertation project. It is still open to question whether it is possible to establish a personal connection of the participants to the Citizen Science project and whether this personal connection leads to an increased engagement. If this could be achieved, it would be a favourable opportunity to make outcomes for participants more beneficial without much effort. Moreover, further research is needed on how to integrate Citizen Science projects into the school day with less effort, so that the training of scientific reasoning can be implemented better and more efficiently into practice. The hypothesis that this could also be successful because of the students' personal connection to the Citizen Science project and therefore, a Citizen Science project participation on the school campus would be sufficient, still needs to be tested in future studies.

6. Conclusion

Scientific reasoning is an important topic for research today and in the future, as it is an essential skill for both students and adults. Hence, the development of a specific teaching approach is an important task for scientists. The results of this dissertation project show, that a Citizen Science approach is principally suitable to foster scientific reasoning skills without being less effective regarding factual knowledge gain, but it needs to be implemented under specific circumstances in order to be beneficial for both adults and students. A merely contributory data collection Citizen Science project is not sufficient to improve scientific reasoning of adults, even when being combined with a high amount of guidance. However, this dissertation project was able to show a functional method in the school context which provides new scientific knowledge, as well as implications for practice at school.

Hence, a curriculum-based Citizen Science project should be implemented more often at schools, as it can be helpful in addressing the lack of good scientific reasoning skills of students and thus can support students in becoming scientifically literate citizens of tomorrow's society.

References

Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology, 103*, 1–18. https://doi.org/10.1037/a0021017

Anthes, N., Götz, H., & Ströhle, G. (2019). *Ornithologischer Sammelbericht für die Region Neckar-Alb (Landkreise Tübingen, Reutlingen, Zollernalb).* Ausgabe 40 (Herbst 2019).

Anthes, N., Götz, H., & Ströhle, G. (2020). *Ornithologischer Sammelbericht für die Region Neckar-Alb (Landkreise Tübingen, Reutlingen, Zollernalb)*. Ausgabe 41 (Winter 2019/2020).

Bamberger, Y., & Tal, T. (2007). Learning in a personal context: Levels of choice in a free choice learning environment in science and natural history museums. *Science Education, 91*(1), 75-95. https://doi.org/10.1002/sce.20174

Bauer, H.H. (1992). *Scientific Literacy and the Myth of the Scientific Method*. Chicago: University of Illinois Press.

Baumert, J., Lehmann, R., Lehrke, M., Schmitz, B., Clausen, M., Hosenfeld, I., Köller, O., & Neubrand, J. (1997). *TIMSS – Mathematisch-naturwissenschaftler Unterricht im internationalen Vergleich.* Deskriptive Befunde. Opladen: Leske + Budrich.

Bhattacharjee, Y. (2005). Citizen scientists supplement work of Cornell researchers: a half-century of interaction with bird watchers has evolved into a robust and growing collaboration between volunteers and a leading ornithology lab. *Science, 308*(5727), 1402–1404. https://doi.org/10.1126/science.308.5727.1402

Bogner, F. X. (2007). Einstellungen und Werte im empirischen Konstrukt des jugendlichen Umweltbewusstseins. In Krüger, D., & Vogt, H. (Hrsg.), *Theorien in der biodidaktischen Forschung* (pp. 221–230). Heidelberg: Springer. https://doi.org/10.1007/978-3-540-68166-3_20

Bogner, F.X. & Wiseman, M. (2006). Adolescents‘ attitudes towards nature and environment: Quantifying the 2-MEV model. *Environmentalist, 26*,247-254. https://doi.org/10.1007/s10669-006-8660-9

Bonney, R. (1991). Amateurs welcome. *Living Bird*, *10*(1), 36-37.

Bonney, R. (2007). Citizen science at the Cornell Lab of Ornithology. In R. Yager, and J. Falk (Edds.) *Exemplary science in informal education settings: Standards-based success stories* (pp. 213-229). Arlington, Virginia: NSTA Press.

Bonney, R., Cooper, C. B., Dickinson, J., Kelling, S., Phillips, T. B., Rosenberg, K. V., & Shirk, J. (2009). Citizen science: a developing tool for expanding science knowledge and scientific literacy. *BioScience, 59*(11), 977–984. http://doi.org/10.1525/bio.2009.59.11.9

Bonney, R., Phillips, T. B., Ballard, H. L., & Enck, J. W. (2016). Can citizen science enhance public understanding of science? *Public Understanding of Science, 25*(1), 2–16. https://doi.org/10.1177/0963662515607406

Bonney, R., Shirk, J. L., Phillips, T. B., Wiggins, A., Ballard, H. L., Miller-Rushing, A. J., & Parrish, J. K. (2014). Next steps for citizen science. *Science, 343*(6178), 1436-1437. https://doi.org/10.1126/science.1251554

Brossard, D., Lewenstein, B., & Bonney, R. (2005). Scientific knowledge and attitude change: The impact of a citizen science project. *International Journal of Science Education, 27*(9), 1099–1121. https://doi.org/10.1080/09500690500069483

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated Cognition and the Culture of Learning. *Education Researcher, 18*, 32–42. http://doi.org/10.3102/0013189X018001032

Bryan, H. (1977). Leisure value systems and recreational specialization: The case of trout fishermen. *Journal of leisure research*, *9*(3), 174-187. https://doi.org/10.1080/00222216.1977.11970328

Chan, T. W., Roschelle, J., His, S., Kinshuk, Sharples, M., Brown, T., Patton, C., Cherniavsky, J., Pea, R., Norris, C., Soloway, E., Balacheff, N., Scardamalia, M., Dillenbourg, P., Looi, C., Milrad, M., & Hoppe, U. (2006). One-to-one technology-enhanced learning: An opportunity for global research collaboration. *Research and Practice in Technology Enhanced Learning, 1*(1), 3–29. http://doi.org/10.1142/S1793206806000032

Charney, J., Hmelo-Silver, C. E., Sofer, W., Neigeborn, L., Coletta, S., & Nemeroff, M. (2007). Cognitive apprenticeship in science through immersion in laboratory practices. *International Journal of Science Education, 29,* 195–213. https://doi.org/10.1080/09500690600560985

Chen, Y., Kao, T., & Sheu, J. (2003). A mobile learning system for scaffolding bird watching learning. *Journal of Computer Assisted Learning*, *19*, 347–359. http://doi.org/10.1046/j.0266-4909.2003.00036.x

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, K. C. (1997). *Internet links for science education.* New York: Plenum Press.

Crall, A. W., Jordan, R., Holfelder, K., Newman, G. J., Graham, J., & Waller, D. M. (2013). The impacts of an invasive species citizen science training program on participant attitudes, behavior, and science literacy. *Public Understanding of Science, 22*(6), 745-764. https://doi.org/10.1177/0963662511434894

Cronje, R., Rohlinger, S., Crall, A., & Newman, G. (2011). Does participation in citizen science improve scientific literacy? A study to compare assessment methods. *Applied Environmental Education & Communication*, *10*(3), 135-145. https://doi.org/10.1080/1533015X.2011.603611

Crowley, K., & Callanan, M. A. (1998). Describing and Supporting Collaborative Scientific Thinking in Parent-Child Interactions. *Journal of Museum Education, 23*, 12–17. https://doi.org/10.1080/10598650.1998.11510365

DDA (n.d.). [Citizen Science und Vögel]. Ornitho. https://www.ornitho.de/

Deci, E. L. (1975). *Intrinsic motivation.* New York: Plenum. http://doi.org/10.1007/978-1-4613-4446-9

Deci, E. L. (1992). The relation of interest to the motivation of behavior: A self-determination theory perspective. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.): *The role of interest in learning and development* (pp. 43–70). Hillsdale, NJ: Lawrence Erlbaum.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.

Deci, E. L., & Ryan, R. M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik, 39*, 223-238.

Deci, E. L., Schwartz, A. J., Sheinman, L., & Ryan, R. M. (1981). An instrument to assess adults' orientations toward control versus autonomy with children: Reflections on intrinsic motivation and perceived competence. *Journal of Educational Psychology*, *73*, 642–650. https://doi.org/10.1037/0022-0663.73.5.642

Deci, E. L., & Ryan, R. M. (2003). Intrinsic motivation inventory. *Self-determination theory*, *267*.

De Jong, T., & Van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research, 68*, 179–202. https://doi.org/10.3102/00346543068002179

DeWitt, J., & Storksdieck, M. (2008). A short review of school field trips: Key findings from the past and implications for the future. *Visitor studies, 11*(2), 181-197. https://doi.org/10.1080/10645570802355562

Dierschke, V. (2017). *Welcher Vogel ist das?* (3rd. ed.). Stuttgart: Franckh-Kosmos Verlags-GmbH & Co. KG.

Druschke, C. G., & Seltzer, C. E. (2012). Failures of engagement: Lessons learned from a citizen science pilot study. *Applied Environmental Education & Communication*, *11*(3-4), 178-188. https://doi.org/10.1080/1533015X.2012.777224

Duschl, R. A., & Grandy, R. E. (2008). *Teaching Scientific Inquiry: Recommendations for Research and Implementation*. Rotterdam, the Netherlands: Sense.

Eccles, J. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives: Psychological and sociological approaches* (pp. 75-146). San Francisco, CA: W. H. Freeman.

Eshach, H. (2007). Bridging in-school and out-of-school learning: formal, non-formal and informal education. *Journal of Science Education and Technology, 16*, 171–190. https://doi.org/10.1007/s10956-006-9027-1

Evans, C., Abrams, E., Reitsma, R., Roux, K., Salmonsen, L., & Marra, P. P. (2005). The neighborhood nestwatch program: participant outcomes of a citizen-science ecological research project. *Conservation Biology*, *19*(3), 589-594. https://doi.org/10.1111/j.1523-1739.2005.00s01.x

Falk, J. H. (1983). Field trips: A look at environmental effects on learning. *Journal of Biological Education, 17,* 137–141. https://doi.org/10.1080/00219266.1983.9654522

Falk, J. H., & Balling, J. D. (1982). The field trip milieu: Learning and behavior as a function of contextual events. *Journal of Educational Research, 76,* 22–28. https://doi.org/10.1080/00220671.1982.10885418

Falk, J. H., Storksdieck, M., & Dierking, L. D. (2007). Investigating public science interest and understanding: Evidence for the importance of free-choice learning. *Public Understanding of Science*, *16*(4), 455-469. https://doi.org/10.1177/0963662506064240

Fischer, F., Kollar, I., Ufer, S., Sodian, B., Hussmann, H., Pekrun, R., Neuhaus, B., Dorner, B., Pankofer, S., Fischer, M., Strijbos, J., Heene, M., & Eberle, J. (2014). Scientific reasoning and argumentation: advancing an interdisciplinary research agenda in education. *Frontline Learning Research*, *2*(3), 28-45. https://doi.org/10.14786/flr.v2i2.96

Ford, M. (2008). Disciplinary authority and accountability in scientific practice and learning. *Science Education*, *92*(3), 404-423. https://doi.org/10.1002/sce.20263

Ford, M. J. (2012). A dialogic account of sense-making in scientific argumentation and reasoning. *Cognition and Instruction*, *30*(3), 207-245. http://dx.doi.org/10.1080/07370008.2012.689383

Frątczak, M., Sparks, T. H., Randler, C., & Tryjanowski, P. (2020). Circadian preferences of birdwatchers in Poland: do “owls” prefer watching night birds, and “larks” prefer daytime ones?. *PeerJ*, *8*, e8673. https://doi.org/10.7717/peerj.8673

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings for the National Academy of Sciences of the United States of America, 111*, 8410–8415. https://doi.org/10.1073/pnas.1319030111

Fujitani, M., McFall, A., Randler, C., & Arlinghaus, R. (2017). Participatory adaptivemanagement leads to environmental learning outcomes extending beyond the sphere of science. *Science Advances*, *3*(6), e1602516. https://doi.org/10.1126/sciadv.1602516

Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching: A meta-analysis. *Review of educational research*, *82*(3), 300-329. https://doi.org/10.3102/0034654312457206

Gerl, T., Randler, C., & Neuhaus B.J. (2021; in review). Vertebrate Species knowledge – an important skill is threatened by extinction. *International Journal of Science Education*.

Gijlers, H., & De Jong, T. (2005). The relation between prior knowledge and students' collaborative discovery learning processes. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching, 42*(3), 264-282. https://doi.org/10.1002/tea.20056

Gilbert, J. (2005). Catching the Knowledge Wave?: The knowledge society and the future of public education in New Zealand. *Education Policy Directions in Aotearoa/New Zealand*, 53-70.

Gilbert, J., & Priest, M. (1997). Models and discourse: A primary school science class visit to a museum. *Science Education, 81,* 749–762. https://doi.org/10.1002/(SICI)1098-237X(199711)81:6<749::AID-SCE10>3.0.CO;2-I

Golumbic, Y. N., Fishbain, B., & Baram-Tsabari, A. (2020). Science literacy in action: understanding scientific data presented in a citizen science platform by non-expert adults. *International Journal of Science Education, Part B, 10*(3), 1-16. https://doi.org/10.1080/21548455.2020.1769877

Grolnick, W. S., & Ryan, R. M. (1987). Autonomy in children's learning: An experimental and individual difference investigation. *Journal of Personality and Social Psychology*, *52*, 890–898. https://doi.org/10.1037/0022-3514.52.5.890

Grolnick, W. S., Ryan, R. M., & Deci, E. L. (1991). Inner resources for school achievement: Motivational mediators of children's perceptions of their parents. *Journal of Educational Psychology, 83,* 508-517. https://doi.org/10.1037/0022-0663.83.4.508

Hartmann, S., Upmeier zu Belzen, A., Krüger, D., & Pant, H. A. (2015). Scientific reasoning in Higher Education: Constructing and Evaluating the Criterion-Related Validity of an Assessment of Preservice Science Teachers' Competencies. *Zeitschrift für Psychologie, 223*(1), 47-53. https://doi.org/10.1027/2151-2604/a000199

Henry, C. K. (1992). Retention and recall of images: Evaluating museum experiences of school children. *Visual Arts Research, 18*, 82-92.

Hill, C. T. (2007). The post-scientific society. *Issues in Science and Technology*, *24*(1), 78-84.

Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science*, *326*(5958), 1410-1412. https://doi.org/ 10.1126/science.1177067

Hooykaas, M. J., Schilthuizen, M., Aten, C., Hemelaar, E. M., Albers, C. J., & Smeets, I. (2019). Identification skills in biodiversity professionals and laypeople: A gap in species literacy. *Biological Conservation*, *238*, 108202. https://doi.org/10.1016/j.biocon.2019.108202

IEA = International Association for the Evaluation of Educational Achievement (1994). *IEA's Third International Mathematics and Science Study (TIMSS), Population 2 Item Pool, Science Items.* Den Haag: IEA.

IEA = International Association for the Evaluation of Educational Achievement (2013a). *TIMSS 2011 Assessment, Released Science Items, Grade 8.* Chestnut Hill, Massachusetts/ Amsterdam: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College/ IEA Secretariat.

IEA = International Association for the Evaluation of Educational Achievement (2013b). *TIMSS 2011 Assessment, Released Science Items, Grade 4*. Chestnut Hill, Massachusetts/ Amsterdam: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College/ IEA Secretariat.

Itzek-Greulich, H., Flunger, B., Vollmer, C., Nagengast, B., Rehm, M., & Trautwein, U. (2015). Effects of a science center outreach lab on school students' achievement–Are student lab visits needed when they teach what students can learn at school?. *Learning and Instruction, 38*, 43-52. https://doi.org/10.1016/j.learninstruc.2015.03.003

Jordan, R. C., Gray, S. A., Howe, D. V., Brooks, W. R., & Ehrenfeld, J. G. (2011). Knowledge gain and behavioral change in citizen-science programs. *Conservation biology, 25*(6), 1148-1154. https://doi.org/10.1111/j.1523-1739-2011-01745.x

Kelling, S., Johnston, A., Hochachka, W. M., Iliff, M., Fink, D., Gerbracht, J., Lagoze, C., La Sorte, F. A., Moore, T., Wiggins, A., Wong, W. K., Wood, C., & Yu, J. (2015). Can observation skills of citizen scientists be estimated using species accumulation curves? *PloS one*, *10*(10), e0139600. https://doi.org/10.1371/journal.pone.0139600

Keselman, A. (2003). Supporting inquiry learning by promoting normative understanding of multivariable causality. *Journal of Research in Science Teaching*, *40*(9), 898-921. https://doi.org/10.1002/tea.10115

Krüger, D., Hartmann, S., Nordmeier, V. & Upmeier zu Belzen, A. (2020). Measuring Scientific Reasoning Competencies - Multiple Aspects of Validity. In: Zlatkin-Troitschanskaia, O., Pant, H., Toepper, M. & Lautenbach, C. (Eds.): *Student Learning in German Higher Education (pp. 261-280)* Wiesbaden: Springer.

Kubota, C., & Olstad, R. (1991). Effects of novelty-reducing preparation on exploratory behavior and cognitive learning in a science museum. *Journal of Research in Science Teaching, 28*, 225–234. https://doi.org/10.1002/tea.3660280304

Kuentzel, W. F., & Heberlein, T. A. (2006). From novice to expert? A panel study of specialization progression and change. *Journal of Leisure Research*, *38*(4), 496-512. https://doi.org/10.1080/00222216.2006.11950089

Kuhn, D., & Franklin, S. (2006). The second decade: What develops (and how). In W. Damon, R. M. Lerner, D. Kuhn, & R. Siegler (Eds.), *Handbook of child psychology: Vol. 2. Cognition, perception, and language* (6th ed., pp. 953–994). Hoboken, NJ: Wiley.

Lazonder, A. W., & Harmsen, R. (2016). Meta-analysis of inquiry-based learning: Effects of guidance. *Review of Educational Research, 86*, 681–718. http://doi.org/10.3102/0034654315627366

Lee, J.H., & Scott, D. (2004). Measuring birding specialization: A confirmatory factor analysis. *Leisure Sciences*, *26*(3), 245-260. https://doi.org/10.1080/01490400490461387

Lee, J. H., & Scott, D. (2006). For better or worse? A structural model of the benefits and costs associated with recreational specialization. *Leisure Sciences*, *28*(1), 17-38. https://doi.org/10.1080/01490400590962461

Lopez, B., Minor, E., & Crooks, A. (2020). Insights into human-wildlife interactions in cities from bird sightings recorded online. *Landscape and Urban Planning, 196*, 103742. https://doi.org/10.1016/j.landurbplan.2019.103742

Lucas, K. B. (2000). One teacher's agenda for a class visit to an interactive science center. *Science Education, 84*, 524–544. https://doi.org/10.1002/1098-237X(200007)84:4<524::AID-SCE6>3.0.CO;2-X

Mäeots, M., Pedaste, M., & Sarapuu, T. (2008). Transforming students' inquiry skills with computer-based simulations. In *2008 Eighth IEEE International Conference on Advanced Learning Technologies* (pp. 938-942). Santander, Cabria: IEEE. https://doi.org/10.1109/ICALT.2008.239

Mayer, J. (2007). Erkenntnisgewinnung als wissenschaftliches Problemlösen. In D. Krüger, & H. Vogt (Eds.), *Handbuch der Theorien in der biologiedidaktischen Forschung* (pp. 177–186). Berlin: Springer. https://doi.org/10.1007/978-3-540-68166-3_16

McFarlane, B. L. (1994). Specialization and motivations of birdwatchers. *Wildlife Society Bulletin*, *22*(3), 361-370.

National Academies of Sciences, Engineering, & Medicine (NASEM). (2018). *Learning through citizen science: Enhancing opportunities by design.* Washington, DC: The National Academies Press. https://doi.org/10.17226/25183

National Research Council (1996). *National Science Education Standards*. Washington, DC: The National Academies Press. https://doi.org/10.17226/4962.

National Research Council. (2000). *Inquiry and the national science education standards*. Washington, DC: The National Academies Press. https://doi.org/10.17226/9596

National Research Council. (2008). *Research on future skill demands*. Washington, DC: National Academies Press. https://doi.org/10.17226/12066

National Research Council (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas.* Washington, DC: The National Academies Press. http://doi.org/10.17226/13165

Njoo, M., & De Jong, T. (1993). Exploratory learning with a computer simulation for control theory: Learning processes and instructional support. *Journal of research in science teaching, 30*(8), 821-844. https://doi.org/10.1002/tea.3660300803

Ortiz, J. L., Conkey, A. A., Brennan, L. A., Fedynich, L. V., & Green, M. (2020). Wild Bird Workshop: A Professional Development Opportunity for Educators. *The American Biology Teacher*, *82*(1), 3-10. https://doi.org/10.1525/abt.2020.82.1.3

Osborne, J. (2013). The 21st century challenge for science education: Assessing scientific reasoning. *Thinking skills and creativity*, *10*, 265-279. https://doi.org/10.1016/j.tsc.2013.07.006

Pant, H. A., Stanat, P., Schroeders, U., Roppelt, A., Siegle, T., & Pöhlmann, C. (Eds.). (2013). *IQB-Ländervergleich 2012: Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I*. Münster: Waxmann.

Parrish, J. K., Jones, T., Burgess, H. K., He, Y., Fortson, L., & Cavalier, D. (2019). Hoping for optimality or designing for inclusion: Persistence, learning, and the social network of citizen science. *Proceedings of the National Academy of Sciences*, *116*(6), 1894-1901. https://doi.org/10.1073/pnas.1807186115

Pedaste, M., Mäeots, M., Leijen, Ä., & Sarapuu, T. (2012). Improving students' inquiry skills through reflection and self-regulation scaffolds. *Technology, Instruction, Cognition and Learning*, *9*(1-2), 81-95.

Pedaste, M., Mäeots, M., Siiman, L. A., de Jong, T., van Riesen, S. A. N., Kamp, E. T., Manoli, C. C., Zacharia, Z. C., & Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational Research Review, 14*, 47–61. http://doi.org/10.1016/j.edurev.2015.02.003

Pfeiffer, V. D. I., Gemballa, S., Jarodzka, H., Scheiter, K., & Gerjets, P. (2009). Situated learning in the mobile age: Mobile devices on a field trip to the sea. *Research in Learning Technology, 17*, 187-199. https://doi.org/10.1080/09687760903247666

Pfeiffer, V. D. I., Scheiter, K., & Gemballa, S. (2012). Comparing and combining traditional teaching approaches and the use of video clips for learning how to identify species in an aquarium. *Journal of Biological Education, 46*(3), 140–148. https://doi.org/10.1080/00219266.2011.584892

Randler, C. (2008a). Teaching species identification–A prerequisite for learning biodiversity and understanding ecology. *Eurasia Journal of Mathematics, Science and Technology Education*, *4*(3), 223-231. https://doi.org/10.12973/ejmste/75344

Randler, C. (2008b). Pupils' factual knowledge about vertebrate species. *Journal of Baltic Science Education*, *7*(1), 48-54.

Randler, C. (2020). *Line transect counts Weggental.* Unpublished dataset.

Randler, C., & Bogner, F. X. (2002). Comparing methods of instruction using bird

Species identification skills as indicators. *Journal of Biological Education, 36*(4), 181–188. https://doi.org/10.1080/00219266.2002.9655830

Randler, C., & Bogner, F. X. (2006). Cognitive achievements in identification skills. *Journal of Biological Education*, *40*(4), 161-165. https://doi.org/10.1080/00219266.2006.9656038

Randler, C., & Bogner, F. X. (2008). Efficacy of Two Different Instructional Methods Involving Complex Ecological Content. *International Journal of Science and Mathematics Education, 7,* 315-337. https://doi.org/10.1007/s10763-007-9117-4

Randler, C., A. Ilg, & J. Kern (2005). Cognitive and emotional evaluation of an amphibian conservation program for elementary school students. *Journal of Environmental Education* 37: 43-52. https://doi.org/10.3200/JOEE.37.1.43-52

Randler, C., B. Kummer, & C. Wilhelm (2012). Adolescent Learning in the Zoo: Embedding a Non-Formal Learning Environment to Teach Formal Aspects of Vertebrate Biology. *Journal of Science Education and Technology, 21,* 384-391. https://doi.org/10.1007/s10956-011-9331-2

Randler, C., Tryjanowski, P., Jokimäki, J., Kaisanlahti-Jokimäki, M. L., & Staller, N. (2020). SARS-CoV2 (COVID-19) Pandemic Lockdown Influences Nature-Based Recreational Activity: The Case of Birders. *International journal of environmental research and public health*, *17*(19), 7310. https://doi.org/10.3390/ijerph17197310

Resnick, L. (1987). Learning in school and out. *Educational Researcher, 16*, 13–54. https://doi.org/10.2307/1175725

Riesch, H., & Potter, C. (2014). Citizen science as seen by scientists: Methodological, epistemological and ethical dimensions. *Public Understanding of Science*, *23*, 107-120. https://doi.org/10.1177/0963662513497324

Ryan, R. M., & Deci, E. L. (2002). An overview of self-determination theory: An organismic-dialectical perspective. In E. L. Deci, & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 3-33). Rochester: University of Rochester.

Savery, J. R., & Duffy, T. M. (1995). Problem based learning: An instructional model and its constructivist framework. *Educational Technology, 35*, 31–38.

Schneiderhan-Opel, J., & Bogner, F. X. (2020). The relation between knowledge acquisition and environmental values within the scope of a biodiversity learning module. *Sustainability*, *12*(5), 2036. https://doi.org/10.3390/su12052036

Scott, D., & Lee, J. H. (2010). Progression, stability, or decline? Sociological mechanisms underlying change in specialization among birdwatchers. *Leisure Sciences*, *32*(2), 180-194. https://doi.org/10.1080/01490400903547203

Seybold, B., Braunbeck, T., & Randler, C. (2014). Primate conservation – An evaluation of two different educational programs in Germany. *International Journal of Science and Mathematics Education, 12*(2), 285–305. https://doi.org/10.1007/s10763-013-9405-0

Shirk, J. L., Ballard, H. L., Wilderman, C. C., Phillips, T., Wiggins, A., Jordan, R., McCallie, E., Minarchek, M., Lewenstein, B. V., Krasny, E., & Bonney, R. (2012). Public participation in scientific research: a framework for deliberate design. *Ecology and society*, *17*(2), 29. http://dx.doi.org/10.5751/ES-04705-170229

Stylinski, C. D., Peterman, K., Phillips, T., Linhart, J., & Becker-Klein, R. (2020). Assessing science inquiry skills of citizen science volunteers: a snapshot of the field. *International Journal of Science Education, Part B*, *10*(1), 77-92. https://doi.org/10.1080/21548455.2020.1719288

The Cornell Lab of Ornithology (n.d.). [Citizen Science and birds]. eBird. https://www.ebird.org/

Thomas, R. L., & Fellowes, M. D. (2017). Effectiveness of mobile apps in teaching field-based identification skills. *Journal of Biological Education*, *51*(2), 136-143. https://doi.org/10.1080/00219266.2016.1177573

Thompson, S. (2007). *BirdSleuth: Final Evaluation Report*. Ithaca, NY: Seavoss Associates.

Trumbull, D. J., Bonney, R., Bascom, D., & Cabral, A. (2000). Thinking scientifically during participation in a citizen-science project. *Science education*, *84*(2), 265-275.https://doi.org/10.1002/(SICI)1098-237X(200003)84:2<265::AID-SCE7>3.0.CO;2-5

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

Whitehead, A. N. (1967). *Aims of education*. Simon and Schuster.

Wilde, M., Bätz, K., Kovaleva, A., & Urhahne, D. (2009). Überprüfung einer Kurzskala intrinsischer Motivation (KIM). *Zeitschrift für Didaktik der Naturwissenschaften, 15*, 31–45.

Wilmsen, C., & Krishnaswamy, A. (2008). Challenges to institutionalizing participatory research in community forestry in the US. In C. Wilmsen, W. Elmendorf, L. Fische, J. Ross, B. Sarathy, and G. Wells (Eds.). *Partnerships for empowerment: participatory research for community-based natural resource management* (pp. 47-67). London, UK: Earthscan.

Wood, C., Sullivan, B., Iliff, M., Fink, D., & Kelling, S. (2011). eBird: engaging birders in science and conservation. *PLoS Biol*, *9*(12), e1001220.

Wünschmann, S., P. Wüst-Ackermann, C. Randler, C. Vollmer, & H. Itzek-Greulich (2017). Learning achievement and motivation in an out-of-school setting – visiting amphibians and reptiles in a zoo is more effective than a lesson at school. *Research in Science Education* 47: 497–518. https://doi.org/10.1007/s11165-016-9513-2

Wüst-Ackermann, P., Vollmer, C., Randler, C., & Itzek-Greulich, H. (2018). The Vivarium: Maximizing learning with living invertebrates—An out-of-school intervention is more effective than an equivalent lesson at school. *Insects*, *9*(1), 3. https://doi.org/10.3390/insects9010003

Appendix

Appendix A

AIC fits for scientific reasoning and factual knowledge.

	Variables included	logLik	AIC	ΔAIC
Scientific Reasoning	All variables: group, time, gender, age, class, group*time, group*gender, time*gender, group*time*gender	3976.85	3982.85	0
	All variables without group*time*gender	3980.04	3986.04	3.19
	All variables without time*gender	3976.85	3982.85	0
	All variables without group*gender	3976.85	3982.85	0
	All variables without group*time	3976.85	3982.85	0
	All variables without class	3979.37	3983.37	0.52
	All variables without age	4105.92	4111.92	129.07
	All variables without gender and without all interactions including gender	3985.99	3991.99	9.14
	All variables without time and without all interactions including time	4012.21	4018.21	35.36
	All variables without group and without all interactions including group	3988.27	3994.27	11.42
Factual knowledge	All variables: group, time, gender, age, class, group*time, group*gender, time*gender, group*time*gender	3839.93	3845.93	0
	All variables without group*time*gender	3848.62	3854.62	8.69
	All variables without time*gender	3839.93	3845.93	0
	All variables without group*gender	3839.93	3845.93	0
	All variables without group*time	3839.93	3845.93	0

All variables without class	3847.22	3851.22	5.29
All variables without age	3966.14	3972.14	126.21
All variables without gender and without all interactions including gender	3853.43	3859.43	13.5
All variables without time and without all interactions including time	4515.94	4521.94	676,01
All variables without group and without all interactions including group	3857.17	3863.17	17.24

Notes. logLik = log-likelihood, ΔAIC = difference between the model given and the model with the best fit.

Appendix B

AIC fits for the KIM scales

	Variables included	logLik	AIC	ΔAIC
Enjoyment	Group, age, class	637.72	641.72	0
	including gender	640.52	644.52	2.80
	including gender and gender*group	639.86	643.86	2.14
Competence	Group, age, class	623.60	627.60	0
	including gender	626.27	630.27	2.67
	including gender and gender*group	623.76	627.76	0.16
Choice	Group, age, gender, gender*group	964.54	966.54	0
	without gender*group	967.80	969.80	3.26
	without gender and gender*group	967.52	969.52	2.98
Pressure	Group, age, class, gender, gender*group	697.01	701.01	0
	without gender*group	700.82	704.82	3.81
	without gender and gender*group	700.89	704.89	3.88

Notes. logLik = log-likelihood, ΔAIC = difference between the model given and the model with the best fit.

Appendix C

AIC fits for the MEV scales preservation and utilization.

	Variables included	logLik	AIC	ΔAIC
Preservation	All variables: group, time, gender, age, group*time, group*gender, time*gender, group*time*gender	1349.38	1353.38	10.61
	All variables without group*time*gender	1348.18	1352.18	8.96
	All variables without time*gender	1349.38	1353.38	10.61
	All variables without group*gender	1349.38	1353.38	10.61
	All variables without group*time	1349.38	1353.38	10.61
	All variables without age	1418.49	1422.49	79.27
	All variables without gender and without all interactions including gender	1372.25	1376.25	33.03
	All variables without time and without all interactions including time	1348.81	1352.81	9.59
	All variables without group and without all interactions including group	1339-22	1343.22	0
Utilization	All variables: group, time, gender, age, class, group*time, group*gender, time*gender, group*time*gender	1417.31	1423.31	11.49
	All variables without group*time*gender	1412.77	1418.77	6.95
	All variables without time*gender	1417.31	1423.31	11.49
	All variables without group*gender	1417.31	1423.31	11.49
	All variables without group*time	1417.31	1423.31	11.49
	All variables without class	1423.70	1427.70	15.88
	All variables without age	1470.95	1476.95	65.13
	All variables without gender and without all	1423.65	1429.65	17.83

interactions including gender			
All variables without time and without all interactions including time	1422.15	1428.15	16.33
All variables without group and without all interactions including group	1405.82	1411.82	0

Notes. Class was not further included in calculations regarding preservation, as the model did not converge and thus, class did not have any effect in the model. logLik = log-likelihood, ΔAIC = difference between the model given and the model with the best fit.