Managing Knowledge Diversity in Computer-Supported Inquiry-Based Science Education

Von der Fakultät für Ingenieurwissenschaften,
Abteilung Informatik und Angewandte Kognitionswissenschaft
der Universität Duisburg-Essen

zur Erlangung des akademischen Grades
Doktor der Ingenieurwissenschaften (Dr.-Ing.)

genehmigte Dissertation von

Sven Manske
aus
Moers, Deutschland

1. Gutachter: Prof. Dr. Heinz Ulrich Hoppe
2. Gutachter: Prof. Dr. Mohamed Amine Chatti

Tag der mündlichen Prüfung: 02.10.2019

DOI: 10.17185/duepublico/71585
URN: urn:nbn:de:hbz:464-20200514-110152-9

Contents

List of Tables

List of Figures

1 Introduction

1.1 Motivation

> "It is time for parents to teach young people early on that in diversity there is beauty and there is strength." (Angelou 1994)

This powerful quote from Maya Angelou highlights the importance of diversity and postulates to educate young people. Although she argues from a civil and women rights perspective, the discussion about diversity has a long tradition in a lot of disciplines and many parallels across them. In the tradition of group learning, researchers discussed about how to perform groupings of learners. One of the most famous approaches, the jigsaw method by Aronson et al. (1978), had the goal to adjust the mixing of ethnicity in classrooms with a high multi-cultural background. Therefore, it employs individual strengths and empowers young students by reducing the resistance to work with each other.

The research in computer-supported collaborative learning (CSCL) picked up the discussion about forming learning groups with respect to a specific goal. Therefore, the jigsaw method has been adopted for (digital) learning environments (Hinze et al. 2002). A lot of attention has been spent on the question whether homogeneous or heterogeneous groupings perform better according to a specific target, for example the learning outcome or the knowledge gain. Although this idea is scientifically correct, it has two flaws: (1) the learning outcome of a group is usually measured as a sum or average of individual scores; (2) heterogeneity is often measured as a performance characteristics, which determines heterogeneity in terms of skills. It can be seen as common sense that heterogeneous groups perform better in terms of knowledge gain or learning outcome. However, if this is reduced to skill, this leads to a stigmatization of weak learners, but also to a lack of fostering of strong learner. Statistically, in a group of weak and strong learners, the maximum (relative) learning gain will be probably observed on the side of the weak learners.

This motivates the question, if there are other ways to form groups without stigmatizing weak learners, and with having a heuristic or goal setting that is beneficial for everyone. Managing diversity has a lot of potential in many disciplines and fields. Organizational or institutional diversity can support companies in many ways. For example, merging different perspectives and utilizing individual experiences from different (cultural) backgrounds can lead to more integrity and better acceptance on all stakeholder parties. The potential of diversity has been highlighted a lot during the last years.

Diversity seems to be a powerful and overloaded term (Fardon 2003), of which its connotation differs from field to field. Exploring diversity in research has been a topic particularly in the humanities or organizational sciences. In educational research, diversity was important in terms of ethnical diversity, historically from the approaches by Aronson. During the last decades, educational research has focused a lot in exploring the 21st century skills. The modern society and the digital age posed new challenges for employers and particularly for employees. In the educational field there has been a shift from traditional teaching to modern, student-centered approaches, which involve critical thinking, creativity, problem solving competencies - metacognitive skills that complement pure memory skills. The primary goal of the educational system is not anymore just mediating and transferring domain knowledge. The goal is to foster competencies that enable learners to extend their knowledge by their own, to ask questions or to scrutinize common things in order to uncover the potential for continuous improvement or growth. It is rather to provide means and mind tools for learners in order to extend their knowledge and construct their own reality.

Some organizations like the National Research Council of the USA identify these 21st century skills as key competencies for the future (Council 1996). Gago et al. (2004) even predict in their report *Europe Needs More Scientists* a major crisis in the production of human resources for science and technology . On the one hand, with respect to OECD data, the amount of technical and scientific jobs in Europe has a tendency to increase (OECD et al. 2003). On the other hand, the number of students with a degree in science, technology and engineering decreases. This can be overcome by implementing science programmes and fostering science education. Particularly schools, which enables young students to experience and study science in a systematic way, by asking questions, carrying out experiments, and collecting results in a scientific manner.

> "They may remember pleasure, joy, success, excitement – or a feeling of failure, boredom, of not understanding counter-intuitive concepts and abstract ideas with no relevance to their daily lives and a constant struggle to find strategies to arrive at exercise solutions without deep thinking or real understanding." (Gago et al. 2004)

To cause learners to stay in sciences needs to connect their positive learning experiences and good outcome with positive affective components. The teachers work to create the learning designing and to orchestrate the learning scenario s/he creates the conditions within the classroom to influence this outcome. Apart from the teacher, we, as experts and researchers in the educational domain, expect from the students that they are the 21st century learners, being able to manage and plan their own learning, to divide task planning, create learning agendas and being aware of social aspects. We expect a lot on the cognitive and metacognitive level. The higher the learning outcome the higher the responsibilities for each party.

With the rise of internet and communication technology (ICT), the demand for using ICT in educational contexts grew. The research area of Technology Enhanced Learning (TEL) is not a synonym for e-learning. The field brings together the disciplines of learning science, pedagogy and computer science in order to create mechanisms and (digital) tools to support learning. The triangular relation between the three disciplines involved highlights that it refers to learning and teaching 'with' technology rather than learning 'through' technology. It exploits individual and social factors from the learning science and defines software tools that are useful for teaching and learning, for cognition and metacognition, digital mind tools and diagnostic functions, and much more.

Scientific learning has shown a lot of potential in the research of TEL. Computer-supported Inquiry Learning (CoSIL) environments follow the idea of inquiry-based learning (IBL) with digital tools, which support scientific inquiry on the part of the learners and teachers. The work of this thesis is contextualized within the Go-Lab project, which aims in promoting inquiry-based science education on a large scale. Go-Lab can be seen as a pedagogical middleware, which provides tools to structure and orchestrate inquiry-based learning. The teacher creates a so called inquiry learning space (ILS), which is a customized learning environment that can be seen as a structured collection of *apps* (embeddable applications), online science laboratories and learning resources. Each ILS is structured as subsequent inquiry phases, which follow best practices of IBL. A variety of online tools and cognitive scaffolds can be added to the ILS. Such tools can help learners, for example, to structure their learning, to ask questions, to externalize knowledge, or to support the scientific processes. The experimentation, as the most distinctive phase in scientific inquiry, is supported by a federation of online laboratories. A high number of virtual and remote laboratories can be embedded into the Go-Lab ecosystem.

The construction of knowledge is often seen as a key ingredient to learning. In addition to the aforementioned aspects of using explicit knowledge representations to support inquiry, possibilities to make use of the assessment of knowledge in a way that it is integrated into the learning design have been investigated as well. One thread of existing research on small group learning has addressed positive effects of grouping learners with complementary knowledge, another one has focused on representing and visualizing knowledge distributions to facilitate cognitive group awareness. Cognitive group awareness can be seen as the perception or awareness of learners about their co-learners' knowledge states, assumptions, opinions or interests. Such information, presented to the learners in form of cognitive group awareness tools, implicitly guides learners through discovery processes and support communication, coordination and reflection of collaborative processes (Janssen and Bodemer 2013; Bodemer and Dehler 2011). Although a combination of knowledge-based grouping, representation of knowledge, and mirroring of knowledge states seems obvious, this has been investigated rarely. As part of this work, visualizations of individual and group knowledge have been created and provided. Such tools have the potential to support metacognition and organization but also help to initialize group processes and learning.

A shift in teaching practices could be observed, but also in the use of intelligent technologies and orchestration of e-learning. With the acceptance of constructivism as the predominant paradigm in teaching, the traditional teaching methods are more and more replaced by student-centered approaches, where students do active exploration and discovery. This shift to more active forms of learning could be observed in the field of e-learning as well. Former applications usually had the function to distribute static materials to the learners, in their role as consumers. Modern approaches demand the learners to externalize their knowledge and to challenge them. Their role changes to prosumers, where they create their own learning objects in more coherent and demanding learning scenarios, with active exploration and engagement.

The aim of this thesis is to explore diversity as a given and observable aspect in classrooms and to develop approaches to make use of knowledge diversity in computer-supported inquiry-based learning scenarios. The practical use of diversity poses new challenges in the frame of this research in the field of inquiry-based science education. This comprises the operationalization of knowledge and knowledge diversity, which is a precondition for algorithmic approaches or the empirical evaluation of these concepts. Externalizations of knowledge and the processed knowledge model can be further integrated into learning scenarios. On the one hand, this work explores methods to support the orchestration of group learning by automatically forming groups based on knowledge models. This work presents a method based on knowledge diversity, where learners are grouped according to satisfy the condition of having complementary knowledge. However, the field of artificial intelligence (AI) is facing the challenge, that many highly optimized algorithms produce output with a good performance in the application domain, but the underlying models and results are not explainable. Therefore, many researchers adopted the topic to develop explainable AIs (XAI). Corresponding to this work, the semantic group formation has the goal to create a human-interpretable model of the group knowledge as a basis for the group formation with the result of a clear model that explains the formation. Cognitive group awareness tools Therefore, as part of this thesis, a group awareness tool has been developed, which takes the output of the group formation and utilizes the learner model as an explicit tool that guides learners through the collaboration phase in inquiry-based learning scenario.

1.2 Problem Statement

Tobin (1990) cited Novak (1988) in order to frame the problems of science learning and knowledge building explicitly:

> "The science laboratory has always been regarded as the place where students should learn the process of doing science. But summaries of research on the value of laboratory for learning science did not favor laboratory over lecture-demonstration [...] and more recent studies also show an appalling lack of effectiveness of laboratory instruction [...] our studies showed that most students in laboratories gained little insight either regarding the key science concepts involved or toward the process of knowledge construction."

The results from Novak highlight the challenges for inquiry-based science education. Particularly the science lab, as crucial and distinctive its role may be, contains a lot of distractors and confusions, which complicates learning and knowledge building on the part of the students.

Web-based inquiry learning environments such as Go-Lab structure the scientific process and support learners by providing cognitive scaffolds and mind tools to externalize and to structure their knowledge. In such environments and associated learning scenarios, learners are more active and produce learner-generated content inside the environment, such as texts, concept maps, hypotheses or observations. Such active productions reflect - to some extent - knowledge on the part of the learners. This thesis explored how these productions in order to infer and represent knowledge can be facilitated. By employing methods of semantic extraction and natural language processing, individual and group knowledge models are created to explore approaches which facilitate these knowledge models. Whether scaffolds can be provided for learners in order to support knowledge creation has been investigated in this work.

Go-Lab promotes a special kind of blended learning, where classroom scenarios are accompanied by a virtual learning environment. The Go-Lab portal provides support mechanisms, teaching materials and the digital tools to generate content in the context of inquiry-based learning. In contrast to pure online courses, such augmented classroom scenarios are more likely to be influenced by social factors. The research on diversity has shown in a lot of fields, how different social, ethnical or other characteristics can be facilitated to impact and to improve learning. Methods like the jigsaw approach from the field of collaborative learning have demonstrated their usefulness for a diverse classroom. However, in the debate on diversity in a classroom context, knowledge is underrepresented as a factor for learning and is often seen as a black box. As part of this thesis, knowledge diversity is defined and operationalized in inquiry-based learning scenarios with online learning environments. Observing that a diversity is given in classrooms, mechanisms to make use of this and to improve learning are explored. This assembly of knowledge diversity as a group characteristic requires to define a knowledge model on the individual or the group level and functions as an operationalized expression of knowledge diversity. Methods such as Jigsaw require the pedagogical model to be adapted in order to prepare the group work. This demands the teacher to resemble and distribute the learning materials according to a specific role schema, for instance when it comes to the formation of experts in Jigsaw. Although this mechanism relies on the idea of grouping learners regarding complementarity, this is only achieved by the specific pedagogical design. A grouping method could be more flexible and easy to orchestrate, when the teachers do not have to care about a distribution of learning materials and roles. In addition, it would be more plausible to facilitate the knowledge diversity that is already given and use this as a criterion for group formation.

Many learning environments such as Go-Lab, show heterogeneity in a lot of ways: the apps used in the scenarios, the structure or the format of the learner-generated content. In order to define individual or group knowledge models, a normalization of the data needs to take place. To quantify diversity of knowledge, representations of knowledge and an operationalization of the (knowledge) diversity need to be discussed and refined. Based on this operationalization, support mechanisms can be defined in an individual or in a group context. These mechanisms can be scaffolds directed to an individual or affecting a whole cohort, for example, automatic group formation. The characteristics of a grouping can affect the learning outcome, for example the debate homogeneous versus heterogeneous groupings. However, this work wants to shift the attention from the heterogeneously skilled groups to a different kind of grouping. Having high- and low-achievers leads to a stigmatizing of learners but also to an unbalanced situation. Besides fostering low-achievers, this thesis advocates for maximizing the learning gain for high-achievers as well, without designating one of them as less valuable. Therefore, knowledge diversity is employed, as a given characteristic of cohorts, to build a semantic group formation algorithm. In this group formation, the knowledge diversity is maximized for all groups in order to stimulate a maximized knowledge gain for everyone. This follows the idea to group learners with respect to knowledge complementarity. Assuming that the learners have a complementary knowledge (and assuming that there is some common ground), each learner may have something distinctive to contribute to the learning.

However, creating algorithms that know about the complementarity of knowledge does not necessarily lead to benefits for learning. Therefore, mechanisms to visualize and mirror these information states about knowledge back to the learners have been investigated and integrated into the learning scenario. The concept of cognitive group awareness is employed in order to create tools that make the knowledge complementarity of learners explicit. This shifts the idea of automatic group formation, that is usually handled as a black box for learners and teachers, to a glass box model. Such explainable models are highly of interest, as they are interpretable and build up trust on the part of the users involved.

The work of this thesis has been evaluated in classroom settings using the Go-Lab learning environment. For every evaluation, Go-Lab inquiry learning spaces have been created. To evaluate the group formation approaches, for each study two ILS have been constructed, one for an individual phase, one for a collaboration phase. To facilitate a diversity of topics and to foster discussion and argumentation, the scenario is situated in the field of renewable energies. In another study, the mechanism of visualizing and feeding back a shared knowledge model to individual learners has been explored. In this case, an ILS about an introduction to cryptography has been created.

1.3 Structure of this Thesis

The first chapter motivates computer-supported inquiry-based learning with online laboratories and highlights the importance of knowledge diversity in this field. Chapter 2 describes the state of the art in the research on inquiry-based learning and learning analytics, with a focus on semantic analyses in order to create individual and group knowledge models.

The approach of this thesis consists of two parts: technical architecture in chapter 3 and knowledge management approaches in chapter 4. The first part is dedicated to elaborate architectures for learning analytics, which have been used during the Go-Lab project. The Go-Lab learning analytics infrastructure was part of two publications, the first described the main components (Hecking et al. 2014), and the second one lined out a rapid prototyping approach for learning analytics apps based on a visual pipes-and-filters framework in Go-Lab (Manske et al. 2014). Finally, this architecture has been used to collect data during the first two years of Go-Lab and analyzed the learning processes and how teachers used Go-Lab for authoring their inquiry learning spaces (Manske et al. 2015a). In the second part (chapter 4), the concept of shared group knowledge models and knowledge diversity is introduced. This model is constructed through an automated semantic extraction. The idea of knowledge diversity is In this chapter, several approaches to facilitate knowledge and knowledge diversity in the context of Go-Lab are presented. This section is based on the publications of the concept cloud (Manske and Hoppe 2016) and group formation (Manske et al. 2015c; Manske and Hoppe 2016).

The subsequent chapters summarize three experimental studies in order to exploit the use of knowledge diversity and heterogeneity in classroom settings, contextualized in the Go-Lab project and its (technical) ecosystem. In the first study, a group formation based on performance characteristics and skills has been created, where the learning outcome of homogeneous groupings has been compared with heterogeneous groupings (chapter 5). In this study, the Go-Lab infrastructure has been used to create performance measures based on the actual learning artifacts that students created individually. The groups have been assigned automatically based on these measures. This chapter is based on the publication in the proceedings of the Conference on Computer-Supported Collaborative Learning (Manske et al. 2015c).

In the second study (chapter 6), the use of a shared knowledge model has been evaluated om a classroom study. The interactive visualization of the shared knowledge model functions as an innovative scaffold, which represents cognitive information of the class in a condensed form. The displayed knowledge elements have been extracted automatically with different semantic technologies. Two different dynamic and automatic extraction methods have been evaluated in this study and compared to a static case and a control group without the scaffold. This chapter is based on the publication in the proceedings of the International Conference on Advanced Learning Technologies (Manske and Hoppe 2016)

In the final study (chapter 7), the knowledge models that have been elaborated in this thesis have been used as an input for the (automatic) semantic group formation. In addition to this, a group awareness visualization based on the output of the group formation is used to support the collaborative processes. This tool-chain that supports facilitates cognitive group awareness and supports the orchestration of group learning through explainable algorithms and models. This was done by showing a visualization as a cognitive group awareness tool (CGAT), which had a dedicated learning phase in the learning design. Based on theses approaches, a classroom experiment has been conducted. This chapter is based on a conference paper for the International Conference on Collaboration Technologies and Social Computing (Erkens et al. 2019).

Finally, chapter 8 summarizes the outcome of this thesis and characterizes the contributions to the field of technology enhanced learning (TEL) and computer-supported collaborative learning (CSCL).

2 State of the Art

The following chapter gives an overview of the scientific background. The thesis is contextualized in the research area of inquiry-based science education and inquiry-based learning with online laboratories. The approach of managing diversity to create conditions that influence learning positively is based on the field of learner modeling and knowledge representations. These are mainly influenced by the research done in intelligent tutoring systems. The extraction of knowledge from the learners in order to perform the modeling of diverse knowledge was lately discussed in the uprising field of learning analytics. Managing diversity is twofold: on the one hand, teachers can be supported through methods that help to orchestrate group learning, particularly through automatic group formation. On the other hand it might be a useful tool to mirror the diversity into the learning situation. The field of cognitive group awareness discusses how visualizations of knowledge can be used in small learning groups in order to reduce organizational aspects of group work. The applications and experimental studies are situated in the European project Go-Lab (2012-2016) and utilize the Go-Lab ecosystem, including the Go-Lab virtual learning environment. Here, learning activities with the respective processes and learner-generated artifacts are tracked through a learning analytics architecture, which has been developed as part of this thesis (compare section 3). This architecture serves as a basic platform for the implemented technologies and applications.

2.1 Scientific Inquiry and Inquiry-based Learning

In contrast to traditional education based teacher-centered methods, in modern pedagogy several other approaches exist, which are student-centered and involve active learning. "Scientific inquiry requires the use of evidence, logic, and imagination in developing explanations about the natural world" (Newman Jr et al. 2004). The National Research Council of the United States of America defines inquiry as the following (Council 1996):

> "Inquiry is a multifaceted activity that involves making observations; posing questions; examining books and other sources of information to see what is already known; planning investigations; reviewing what is already

> known in light of experimental evidence; using tools to gather, analyze, and interpret data; proposing answers, explanations, and predictions; and communicating the results. Inquiry requires identification of assumptions, use of critical and logical thinking, and consideration of alternative explanations."

The role and importance of science laboratories have been pointed out quite frequently and developed throughout the majority of science curricula (Hofstein 2004; Eilks et al. 2013; Lunetta et al. 2005). Experimentation in a science lab incorporates social processes, senses and active learning. Many science educators understand constructivism as a theoretical model for science teaching (Lunetta 1998) and consider it as a dominant paradigm in the field of science education research (Lakatos 1970). Sigel (1978) defines constructivism the following:

> "Constructivism refers to that process of constructing, in effect, creating a concept which serves as a guideline against which objects or people can be gauged. During the course of interactions with objects, people, or events the individual constructs a reality of them... This mental construction then guides subsequent actions with the object or events."

In a science laboratory, the students interact with the lab equipment, investigate and build their own understanding based on their experience. These concepts were build historically on constructivist theories. They emphasize that learning is, according to their view, an individual, active and interpretative process. Tobin (1990) highlights the positive opportunities that come from a constructivist perspective: "Laboratory activities appeal as a way to learn with understanding and, at the same time, engage in a process of constructing knowledge by doing science". However, the science laboratory has been criticized in science education research. Although the distinctive role for scientific discovery has been acknowledged, the lab has been seen as confusing and unstructured (Hodson 1993). Research has found evidence that there is a gap between learning goals and learning outcomes (Goodlad 1983; Lunetta et al. 2007; Hodson 1993, 2001; Wilkinson and Ward 1997). Many teachers do not find that their stated lesson aims to be addressed during the lessons that involved laboratory work (Hodson 2001). This criticism highlights the need for support mechanisms tailored to scientific inquiry (see section 2.1.1).

Barrow (2006) described that during the historical development of the term inquiry and the concept describing it, there have been different discussions and a thematic shift towards scientific inquiry. Several attempts to define the term have been made, complemented by operationalizations and pedagogical models of inquiry-based learning and inquiry-based science education.

In an early historical perspective, John Dewey recommended the inclusion of inquiry into science curricula. Dewey points out that "science teaching has suffered because science has been so frequently presented just as so much ready-made knowledge, so much subject-matter of fact and law, rather than as the effective method of inquiry into any subject-matter" (Dewey 1910). At this time, the common practice in science education was teacher-centered, passing instruction in a passive, lecture-style way. He further states out that "when our schools truly become laboratories of knowledge-making, not mills fitted out with information-hoppers, there will no longer be need to discuss the place of science in education."

Jerome Seymour Bruner, one of the pioneers in cognitive psychology, viewed sensation and perception as being active processes. Bruner (1961) outlined in his early works *the act of discovery* the positive aspects of "the experience of learning through discoveries that one makes for oneself". In his paper, he focused on learning through making discoveries on an individual level. Bruner illustrated this in an experiment with children, who had to remember pairs of words and recall them later. A second group of children were told to remember them by producing a word-pair or idea in a way that it makes sense to them. The word pairs included juxtapositions of words that reflected different individual preferences of the children. Bruner points out that in this second group, the recall was much higher. The design of this experiment, particularly the characterization of the groups implies criticism to traditional, teacher-centered pedagogy, which does not put emphasis on individual construction of knowledge. Bruner has been criticized for excluding social processes from learning in his work. However, it can be seen as one of the early influential works for IBL as it shows the potential of shifting the role of teaching and learning towards a constructivists perspective (Bruner 1987; Dobber et al. 2017).

These "historical" constructivist approaches tended to have limitations. While the work by Bruner focused on individual learning excluding social factors, the work by Dewey mainly put emphasis on active learning and excluded passive stimuli for learning. It is important to state out that modern inquiry-based learning is based on socio-constructivist approaches or socio-cultural models that do not exclude passive learning or social factors (Kruckeberg 2006; O'loughlin 1992). Saunders (1992) lists four important features of effective science programs that directly stem from the constructivist perspective and that have shown to enhance learning: hands-on investigative labs, active cognitive involvement, group work, and higher-level assessment. This sets a baseline for the design of learning scenarios and tasks that have been used in the experimental studies as part of this thesis.

The field of epistemology is a central philosophy of the foundation of knowledge that underlies scientific methods. This applies both for the foundation of scientific approaches and for scientific learning. Scientific discovery learning is defined by de Jong and Lazonder (2014a) as "a method of learning in which knowledge acquisition is based on the induction of domain rules through structured experimentation." This definition connects socio-constructivists perspective with epistemic aspects of knowledge building in a modern approach for scientific learning (Scardamalia and Bereiter 1991).

The emphasis on knowledge about a domain to be investigated requires the learners to be able to express, construct and externalize their knowledge (van Joolingen and Zacharia 2009; Scardamalia and Bereiter 1991). Particularly in computer-supported inquiry activities, this involves tools for expressing knowledge. Usually, these tools are specific to inquiry activities, such as conceptualization, hypothesis formulation, planning and observing of experimental activities, or drawing conclusions. These tools provide representations that are a medium of the inquiry process for the learners and can help structuring the (inquiry) learning activities (cf. section 2.1.1).

van Joolingen and Zacharia (2009) summarized the elements in the learning environment that can sustain inquiry processes:

- The *mission* of an inquiry activity motivates learners and provides a goal. This does not only involve a domain, but also incentives to ask questions and to illustrate a goal setting.

- The *source of information* for experiments, such as simulations or science laboratories.

- The *tools for expressing knowledge* which enable learners to externalize, communicate and negotiate their knowledge, for example through creating models or explanations.

- The *cognitive and social scaffolds*, which help students to perform inquiry processes that they usually would not be able to perform.

These four elements are relevant aspects for the underpinning of this thesis. The first and the second aspect influence the design of learning scenarios for the experimental studies as part of this work. The third and the fourth aspect, tools for expressing knowledge and cognitive scaffolds, are crucial for the conceptual part of this work. The proposed approaches for knowledge diversity take such knowledge representations and models from inquiry activities as an input in order to generate results that can enhance learning through facilitating knowledge diversity. Frameworks for inquiry-based learning suggest best practices for the design of scenarios, they define how to structure IBL activities, and help to orchestrate inquiry-based science education. The following section describes frameworks and models for IBL with a focus on the inquiry cycle, which has been adapted for Go-Lab.

2.1.1 Inquiry Cycle

Contrasting to traditional teaching methods, in IBL learners pose questions and answer them through scientific discovery and with the use of scientific methods. The activities and phases which are typical for scientific approaches are connected to epistemic practices and are reflected in cyclic models. Inquiry-based learning ("IBL") is a pedagogical method and form of active learning that combines scientific processes with discovery learning based on epistemic aspects such as observation, evaluation and knowledge building. One of the goals is to encourage young people to work in scientific jobs and thus become the next generation of researchers and scientists (Gago et al. 2004; Council 1996). Research has shown that discovery-led, structured teaching, incorporating feedback, working examples, scaffolding, and elicited explanations, goes along with an improved learning gain compared to other teaching practices, such as explicit instructions or unassisted discovery (Alfieri et al. 2011). Several models of inquiry-based learning have been created in order to define best practices for teaching.

Dewey described education as the collaborative reconstruction of experience. Furthermore, he states out that scientific learning should be authentic to science practice (Dewey 1959, 1910). Garrison et al. (1999) picked this idea up and presented the model of practical inquiry. This model incorporates socio-constructivist principles ("shared world" and "social presence") with a model of guidance (discourse and reflection). In contrast to learning by design, it contains a view on "teacher presence", which advocates adaptive planning by the teacher and emphasizes the role of the teacher as an mediator.

The knowledge-building community model can be seen as a much more open ended version, where knowledge construction through collaborative inquiry is a collective goal of the learners. Scardamalia and Bereiter (1994) incorporated principles of constructionism, socio-cultural activities and apprenticeship into this model. The idea of Computer-Supported Intentional Learning Environments (CSILE) is that schools use internet and communication technology to function as knowledge-building communities comparable to knowledge-advancing enterprises (Scardamalia and Bereiter 2006). However, this model was not intended to be a pure model for inquiry-based learning, although it shares a lot of common aspects.

A widely adopted model for web-based IBL is the *Inquiry Cycle*. This model has been used as a pedagogical foundation in the Go-Lab project (cf. section 2.2.3), which is the target web-based IBL platform for the work of this thesis. The inquiry cycle (IC) structures learning activities in a circular flow. Figure 2.1 shows such a circular inquiry model by White and Frederiksen (1998). Each sub-activity (phase) is called inquiry phase. Each phase in this cycle has certain activities and subphases bound to pedagogical decisions about the particular inquiry-related task or problem. The goal of this cycle was to create an instructional model that corresponds - in essence - to scientific methods:

> "This instructional cycle of getting students to make predictions, do experiments, formulate laws, and investigate the generality of laws resembles the classic conception of the scientific method. We thereby created a correspondence between the phases of the instructional cycle and the process of scientific inquiry. As students participate in the instructional cycle, they are introduced to the construction of scientific theories as well as to a conception of what scientific inquiry entails." (White 1993)

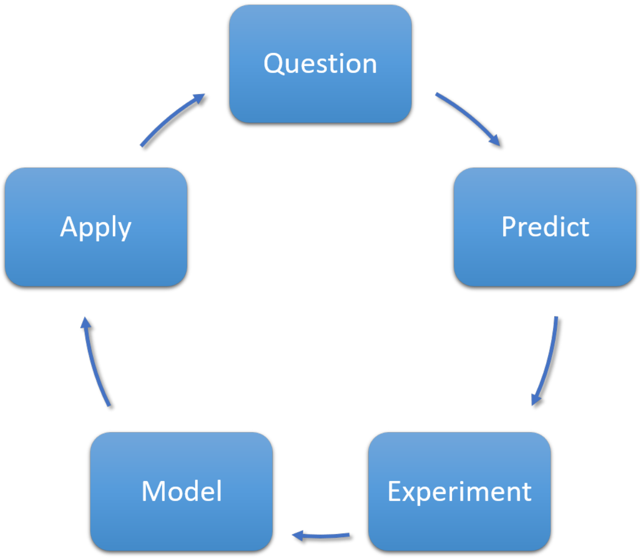


Figure 2.1: The inquiry cycle by White and Frederiksen (1998).

The structure of the inquiry activities can vary depending on the concrete model that has been chosen. In online learning environments, these models differ depending on the concrete approach. For example, in Go-Lab, inquiry activities are structured as subsequent phases and represent a linearization of a cycle and do not contain multiple cycles. There is not a single inquiry cycle that has been accepted as a predominant model - More than 32 different cycles exist in the literature. Pedaste et al. (2015) provide a review of 32 articles that described inquiry phases or cycles. By analyzing these articles, five distinct phases have been identified: *orientation*, *conceptualization*, *investigation*, *conclusion* and *discussion*. This synthesized framework is used as a pedagogical basis for structured and scaffolded inquiry activities, such as the Go-Lab project. They described the subphases and suggested for each phase specific cognitive scaffolds are typical for supporting learners in their inquiry processes. A cognitive scaffold can be a (pedagogical) tool or an app as the digital counterpart of such tool, which serves as a support function in the inquiry process. Such cognitive scaffolds can be, for example, texts, hypotheses, or concept maps.

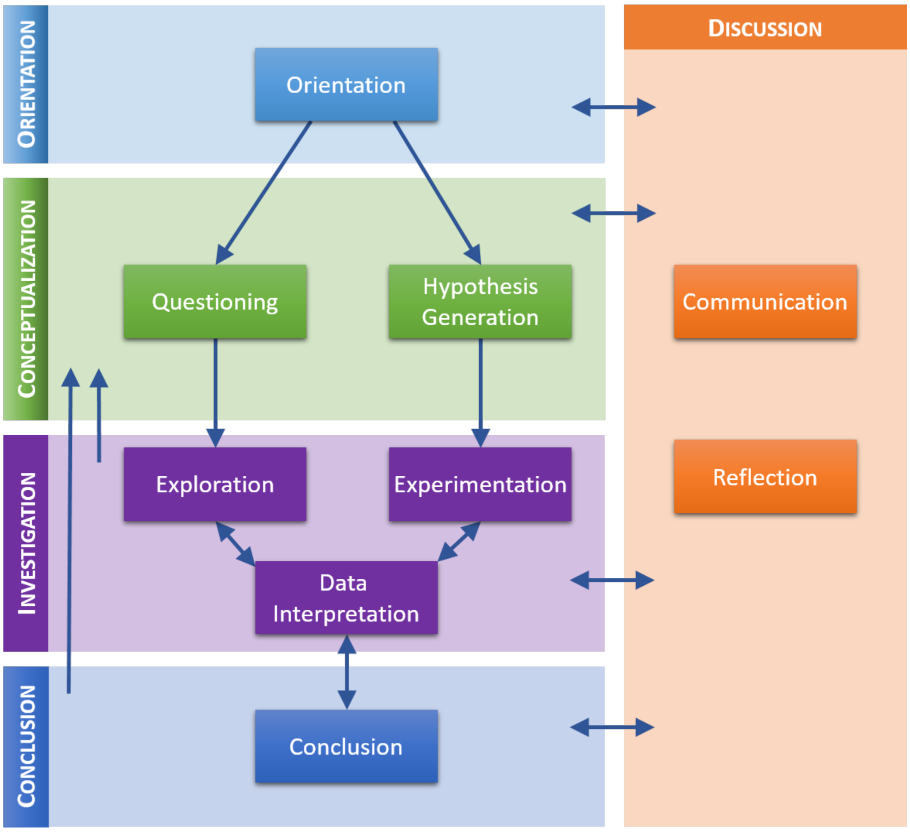


Figure 2.2: The synthesized model of inquiry cycles according to Pedaste et al. (2015).

Figure 2.2 shows the model that has been created as a synthesis of existing approaches of inquiry cycles by Pedaste et al. (2015). Each arrow lines out a possible transition between phases in the design of an inquiry activity. It consists of the following phases:

Orientation. The phase of orientation mainly serves the purpose to catch the attention and to attract the curiosity of the learners. By introducing big ideas of science or outlining challenges in a problem statement, students' interest on a particular topic will be raised. Typical for such a phase are learning materials that engage learners to further explore a topic and to get in touch with scientific-oriented questions. These questions can be picked up or concretized in the next phase, the conceptualization.

Conceptualization. One of the key aspects of scientific learning is *asking questions* (Hofstein et al. 2005; Windschitl et al. 2008; Crawford 2000). The phase of conceptualization consists of steps that are preceding the experimentation, particularly in asking questions and generating hypotheses. In this phase it is helpful to use mind tools which help structuring and externalizing the learners' knowledge. Tools which support the externalization of knowledge are concept maps or learner-generated texts. In an ideal case, the transition to the experimentation phase is smooth when the learners have the ability to articulate or externalize their (domain) knowledge. One of the challenges for learners in this phase of IBL is the formulation or verification of hypotheses. This results from a lack of scientific vocabulary, missing operators or quantifiers, the difficulty to express scientific processes, and relations between variables (Dunbar 1993; Sandoval and Reiser 2004; Quintana et al. 2004). Windschitl et al. (2008) found out that many students believed "that hypotheses function as guesses about outcomes, but are not necessarily part of a larger explanatory framework. They believed that science studies culminate in 'conclusions' that merely summarize trends in the data, and many thought that making claims that attempt to link data with unobservable processes was recklessly speculative."

Investigation. The investigation phase is a central part of IBL: learners conduct an experiment in order to test their hypotheses. Figure 2.3 shows the subphases and activities during the investigation phase. Depending on the openness of inquiry, the planning can involve the design of an experiment, defining the parameter variations or methods to explore specific aspects. The experimentation can take place in a real science laboratory, but the advances in computer-supported inquiry learning also introduced virtual and remote science laboratories. Alternatively, a variety of prepared data sets from real experiments, which are significant to science, can be a embedded. Including data sets instead of online laboratories puts a focus on analytical and interpretative activities. For example, the data from the ATLAS experiment at CERN have been made accessible for learning (Barnett et al. 2012).

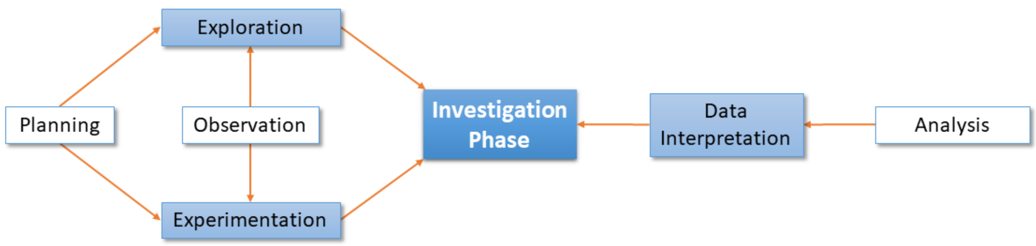


Figure 2.3: Activities of the investigation phase according to Pedaste et al. (2015).

A central part of IBL is experimentation and investigation. In traditional IBSE this was done using real laboratories. With the advent of online learning, web-based inquiry learning environments cover the digital counterparts of science laboratories, particularly virtual and remote laboratories. Broader definitions also include data sets (de Jong et al. 2014), which can be useful to support methodological competencies and evaluation skills, without demanding all of the skills needed for open discovery from the learners.

Conclusion. Based on the previous phase, the learners connect explanations with scientific knowledge in order to draw conclusions from the inquiry activity. Such explanations or conclusions might lead to new insights, a new or refined theory. The inferences from the data might be based on hypotheses, data, models or research questions.

Discussion. The discussion can contain the two subphases communication and reflection. Communication involves activities to disseminate the findings of certain phases or the conclusions of the inquiry to the other learners. Reflective activities are mainly viewed "as an internal process (...) focused on the inquiry-based learning process and communication on domain-related outcomes of this process" (Pedaste et al. 2015).

Scaffolding and Guidance in IBL

The role of supporting learners in inquiry-based learning has been highlighted quite frequently (Kirschner et al. 2006; Reid et al. 2003). Mayer (2004) points out drastically "that the formula *constructivism = hands-on activity* is a formula for educational disaster." Without a thought-out concept of how to guide learners through their inquiry or discovery, they might not be able to find or filter the appropriate information or approach the content to be learned. However, there is a difference between discovery learning and inquiry-based learning, where the latter is already highly scaffolded, for example, through the inquiry cycle (Hmelo-Silver et al. 2007).

Although computer supported inquiry learning (CoSIL) has demonstrated its usefulness, the research in this area points out that learning with such environments poses challenges for many students due to the complexity of these environments. This depends on - but does not exceed - the difficulty to find all variables involved in physical phenomena, to identify all relations corresponding to these variables, or to select the right parameters to manipulate and vary in an experiment (Azevedo 2005; Scheiter and Gerjets 2007; Swaak and de Jong 1996; Zacharia and Olympiou 2011; Marshall and Young 2006).

Apart from the challenged posed through the complexity of the learning environment, there are cognitive and metacognitive challenges for the learners (Zacharia et al. 2015). Pedaste and Sarapuu (2006) proposed a support system for inquiry learning in such an environment. The web-based inquiry-learning environment "Hiking Across

Estonia" was enriched by a support system for learners that increased the problem-solving ability and analytical skills. Due to the nature of scientific discovery, students are facing problems particularly in the fields of hypothesis generation, the design of experiments, interpreting data collected during the experiments or in advance, and the regulation of learning (de Jong and van Joolingen 1998).

Vygotsky developed - with respect to Piaget's stages of child development - the *zone of proximal development* (Vygotsky 1980). While Piaget sees more the individual exploration as the main part in childrens' learning, Vygotsky incorporates social factors and knowledge co-construction into his theory. Assuming that there are certain things a learner can achieve with the own knowledge, and there are certain things that cannot be achieved, the zone of proximal development is usually outlined as in between these two areas. With help or guidance, the learner can extend the own abilities to achieve the goal or to solve a task. Although this concept has caught a lot of attention from cognitive and developmental psychologists, it has been criticized a lot for a weak operationalization of the main constructs of the zone, such as guidance or the term development itself (Wertsch 1984; Tudge 1992; Bruner 1984). However, the idea behind the zone of proximal development underlines the role and effect of scaffolding for learning.

The Cambridge handbook (Sawyer 2005) of the learning sciences has a broad definition of scaffolding: "Scaffolding is the help given to a learner that is tailored to that learner's needs in achieving his or her goals of the moment." Furthermore, it states that "effective learning environments scaffold students' active construction of knowledge in ways similar to the way that scaffolding supports the construction of a building." The term scaffolding raised a lot of discussions about the less precise meaning and the divergent use in different contexts. Puntambekar and Hubscher (2005) found out that the notion of scaffolding changed from the aforementioned metaphor of the construction site to features of "diagnosis, calibrated support, and fading." In latest research in the field of computer-supported inquiry-learning, there is a consensus that guidance is the more general term in this typology (de Jong 2006; Lazonder and Harmsen 2016; Lazonder 2014). Scaffolds are seen as more specific features: "scaffolds explain or take over the difficult parts of the activity; when the learner's skill level increases, the scaffolding is gradually removed so that the learner eventually performs the activity without assistance" (Lazonder and Harmsen 2016). The frequently cited classification by Quintana et al. (2004) has been consolidated and refined by de Jong and Lazonder (2014a). Lazonder and Harmsen (2016) amended this schema and conducted a meta-analysis to create a typology of guidance:

Process constraints. Process constraints are a less specific type of guidance, where the learning process itself runs under certain restrictions. Such guidance mechanisms restrict the number of pathways, of features, or the series of subtasks or elements the learner investigates (Lazonder and Harmsen 2016; White 1993; Rieber and Parmley 1995). For inquiry-based learning, this could mean that the learner takes part in certain sub-activities without any possible alternatives or "detours" through a given and fixed pathway. For Go-Lab, this is enforced (a) by the learning design, and (b) through particular apps and their configuration. By designing an ILS with respect to specific inquiry phases and the apps that define the corresponding activities, a structure of the learning process is already given. The constraints do not restrict all pathways as the learner is allowed to jump forth and back in the navigation that comes from the phase structure. Besides this, the apps can be configured further. For example, a concept mapping tool can be configured to contain predefined concept and relation labels.

Status overviews. Learner progress, states of inquiry or knowledge can be displayed to the learners in the form of status overviews. Typically, such information does not interfere with the idea of inquiry learning, so it should not display any instructions or direct learning materials, but make the learning progress visible. These tools are usually embedded "on demand", so the learners decide how to use the information from the status overview for further processing. For example, a tool that visualizes the contributions to a certain learner-generated content by the participants in a collaborative scenario. Status overviews can be realized as performance dashboards, which are monitoring tools displaying progress information about learners, for example knowledge evolution or learning progress. SimQuest offered a monitoring tool in which learners could review, compare and replay their experiments (van Joolingen and de Jong 2003).

Prompts. Lazonder and Harmsen (2016) describe prompts as "timed cues, either given by a human being or embedded within the learning environment, that remind the learner to perform a particular action." In contrast to status overviews, prompts are more specific and direct the learners action to perform (what to do) a certain task without giving too detailed instructions (how to do it). Typically, such directive instructions are timed and preceded by a certain event or action pattern. Prompts are more specific than status overviews because they tell the learner what to do (but not how to do it) at appropriate moments during the inquiry. They have a long tradition in inquiry-based learning. A typical distinction is between cognitive and metacognitive scaffolds (Wichmann and Leutner 2009; Wichmann et al. 2010).

Heuristics. According to Holton and Clarke (2006), guidance through heuristics “relates to the development of heuristics for learning or problem solving, that transcend specific content." In addition to prompts, "heuristics remind learners to perform an action and point out possible ways to perform that action" (Lazonder and Harmsen

2016). An example for heuristics is VOTAT ("vary-one-thing-at-a-time"), which suggests strategies for parameter variation in the experimental design or the experimenting behavior of the learner. In the case of VOTAT, research has shown that presenting the heuristics explicitly can facilitate more self-regulation in students (Veermans et al. 2006).

Scaffolds. Scaffolds can restrict the comprehensiveness of tasks or can reduce the cognitive load and complexity in a domain, give hints towards possible solutions as well as provide affordances to perform actions (Podolefsky et al. 2013; Hmelo-Silver et al. 2007). In summary, cognitive scaffolds bridge the gap between open inquiry and strict teacher-centric education, weakening the barriers and challenges that scientific learning poses upon the students. The "hypothesis scratchpad" is one example of a (cognitive) scaffold helping learners to generate hypotheses. This is achieved by providing a certain prestructuring of hypotheses that includes already predefined conditions, variables and relations (de Jong 2006). Particularly cognitive scaffolds help to structure learners' knowledge and therefore are externalizations of their knowledge. Section 2.1.2 presents some cognitive scaffolds that relate to knowledge models.

Explanations. Explanations are a low-level type of guidance, which offers learners a concrete specification of how to perform an action. An explanation can be timed or not, integrated into a specific phase, or could be performed before the use of complex systems. Therefore, explanations offer the most specific type of guidance. The SIMQUEST notebook was a tool to display (background) information directly to the learners without directly spoiling the inquiry (Veermans 2002). However, such a scaffold can be useful, if the learner has no prior or just a little knowledge in the domain of the inquiry.

Apart from this typology of guidance, other characterizations of learner support exist in the literature. Podolefsky et al. (2013) subdivided mechanisms of guidance into implicit and explicit scaffolds. Applied to this scheme, prompts, heuristics, and explanations count as explicit, while process constraints are more implicit. Other distinctive features of guidance mechanisms have been introduced by de Jong and Njoo (1992): They differentiate support regarding the unobtrusiveness and subtlety. Overt support is presented directly to the learners in a way that it is immediately recognizable as a means of guidance, for example in a help file. Covert mechanisms are more subtle and unobtrusive. This could be realized through slight and gradual changes to the learning environment, for example a fading out of the guidance to increase the complexity of the learning environment. Particularly for scientific discovery learning, Reid et al. (2003) extended the typology through experimental support, interpretative support, and reflective support. Interpretative support targets to support the access and activation of knowledge in order to build a coherent understanding. Particularly, this has an effect on the construction of hypotheses, which requires learners to grasp the systematic dependencies and relations of variables and parameters in a certain domain for a successful inquiry. Experimental support provides tools to guide learners through the challenging task of predicting and observing outcomes of an experiment, or to draw reasonable conclusions. The knowledge integration and abstraction of the discoveries can be improved and ex-post through reflective support. In order to abstract and reflect, it facilitates mechanisms to target the learners' self-awareness and the awareness of the learning process.

Levels of inquiry

One of the main challenges for the orchestration of IBL is the degree of freedom exposed to the learners. While they are far away from a research-oriented daily practice, it is quite difficult for them to pose questions, choose appropriate methods for investigation and most obviously to verbalize and formalize their approaches and results. The degrees of freedom in IBL, the so-called levels of IBL range from "confirmation inquiry" to "open inquiry" (Banchi and Bell 2008). *Confirmation inquiry* aims for introducing inquiry skills such as collecting data. Therefore, the question, the method and also the results are known in advance to focus on the particular inquiry skill. This kind of inquiry activity is suitable for novices in laboratories as it is limited to a single component in a scientific processes, particularly the data collection. *Structured inquiry* adds another degree of freedom on part of the learners when they have to find their own explanation for the evidence from their data collection. The next level, *guided inquiry*, does not prescribe the procedure - students choose their own method for the investigation. This involves, for example, hypothesis creation, data analysis and drawing conclusions. The *open inquiry* is the most challenging level, where the learners have to pose their own questions and develop their own procedure. The openness in both the learning and scientific context, the lack of cognitive tools, and the missing cross-domain knowledge demanded in scientific processes limits the usefulness of such IBL. The irregularity to traditional teaching approaches and the variety of skills, on a cognitive as well as on a metacognitive level, create a tension for teachers' practice. Yoon et al. (2012) highlight the role of hypothesis-based inquiry, which is an important approach to promote science skills, critical thinking and problem solving skills. Difficulties in inquiry-based science education "on the lesson" and "under the lesson" occur due to the complexity of IBL, but these issues have to be addressed in teacher education programs. Therefore, open inquiry is not common in regular science curricula (Roth and Bowen 1995; Hofstein 2004). However, science fairs are to some extent forms of open inquiry in science classrooms. Some examples of these science fairs are the "European Union Contest for Young Scientists"[1], the "Intel Inter-

[1] EU Contest for Young Scientists (EUCYS): `http://ec.europa.eu/research/eucys/index.cfm`, retrieved 2018-08-27.

national Science and Engineering Fair"[2], or "Jugend forscht"[3], a national competition in Germany. Roth and Bowen (1995) have shown an successful embedding of open inquiry into science classes and investigated in knowledge construction and how students' understanding changed over time. The students were introduced to the topic following the model of cognitive apprenticeship and were assigned to small groups. Therefore, their work is quite inspirational for this thesis, as Roth and Bowen investigated in small groups in inquiry-based learning scenarios and found out that the knowledge construction was influenced positively. When students had different understandings of concepts, they negotiated about it. However, such IBSE do not rely on virtual learning environments and do not facilitate ICT. These results and conclusions form a conceptual baseline, but the transfer to the digital world can be done to some extent only, with limitations.

2.1.2 Cognitive Scaffolds and Guidance in IBL

Knowledge is often seen as a necessary requirement for doing inquiry and tools to externalize mental models, such as concept mapping, have the potential to scaffold the learning activity and the particular learning process. The so-called cognitive scaffolds support the learners in their inquiry process (see section 2.1.1). According to van Joolingen (1998), cognitive tools, "defined here as instruments that support or perform cognitive processes for learners in order to support learning, can bridge the difference between open learning environments, like discovery learning environments and traditional supportive instructional environments."

Typically, such cognitive scaffolds do not simply (pre-)structure the process, but furthermore integrate into more atomic activities, for example the creation of hypotheses, the structuring of knowledge through concept mapping or in the form of a free text. Beyond this, for the learners these are powerful tools to express their knowledge. As tools to externalize the knowledge, they support the verbalization and operationalization of the target representation. They are specific in the semantics and syntactic structure of each tool, as they are connected to particular scientific inquiry activities and enable the learners to operate on domain knowledge.

Concept Mapping

Concept mapping (Novak 1984) is a technique for externalizing knowledge structures in the form of semantic networks. Combined with computer-based representations

[2] About Intel ISEF: `https://student.societyforscience.org/intel-isef`, retrieved 2018-08-27.

[3] Jugend Forscht: `https://www.jugend-forscht.de/`, retrieved 2018-08-27.

and tools (Novak and Cañas 2004; Cañas et al. 2004), concept mapping has permeated many areas and various scenarios of technology-enhanced learning. Schwendimann (2015) provides a quite comprehensive overview of pedagogical applications and functions of concept mapping with a special focus on knowledge integration. He distinguishes the activities of map generation, interpretation and revision and discusses the correspondence of knowledge integration processes with certain concept mapping activities. E.g., concept mapping can be used for "knowledge elicitation" and thus as a test of the learner's understanding of a certain knowledge domain. Accordingly, the extension and refinement of concept maps corresponds to a further differentiation of knowledge.

Concept maps reflect the structure of domain knowledge of individual learners. According to Stoddart et al. (2000) these artifacts are particularly well suited as an add-on to other types of test to identify and diagnose students' knowledge. Figure 2.4 shows a concept map that has been created in Go-Lab. In the analysis of concept maps, we have to distinguish semantic and structural aspects. The "semantic richness" of a concept map (possibly in terms of concepts and relations) could be determined by comparison to a domain ontology. In the structural perspective, we would use graph-based measures to characterize features such as the complexity, cohesion or density of a map. We employed several of these measures for the assessment of concept mapping skills in section 5. The concepts and relations used in a concept maps can be prescribed through the learning design as an approach to guide the learner.

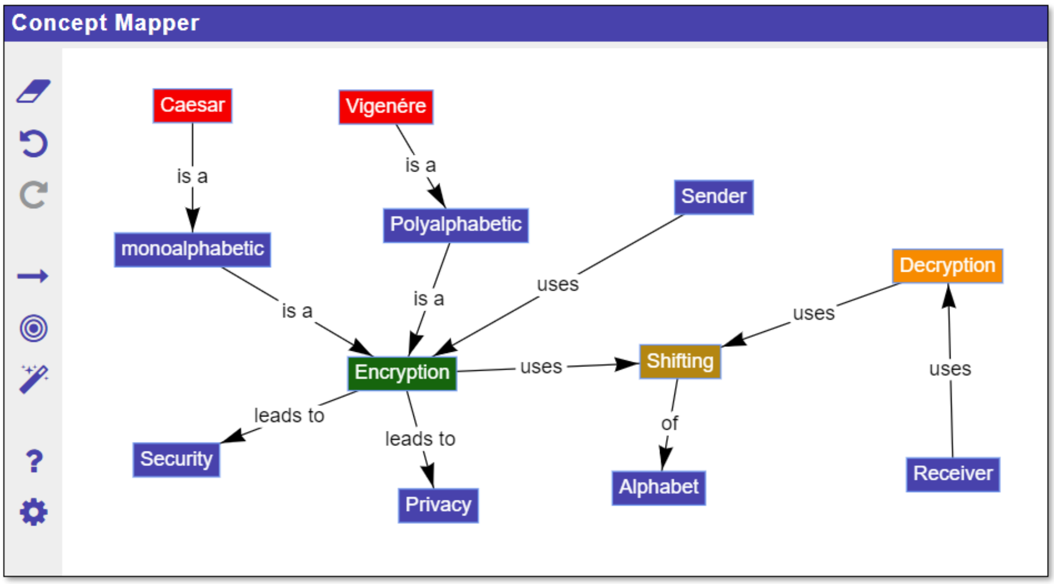


Figure 2.4: A concept map created in Go-Lab.

Hypothesis

Klahr and Dunbar (1988) describe scientific discovery as a dual search problem in the two spaces, "hypothesis space" and "experimentation space". Furthermore they state out that "search in the hypothesis space is guided both by prior knowledge and by experimental results. Search in the experiment space may be guided by the current hypothesis, and it may be used to generate information to formulate hypotheses." This highlights that both formulating and reformulating hypotheses encode learners' knowledge.

Van Joolingen and de Jong (1993) describe the process of hypothesis generation as "traversing variable and relation space", which are the two sub-spaces of the hypothesis space according to the model by Klahr and Dunbar (1988). For this purpose, they created a software tool, the *hypothesis scratchpad*, to support the process of hypothesis generation by offering a prestructured format of hypotheses. Additionally, the hypothesis scratchpad displays the variables and qualitative relations in the given domain (van Joolingen et al. 2005). An example of such a hypothesis in the context of buoyancy could be: "If object's density is smaller than liquid's density, then the object will float."

In contrast to previously specified hypothesis, this enables the learner to explore the relations between variables and also to define key variables for the experimentation. The recent research using the scratchpad also involved the students' propositions regarding the hypotheses (van Joolingen et al. 2005) or the respective confidence. Such belief meters are relevant to the research in knowledge construction (Lajoie et al. 2001). However, we exclude beliefs from our optimistic knowledge modeling approach (cf. section 4), as we assume that learners only express their knowledge if they have a certain degree of confidence about a concept. In essence, we identify the two spaces "variables" and "relations" as a relevant part of the knowledge model behind a hypothesis. In the context of IBL and the inquiry cycle, the generation of a hypothesis is specific to the conceptualization phase. Figure 2.5 shows a hypothesis in the hypothesis scratchpad.

Texts by Learners

Wiki tools are usually asynchronously collaborative tools for writing and editing texts. Wiki tools are platforms for the collaborative editing and publishing of the texts, the published web site is called a *wiki* and the text is called *wiki article*. The most popular example of a wiki is Wikipedia[4], which is an encyclopedia that has been created by around 35 million users with the MediaWiki software. MediaWiki is the underlying

[4]Wikimedia Foundation, Wikipedia, `https://www.wikipedia.org/`, retrieved 2018-11-09.

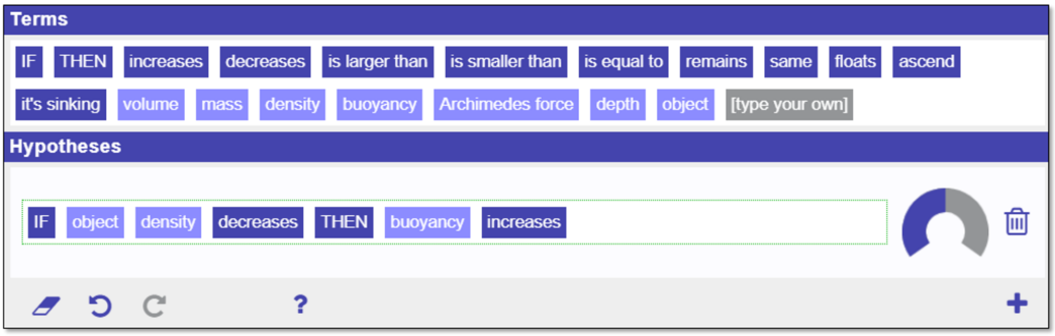


Figure 2.5: The hypothesis scratchpad displaying a single hypothesis: "IF object density decreases THEN buoyancy increases". The confidence meter next to the hypothesis is set to 50 %.

technical platform. However, as part of this work we do not investigate the technical challenges for wikis, as we are focusing on knowledge representations and tools for learners to express their knowledge.

Such wikis have been used as knowledge management platforms (Girard and Girard 2015; Andrus 2005), encoding explicit and tacit knowledge. A lot of research investigated wikis as educational tools, both for teaching and learning (Parker and Chao 2007; Boulos et al. 2006; Duffy 2006). They have shown to improve collaborative learning (Augar et al. 2004). Particularly in the context of communities of practice, wikis demonstrated their usefulness when functioning as a knowledge management tool for the community. Inquiry-based learning has many parallels to this approach and particularly in the inquiry phase of conceptualization, they can be used to structure the learners knowledge, both in collaborative and individual learning phases. Figure 2.6 shows a learner-generated text in the wiki tool that is used in Go-Lab. Similar to other wiki tools, the Go-Lab wiki supports multiple pages. The data can be shared across a learning group or separated for an individual use.

One of the key concepts of wikis is to create an explicit revision of a document. In asynchronously collaborative platforms such as MediaWiki, a user who edits a document locks the document. When the user finished editing the document, the new revision is stored inside the platform and the lock is released. For simple texts in an individual learning scenario, this is sometimes seen as a lot of overhead, particularly for quick notes by learners. Therefore, other alternatives to wiki tools, can be provided.

In contrast to the wiki texts, simple free texts usually do not necessarily have a structure of revisions. Therefore, it is difficult to track progress of learners or assign contributions in a collaborative setting. However, the main relevance for this work is situated in the knowledge representation itself. A text contains a certain number of key

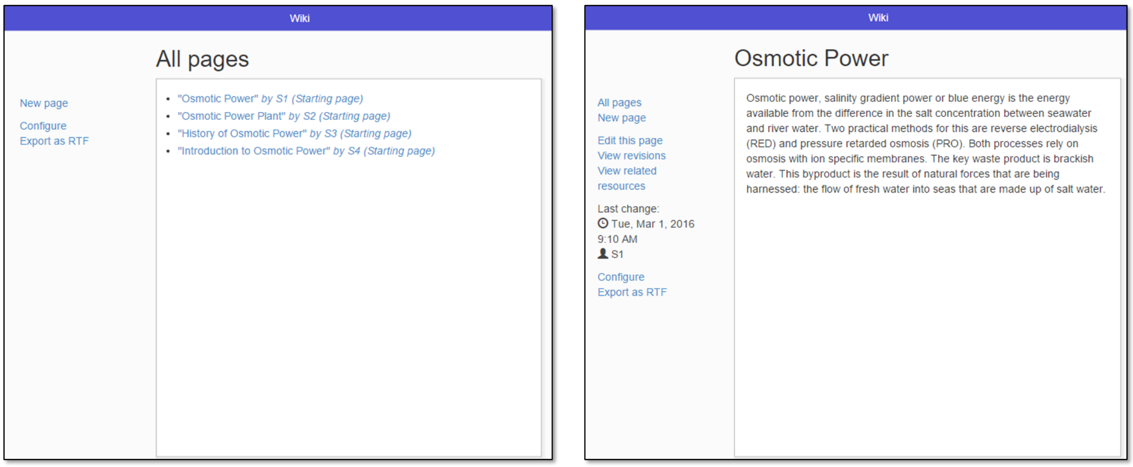


Figure 2.6: Overview of the wiki pages created by a learner (left); one of these wiki pages about osmotic power plants in detail (right).

concepts the learner wants to express. One of the challenges for this work is to extract the key concepts from a text, despite of the nature as a wiki or a free text, and use this as an appropriate knowledge representation.

2.1.3 Experimentation and Online Laboratories

Online laboratories have a special role in the field of inquiry-based learning. The use of simulations and virtual science laboratories has a lot of benefits. The non-exhaustive list of positive aspects comprises (1) an easy setup of experiments he broad availability of complex; (2) expensive laboratories, which can be simulated; (3) the integration into online learning environments; (4) enabling self-regulated and lifelong learning. On a superficial level, online labs are the digital counterpart of real, physical laboratories. But from a constructivist perspective, they do not fulfill the same needs. The degree of interaction is influencing the learning outcome. Wise and Okey (1983) have shown that the use of manipulative activities have been far more effective in terms of learning gain and achievement than observing and reading about phenomena in books.

Online laboratories can be categorized into *virtual laboratories, remote laboratories* and *data sets* (de Jong et al. 2014). A virtual laboratory is the simulation of a science laboratory. It can be distinguished from a pure simulation as it also simulates the laboratory equipment. Typically, the learners have to deal with similar equipment as in a real science laboratory. Figure 2.7 shows the "splash lab" from Go-Lab, which enables learners to conduct experiments about buoyancy. The equipment in this case contains different materials and different liquids. A virtual lab like this lab for buoyancy experimentation simplifies the parameter variations on part of the learners. It also contains a tool to capture and visualize the measurement data, which supports the data interpretation and enables the learner to capture bigger data sets in the same time. The collection of bigger data sets and the reduction of experimental errors make the physical phenomena more transparent.

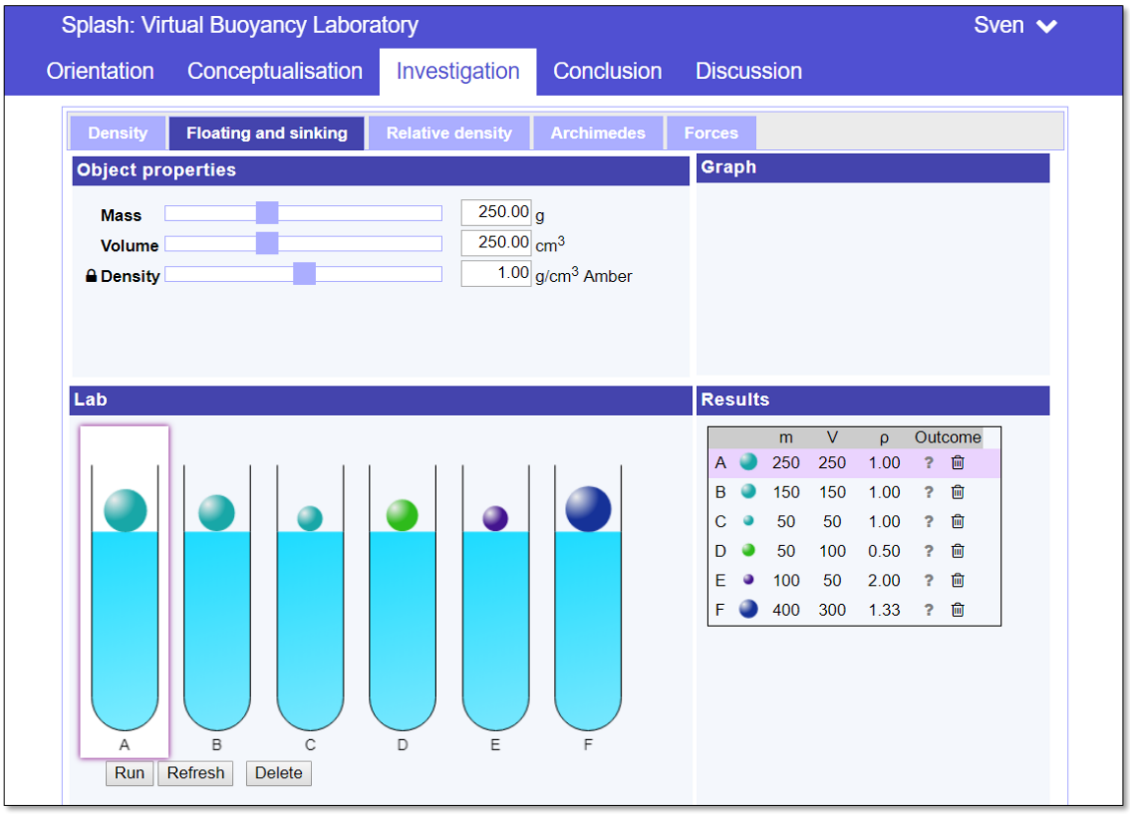


Figure 2.7: The splash lab for buoyancy.

A remote laboratory (sometimes also called "web lab") is the digital extension of a real, physical lab, which is made accessible for the learners through a remote access, typically through a web interface. This scenario of how to operate a remote lab is comparable to the "control of robots used in remote manufacturing" (Ma and Nickerson 2006), where the learners rather operate the robot than the things to be manufactured. A reason for this is an encapsulation of operations for each lab, as it is necessary to minimize or to sandbox specific error sources in order to preserve the maintainability of a lab. In a completely remote and autonomous setting, no one will have the chance or will be responsible to change a light bulb or to repair certain things. Additionally, a remote lab is a dedicated physical resource, which allows only for a unique access. Remote laboratories pose many pragmatic challenges for the maintenance, the deployment and the access of laboratories (Garcia-Zubia et al. 2006; Gustavsson et al. 2007a; Garcia-Zubia et al. 2009; Sancristobal et al. 2011; Orduña et al. 2012; Santana et al. 2013). These pragmatic issues also affect the software engineering practice. Therefore, it has been elaborated as a technical approach to deploy remote laboratories through a specific gateway in order to scale it up (Melkonyan et al. 2014). Figure 2.8 shows a radioactivity lab with an embedded view of the experimental results. This example highlights another positive use case for remote experimentation, which regards safety issues and the expense of equipment.

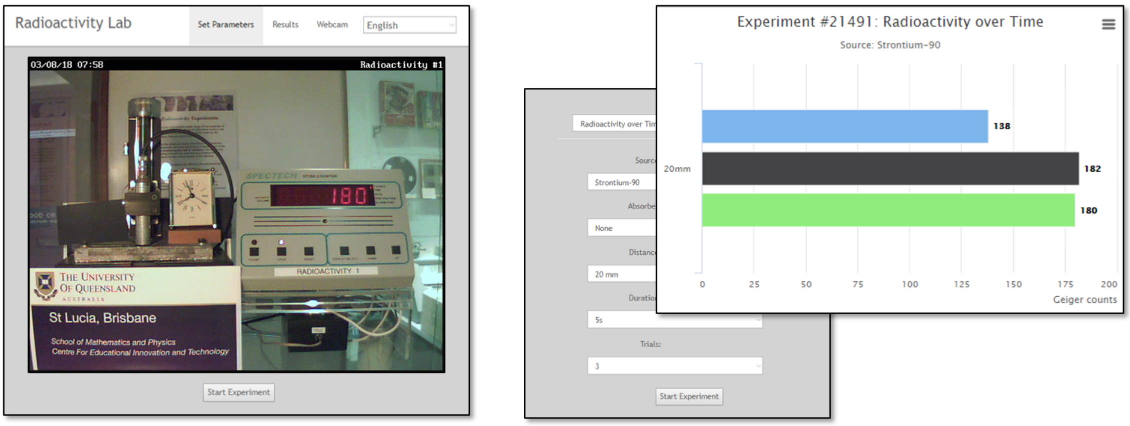


Figure 2.8: A screenshot of the radioactivity lab by the University of Queensland. It contains a control to specify the experimental parameters and an embedded result view of the experimental data.

Data sets are static experimental data that are provided to learners. Although this seems to be the case of an online laboratory with the least interest, it also has an interesting use when it comes to real world experimentation. For example, the ATLAS experiment from the Large Hedron Collider at CERN provided data sets that are highly relevant for recent research (Barnett et al. 2012). By handing this out to students, who are not able to conduct such -in a lot of ways- extremely demanding experiments, they can still draw conclusions and learn about data interpretation in this motivating context. Figure 2.9 show the HYPATIA analysis tool, which is used to analyze the data from the ATLAS experiment at CERN.

Apart from these obvious and pragmatic reasons to use a virtual laboratory, research could confirm the positive effects on students' science learning through simulations (Chang et al. 2008; Huppert et al. 2002; Ingerman et al. 2009; van der Meij and de Jong 2006; Zacharia and Constantinou 2008; Zacharia and Anderson 2003; Zacharia 2005). Computer simulations successfully enhance traditional instruction and are an effective way to prepare laboratory activities (Rutten et al. 2012). In combination with support mechanisms, the outcome of learning with simulations and virtual labs could be even improved. By actively relating multiple representations and integrating external sources of information before using interactive visualizations and simulations, the learning performance can be improved (Bodemer et al. 2005; Bodemer and Dehler 2011).

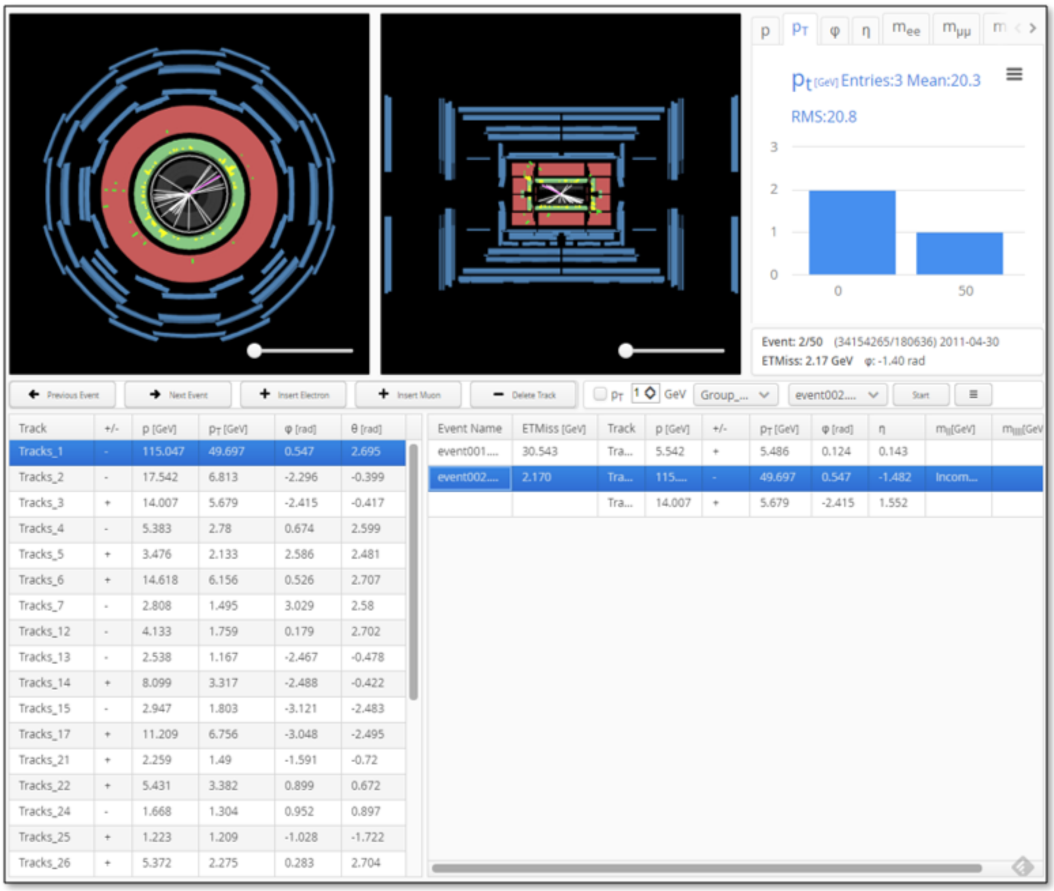


Figure 2.9: A screenshot of the HYPATIA tool to analyze the data set from the ATLAS experiment at CERN.

Compared to traditional instruction, learning with simulations can lead to a better conceptual learning and potentially helps the students to maintain their knowledge over a longer period (Deslauriers and Wieman 2011). In contrast to real, physical science laboratories, there are many positive aspects in conjunction with learning processes. Virtual labs have the potential to improve the conceptual understanding by making invisible things visible, for example magnetic or electric fields. Such visual representations of abstract objects make physical phenomena more transparent for the learners (Olympiou et al. 2013). Particularly, simulations promote scientifically accurate concepts, which emphasizes certain phenomena and supports students in their data interpretation (Zacharia and Anderson 2003).

Scalise et al. (2011) remark that the "development of science simulations and virtual laboratory software is in its infancy". The use of ICT has a lot of benefits, proven by a broad number of studies in research around simulations and inquiry-based learning. Still, the desired learning outcome will be achieved through a diversification of teaching methods, which goes hand in hand with applying frameworks of IBL, incorporating the positive effects of ICT with effective guidance.

However, experimentation, discovery and inquiry are challenging for learners (cf. section 2.1.1). While unassisted discovery does not seem to be beneficial for learners, feedback, worked examples, and guidance mechanisms seem to improve learning (Alfieri et al. 2011). In contrast to traditional self-guided learning approaches such as the Piagetian theories (Piaget and Cook 1952; Piaget 1970), rich information and inquiry environments have the potential to enhance science learning and improve scientific inquiry. Marušić and Sliško (2012) argue that the Piagetian and Vygotskian visions on learning can go hand in hand through class management, highlighting the importance of collaborative learning. Smetana and Bell (2012) published a critical analysis of the research literature about computer simulations in science education and concluded:

> "Computer simulations are most effective when they (a) are used as supplements; (b) incorporate high-quality support structures; (c) encourage student reflection; and (d) promote cognitive dissonance. Used appropriately, computer simulations involve students in inquiry-based, authentic science explorations."

The upcoming section shows examples of computer-supported inquiry-based learning environments that encompass such supportive features and mechanisms of guidance.

2.2 Web-based Learning Environments for Inquiry-based Learning

The technical focus of the thesis is in the Go-Lab system, which provides flexible and powerful authoring tools for creating and enrolling virtual learning environments in the context of inquiry-based science education with online laboratories. However, the work of Go-Lab has been inspired by other projects for computer supported inquiry learning (COSIL), in which other software platforms have been created. The consortium of the Go-Lab project consists of institutes that took part (beside other projects) in the ROLE project (Faltin et al. 2013), the SCY project (de Jong et al. 2010), and different activities using and creating online laboratories. ROLE was mainly focused around personalized learning environments with a powerful sandbox, in which learners and educators could create these learning environments. This mainly inspired and influenced the authoring and run-time of the Go-Lab learning environment, mainly the technical platform Graasp, which has its origins in the ROLE project.

Inquiry-based learning activities are usually structured learning activities that implement an inquiry cycle model (cf. section 2.1.1) (Kuhlthau et al. 2007; Pedaste et al. 2015). Co-Lab was a collaborative inquiry-learning environment using remote and virtual experiments. The inquiry process has been structured in five phases close to scientific processes: analysis, hypothesis generation, experiment design, data interpretation, and conclusion (van Joolingen et al. 2005). Similar to Co-Lab, environments such as WISE (Slotta and Linn 2009), SCY-Lab (de Jong et al. 2010), or JuxtaLearn (Haya et al. 2015) have explicit and rarely flexible IBL models. Despite following a well-known predefined model may be helpful, especially to support novice teachers or students, the use of a rigid model constrains the teachers' chances to customize it to their learning contexts, and some students may have difficulties to adapt it themselves (Dillenbourg 2002).

2.2.1 SCY

The SCY project ("Science Created by You") was much more specific for inquiry-based learning, as it featured an innovative web-based platform, the SCY-Lab, as a rich toolbox for IBL. Figure 2.10 shows the user interface for the learner. The Lab consists of different tools to support the inquiry process through web quests, concept mapping, system dynamics, hypothesis construction, and more. In this active learning approach, the students create artifacts by using the web-based tools. These learning objects emerge through the whole learning process. As the learning objects emerge throughout the process and made available for other tools and also for other learners, they follow the idea of "emerging learning objects", so-called ELOs. The SCY-Lab

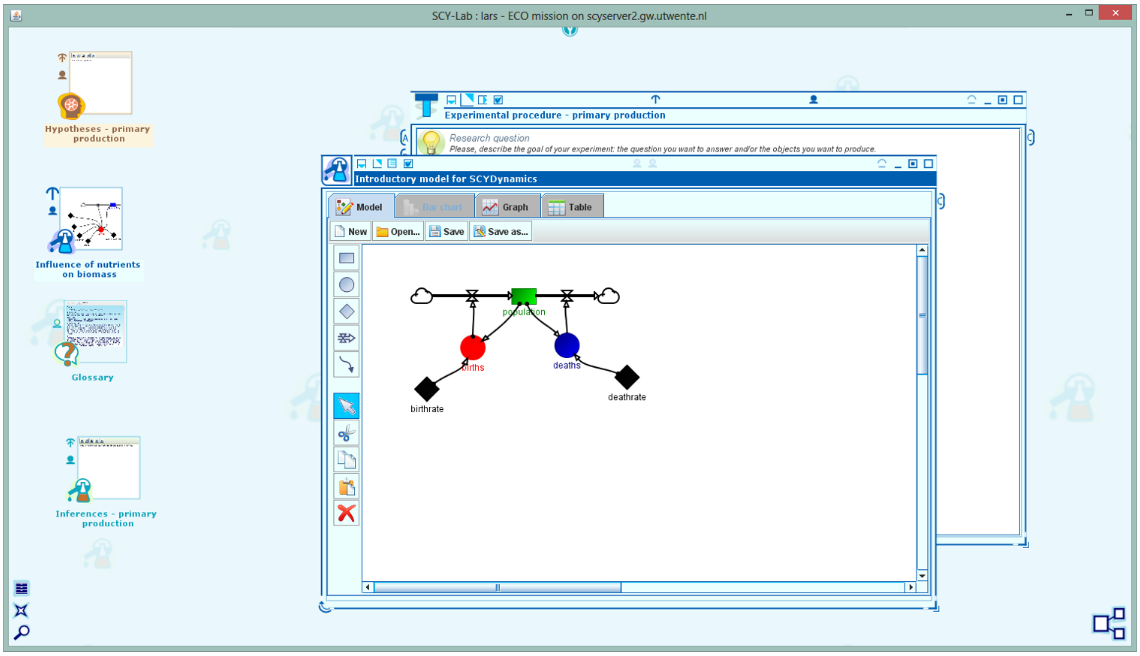


Figure 2.10: The user interface of the SCY-Lab, the user interacts with a SCY-Dynamics ELO (Geraedts et al. 2011).

was not limited to individual learning, it also supports collaborative learning and collaborative construction of ELOs. An example of such a tool that has been evaluated in the SCY project was a real-time synchronous collaborative concept mapping tool ("SCY-Mapper"). Besides performing tasks in synchronous collaboration tools, learners were able to collaborate through artifacts. The ELOs can be found via search and then be forked by other individual learners, which is an asynchronous collaboration through the artifact. On the software level, each ELO could be accessed through a particular tool. Each ELO consists of data and metadata, based on the LOM standard, that describe the learning context of the LO (de Jong et al. 2010; Hodgins and Duval 2002). In total, an instance of SCY-Lab was a collection of ELOs and resources for each learner, while the learning activities are prestructured and visualized in SCY-Lab through a so-called "mission map". The mission map is a special kind of implementing an inquiry-cycle, which explicitly outlines different pathways through the cycle. This preserves a certain degree of freedom for the learner without limiting them to a linear sequence of activities. However, it also scaffolds the process as it explicitly outlines the order of certain activities in a predecessor-/successor-relation.

In SCY, learning activities are grouped in Learning Activity Spaces (LASs), which are a conceptual unit that combine activities and ELOs (de Jong et al. 2010). "A Learning Activity Space (LAS) is defined as a coherent and intuitive set of activities supported with specific tools and scaffolds. The input and output of a LAS are described in terms of a set of artifacts created by students (further called emerging learning objects (ELOs)" (Lejeune et al. 2009a). Such LASs can represent inquiry (sub-)phases, for example "Experiment", where the learner collects data from an experiment and creates an ELO that contains the data set. Such an ELO, which is central for a LAS, is called "anchor ELO". The LASs are combined into a learning scenario, the mission, while all possible sequences are outlined in the mission map, which is a directed graph and also visualized in the user interface to guide learners through the inquiry process. Additionally, it is possible to have different iterations of inquiry cycles or even completely different cycles. An adapted version of the "Eco Mission" was using four different inquiry cycles in different domains, namely a) nutrients and primary production, b) the role of light in ecosystems, c) relationships between trophic levels, d) pH and aquatic ecosystems (Pedaste et al. 2013).

The first SCY mission was the design of a CO_2 friendly house. Figure 2.11 shows the learning scenario of this mission as a composition of LASs, where each LAS groups collaborative or individual learning activities. The activities are related to the scientific inquiry process, for example, the creation of a hypothesis. The duration of the whole learning scenario is approximately 20 hours and the learner productions are mainly done on the SCY-Lab. Since the experiment is conducted offline, data has to be captured through dedicated data collection tools and imported into SCY-Lab. The group activities in this scenario are based on the jigsaw approach (Aronson et al. 1978; Geraedts et al. 2010).

The definition of learning activity spaces is not limited to inquiry-based learning, it can be seen as an abstraction and grouping of activities, emerging learning objects, tools, and scaffolds in the field of science learning. Other examples are design-based learning or argumentative knowledge construction that can be modeled through LASs (Lejeune et al. 2009b). The creation of learning activity spaces has been supported by software tools for graph-based modeling, such as SCY-SE based on FreeStyler (Weinbrenner et al. 2009; Lejeune et al. 2009a). This approach integrates with an IMS-LD editor for scripting the learning design (Harrer et al. 2007), which mainly targets the community of educational designers (Wichmann and Leutner 2009). The scripting of learning design has a long tradition in the field of CSCL, particularly with the purpose of reusability and exchangeability of learning scripts, but it is out of focus of this thesis, as we employ more implicit and intuitive mechanism of learning design as facilitated by Go-Lab (cf. section 2.2.3).

However, the SCY-Lab lacks in the flexibility that instructors or teachers can easily create their own learning spaces with reasonable efforts. The absence of graphical authoring tools for the SCY missions and missing facilities to deploy learning spaces (which involves Java packaging mechanisms and server redeployments) hinder teachers to create custom learning environments. The existing learning scenarios do not cover a variety of topics that align well with the curricula across the countries of the

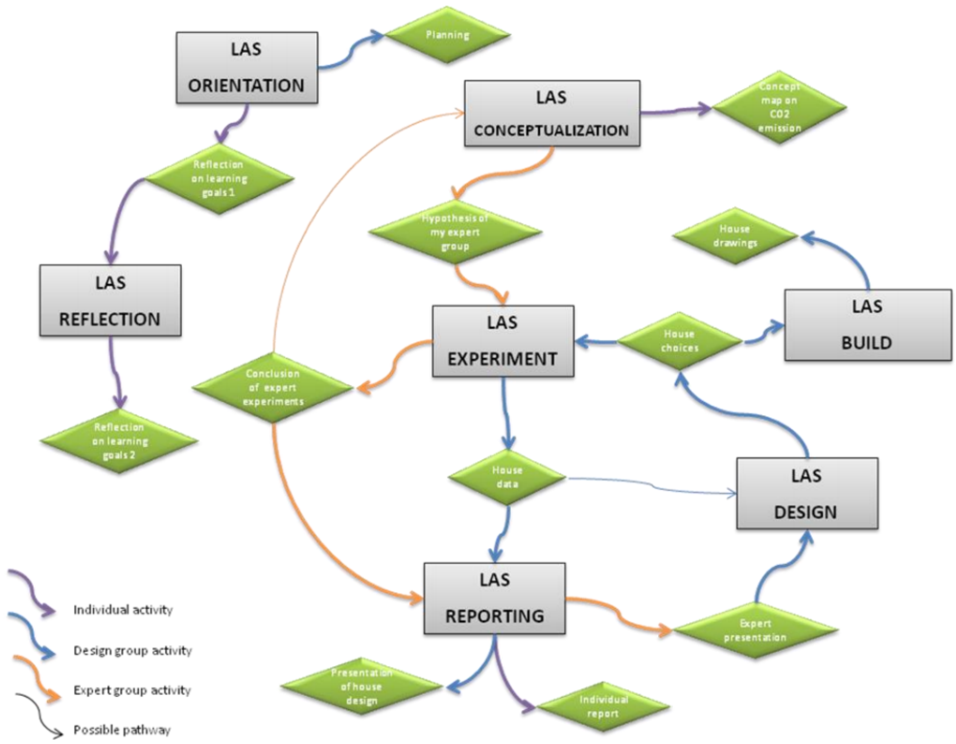


Figure 2.11: The learning design of the first mission in SCY (Geraedts et al. 2010).

institutes of the SCY consortium. During the project, 13 different LASs have been created as part of the pedagogical specification that can be grouped and composed to create a mission (de Jong et al. 2010). By the end of the project, 4 missions have been delivered officially. For accomplishing the first SCY mission, students have to spend around 20 hours on this mission, which shows the drawbacks that might prevent a wider curricular support and a large-scale implementation of inquiry-based science education. Apart from the difficulty of educational design, the experimentation was not supported through the system in a way that it can be extended easily by other laboratories than the ones created by SCY. A lack of interfacing and standardization prevents the extension through new online laboratories and simulations and makes this approach less generic.

2.2.2 ROLE

The technical platform for the Go-Lab learning environment is Graasp, developed by the REACT group at École Polytechnique Fédérale de Lausanne (EPFL). Graasp has its origins in the ROLE project:

> "Graasp is a ROLE Tool developed to enable Agile Knowledge Management in general and Personal Learning Environment creation in particular. It enforces the concept of contextual shared spaces with fine privacy settings and recommendation of relevant peers, resources and apps from the cloud for teamwork or self directed learning activities. It also enables the construction and exploitation of Web app bundles." [5]

This description highlights the nature of the platform, which focuses on flexibility and personalization. Each personal environment, a *space*, is a structured collection of apps, assets and actors. The space itself can be structured by adding sub-spaces. The apps are applications that support the specific context of the space, for example calculators or graphic formula widgets in scientific contexts. The available widgets were listed in the ROLE widget store and could be immediately integrated into the own (learning) spaces or added to widget bundles that could be distributed externally in a compound format. Figure 2.12 shows an example of the widget store. Although ROLE primarily focused the research area of personalized learning environments (PLEs), many use-cases were in scientific domains and integrated remote science laboratories through its advanced, flexible and pluggable widget architecture based on the Open Social standard. Therefore it was possible, to develop learning spaces for inquiry-based learning scenarios through the main platform of ROLE, the ROLE sandbox. An example of the ROLE sandbox can be seen in figure 2.13.

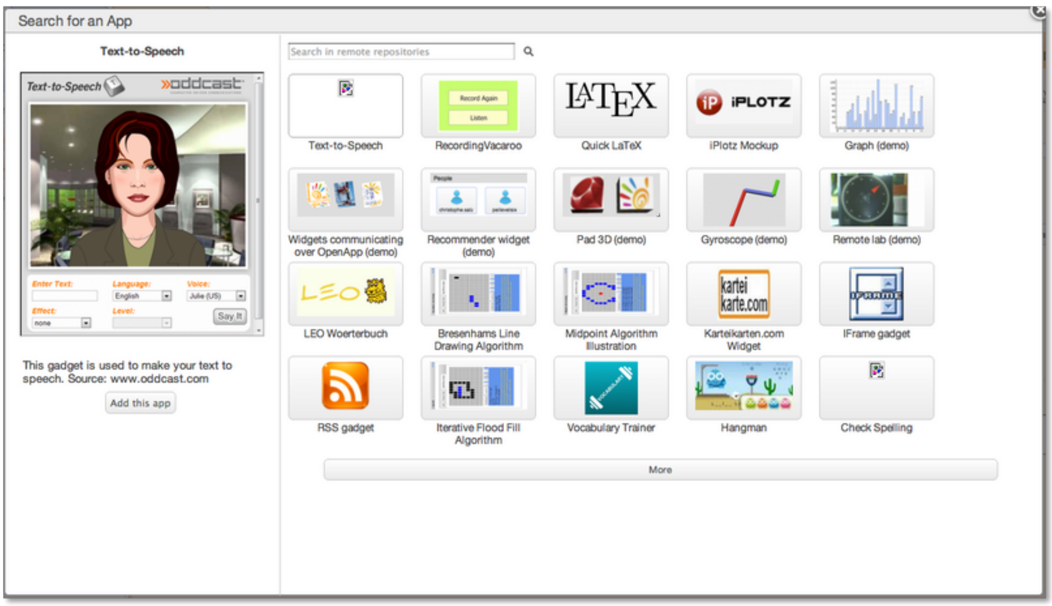


Figure 2.12: ROLE widget store.

[5] ROLE Consortium, Graasping the Basics of ROLE Graasp (2012-02-16), `http://role-project.archiv.zsi.at/index.html%3Fp=2413.html`, retrieved 2018-11-01.

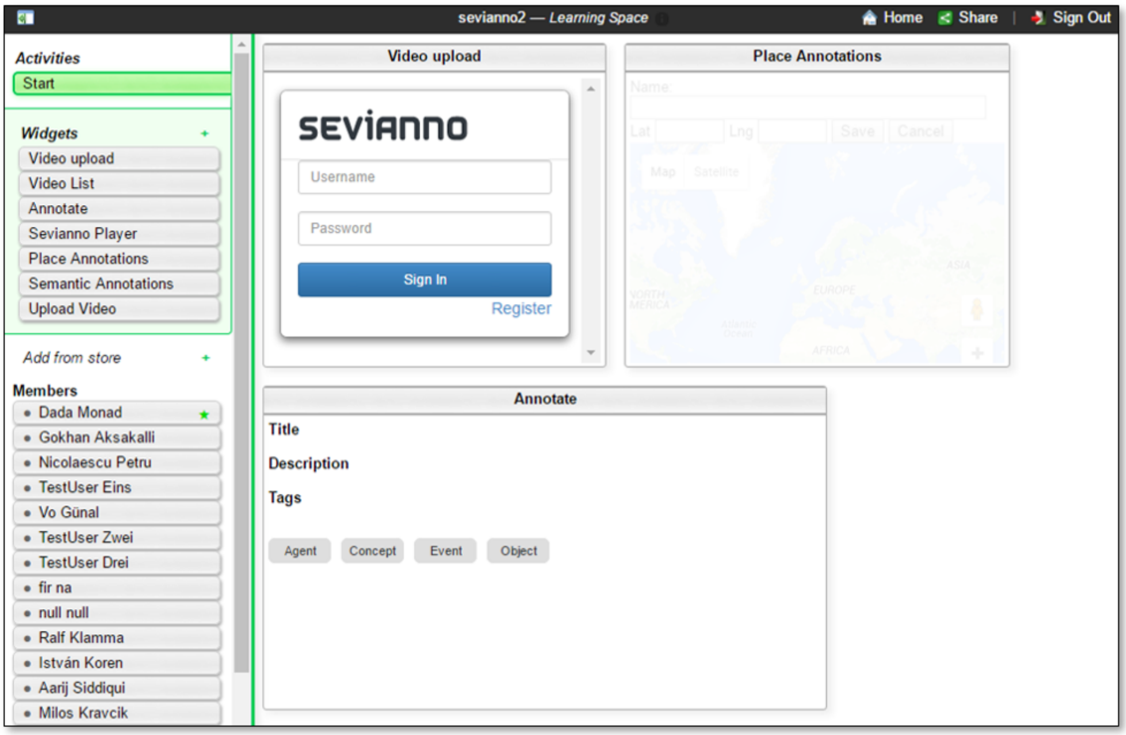


Figure 2.13: ROLE sandbox: the left bar contains a list of widgets and members of the space. The main part contains widgets that can be re-arranged on the canvas.

One of the relevant aspects of the ROLE sandbox was that it was highly customizable. The user could rearrange the space with the embedded widgets, plug in and add new widgets as they were necessary, but also share the environment with other users. This lifts the idea of a personalized learning space to a higher abstraction as the users produce collaborative (shared) spaces that can be organized in a flexible way in different use cases. Another purpose could be web conferencing. The widget store contained a plugin to a video conferencing tool and the whole space can be used to incorporate tools for the specific domain the conference is about, but also organizational tools as a collaborative notepad or a space to share files. Many research challenges throughout the ROLE project were on a software engineering level in the fields of personalized learning environments, pluggable architectures (widgets), software interfaces (open standards, such as OpenSocial), and mobile technologies. Also the field of contextual metadata for the tracking of user activities has been exploited throughout the project, resulting in the so-called *contextual attention metadata* format ("CAM"). This format was a relational logging format that integrates contextual information, for example about the learning environment.

2.2.3 Go-Lab

Go-Lab is a European project aiming at promoting IBL with online laboratories (labs) for STEM education at school (de Jong et al. 2014). During the project (2012 - 2016), Go-Lab implemented inquiry-based science education in more than 1000 schools all over Europe. The Go-Lab online learning environment is a single entry point to access online laboratories and to create inquiry learning spaces (ILS). ILS are rich open educational resources that can be collaboratively created in the Graasp social media platform, shared in the Golabz open repository, and exploited by the students either as standalone resources or embedded in open social or educational platforms (Rodríguez-Triana et al. 2014).

Go-Lab provides a single entry point to create and run learning spaces. Figure 2.14 shows the connection of the different subsystems of Go-Lab. First, the user browses the Go-Lab portal, which is an inventory of embeddable applications ("apps"), online laboratories ("labs") and inquiry learning spaces ("ILS"). The user, typically a teacher who wants to facilitate inquiry-based science education, identifies either a pre-made scenario ready-to-use in the form of an ILS, adapts it to his or her needs, or creates an own ILS from scratch. When the teacher browses the labs, s/he can select a lab according to the individual preference regarding the curriculum, target group or age. When selected, the Go-Lab portal has a button to create an ILS. The authoring process then takes place in another system, Graasp. The systems are seamlessly connected through and the entities to be shared described through RDF. The authoring platform allows the teacher to share the ILS with the learners in order to run the learning scenario. This sharing is in form of a URL that can be distributed through any established communication channel that the teacher wants to make use of.

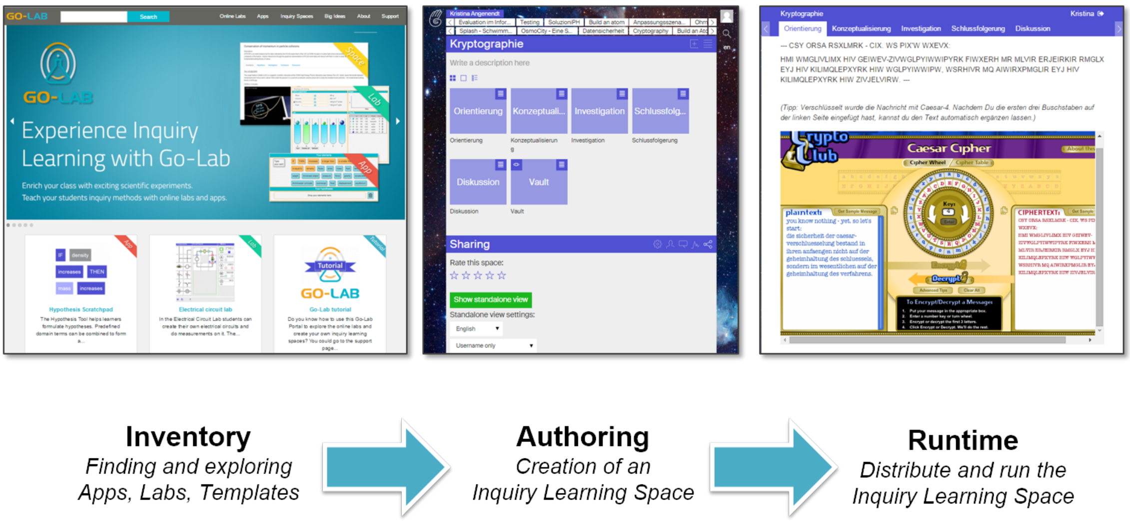
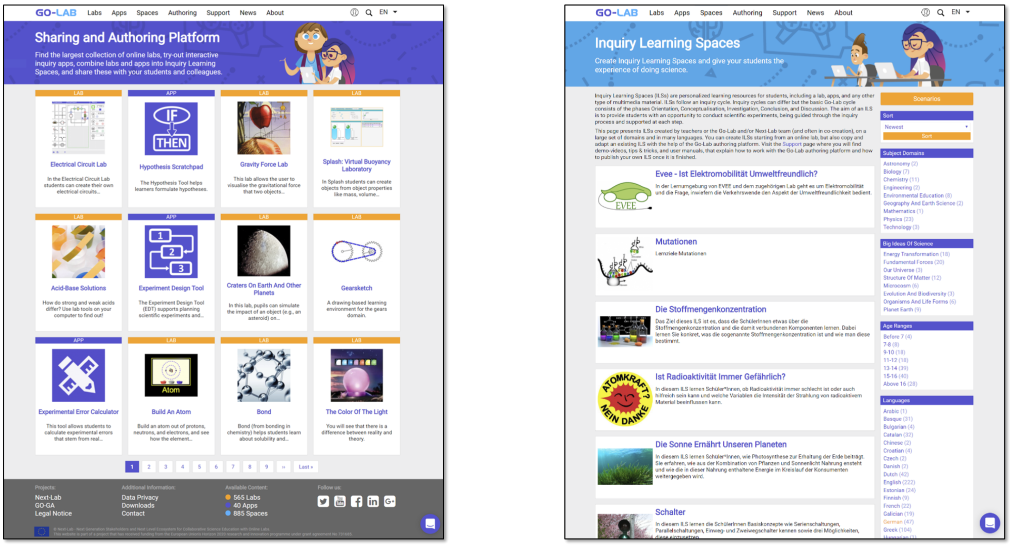


Figure 2.14: Go-Lab provides a single entry point to create and run learning spaces.



(a) Golabz landing page (b) German ILS published to Golabz

Figure 2.15: The Go-Lab Portal is an inventory for apps, labs and ILS.

The Go-Lab Portal

Golabz.eu, the Go-Lab portal, is the entry point for the creation of inquiry learning spaces. Technically, the platform is a content management system based on Drupal[6], which lists apps, labs and published ILS in its inventory. It is directly connected to the Go-Lab learning environment (Graasp, see section 2.2.3) in order to create own instances as copies from published ILS or to use apps and labs in a standalone mode. A faceted search based on a metadata scheme that has been developed through the projects enables filtering of the displayed items. Therefore it is possible, for example, to view all German ILS, or even to narrow that down to a certain subject, domain or target age range (see figure 2.15b).

By the end of October 2018, the inventory of Go-Lab listed 565 labs, 40 apps, and 885 inquiry learning spaces published to the platform. Besides the inventory, unofficial or external apps and labs can be embedded into the learning environment without being listed on Golabz. Also the number of the ILS used is much higher than the number of ILS published.

[6]The Drupal Association, Drupal Open Source CMS, `https://www.drupal.org/`, accessed: 2018-11-01.

The Go-Lab Learning Environment

The Go-Lab learning environment is conceptually and technically based on the ROLE sandbox, which has been described in the previous section 2.13. It is mainly organized as a development to create personal learning spaces, comparable to a folder or directory with arbitrary content. Each learning environment is structured as a *space*, which is in line with the Graasp / ROLE concept of spaces. However, the personalization features are on the level of the creator of the space, usually the teacher or instructor of the educational or learning scenario.

Graasp in the context of Go-Lab can be seen as a continuous development of the ROLE tool with respect to inquiry-based learning. Go-Lab adds a pedagogical middleware to the Graasp environment, which helps to implement and scale up inquiry-based science education with online laboratories. Figure 2.16 shows how this middleware bridges the gap between platforms and inquiry-based science education.

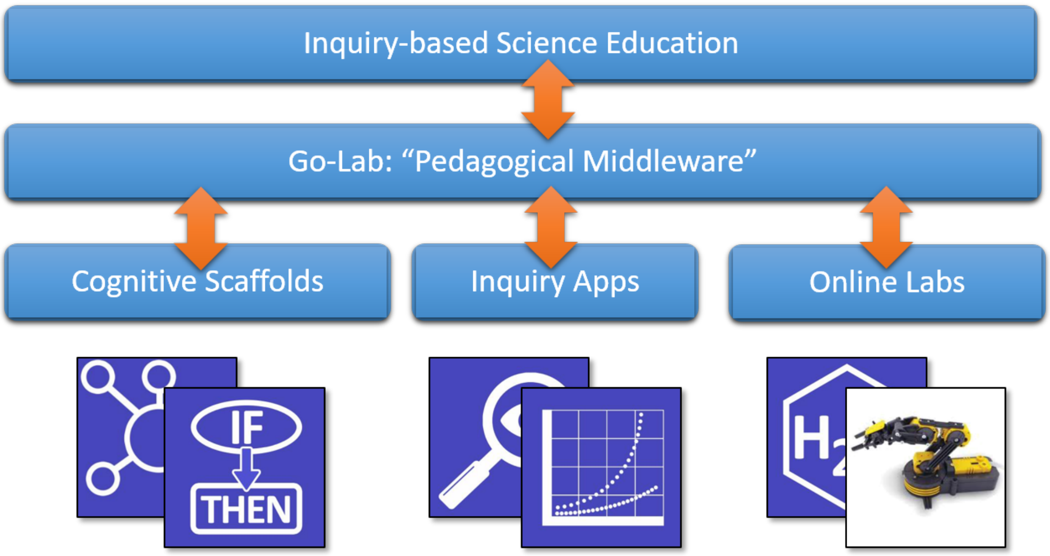


Figure 2.16: Go-Lab acts as a pedagogical middleware for IBSE.

The model for inquiry-based learning in Go-Lab is a synthesized model of inquiry cycles (see section 2.1.1) that has been flattened to subsequent inquiry phases. Although the literature identifies different sequences, models or names of phases to structure the inquiry learning process, the main idea is always to encourage the students to develop their own questioning, figure out their own responses by making proper hypotheses and designing proper experiments, and reflect on the observations. A recommended model within Go-Lab takes the following phases into consideration: Orientation, Conceptualization, Investigation, Conclusion, and Discussion (Pedaste et al. 2015).

Graasp provides teachers with a recommended process model and a collection of supporting tools that can be adapted according to the needs of the learning scenario. Still, during the learning activity, a student might or might not follow this structure. The sequence of phases followed by the student corresponds to a model that represents the learning process. The degrees of freedom given to teachers and students in the design and enactment of ILS make Go-Lab an interesting example for applying analytic methods of process discovery. One of the tasks in this field is to extract process models and find deviations from this model in concrete process instances (van der Aalst 2011). These deviations might be used as indicators to enforce process awareness and reflection on part of the learners, which is seen as beneficiary in inquiry learning (Garrison 2003) besides process-oriented guidance mechanisms (Zacharia et al. 2015).

Cognitive Scaffolds in Go-Lab

In Go-Lab, the degree to which scaffolding is added and faded is defined by the teacher and not restricted to or by the Go-Lab environment. The term of cognitive tools or scaffolds has been defined in a general way: "The basic idea of most cognitive tools is to boost the performance of learning processes by providing information about them, by providing templates, or by constraining the learner's interaction with the learning environment" (van Joolingen et al. 2007). Some of the apps in the Go-Lab inventory serve the purpose to support and to scaffold the (inquiry) learning process. When we refer to cognitive scaffolds, we point at concrete, embeddable apps to support the inquiry process, knowledge representation or construction. In the Go-Lab inventory, these apps are called "Inquiry Learning Apps" if they are connected to inquiry activities for learner support. The non-exhaustive list of inquiry learning apps in Go-Lab contains the following apps that are useful during the according inquiry phases:

Conceptualization: Concept Mapper, Hypothesis Scratchpad, GoModel (System Dynamics), Wiki Tool, Question Scratchpad

Investigation: Experimental Design Tool, Observation Tool, Table Tool, Experimental Error Calculator

Conclusion: Conclusion Tool, Report Tool, Wiki Tool, Data Viewer

Discussion: Question Scratchpad, Wiki Tool, Input Box

Not all of these apps directly represent knowledge. The Experimental Design Tool is highly specific and bound to the design of an experiment. It uses variables that are already defined through other scaffolds, for example the hypothesis scratchpad, and it provides the user with means to specify parameter variations. The Data Viewer is a tool intended to display data sets that have been collected during the experimentation. This is quite useful for learners to have an visualization of the experimental data in order to relate multiple representations (Bodemer and Dehler 2011).

Another category of cognitive tools to support learners in doing inquiry are monitoring tools to help students keep track of their learning process or experiments (de Jong 2006). In Go-Lab, such tools for process awareness exist (Manske et al. 2015a), but are not highlighted in this section, as they do not directly influence or rely on the concept of knowledge diversity, but possibly complement the effects regarding group awareness (cf. section 7).

The following apps are highlighted in this section as they are tools to express learners knowledge (cf. section 2.1.2) and were used for the approach of this thesis (see section 4). The Go-Lab *concept mapper* is typically used during the conceptualization phase as an externalization of learners' knowledge. Concepts and relation names can be predefined by the teacher, but are not restricted to it. In an optimal case, they encode key concepts and their respective relations, but we saw in our experiments, that sometimes learner use more natural language and whole sentences for relations or concepts. This makes the identification of knowledge more difficult and shifts the role of this scaffold to the direction of a note pad. The *hypothesis scratchpad* can be used by learners in order to create hypotheses using a drag and drop approach with prescribed or free text blocks. Predefined quantifiers and operators help to prestructure hypotheses. However, we could observe that learners use free text blocks, similar to the concept mapping, to write full sentences in the hypothesis editor. Therefore, the identification of key concepts is quite similar to the extraction from learner-generated texts. In our Go-Lab scenarios, we used two tools to let learners write texts: an input box, which is a very minimalistic plain text input field, and a wiki tool. The input box implicates learners to write short answers or texts. The wiki tool (see figure 2.6) has a document-oriented user interface and wiki markup to format and connect texts. In contrast to a plain text tool, this implies and motivates to write more elaborated texts, for example reports, conclusions or reflections. The use of this tool can vary depending on the use case. Go-Lab supports, similar to other apps, different modalities to embed the tool, for instance in a specific inquiry phase or as a general tool in the "tools" panel (bottom bar in an ILS).

2.3 Learning Analytics

The interdisciplinary field of the learning sciences exposed a rich theoretical underpinning and understanding of learning. Many theories have been advanced to frame conditions for effective learning, for example how effective learning design impacts learning. Apart from this area, learning analytics (LA) emerged from different disciplines as an educational research field and set up a new paradigm in this area. In contrast to the learning sciences, the roots in artificial intelligence, statistical analysis and business intelligence foreshadow data-driven approaches and narratives to explain aspects of learning. Often, LA stands for the usage of computational (analysis) methods on learning data to inform different stakeholders with the aim to improve learning processes and environments (Ferguson 2012). Three types of computational methods used in LA are distinguished to categorize analytical approaches and objects of interests (Hoppe 2017):

1. *Product analysis* is focused on the products of learning, particularly learner-generated content. The analysis of content can be performed using text mining or other techniques of artifact analysis.

2. *Process analysis* is based on action sequences and traces that help to reenact the (whole) learning situation. The analysis of processes can be performed using methods of sequence analysis, for example sequential pattern mining or sequence alignment.

3. *Social network analysis* (SNA) relies on social structures, for example interactions between learners or learners and artifacts in collaborative learning scenarios. These structures are typically represented as graphs, where SNA employs methods of graph theory for the analysis.

Figure 2.17 presents an overview of the three types and examples of the respective methods for each category. In the context of this dissertation, which is situated in the Go-Lab project, network analysis has minor relevance due to its design principles and the lack of explicit (social) relations. The analysis of learning Processes following a model of learning phases in IBL has been applied using methods of sequential pattern mining (Manske et al. 2015a). Content-based analyses (1) have so far received less explicit attention from an LA perspective. Although we employ some techniques based on network analysis, they primarily target the content of certain objects. For example, concept maps have a structure that is comparable to a network of concepts, which makes it possible to employ network measurements to assess concept map quality. Additionally, we used network text analysis to create a text network from a learner-generated text in order to identify key concepts.

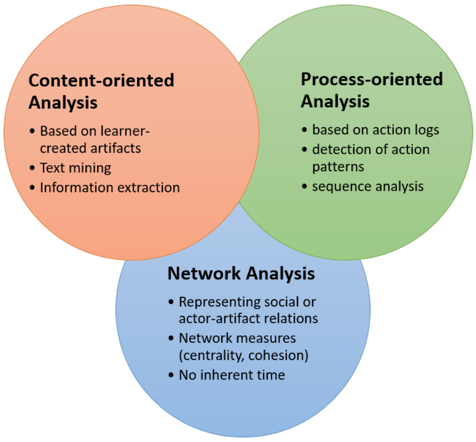


Figure 2.17: The trinity of LA according to Hoppe (2017).

2.3.1 Process Analysis

Although process modeling has its origin in the field of business processes, it has been also applied to learning contexts. For instance, learning processes can be modeled and structured through learning design and process mining, similar to budiness process models (Bergenthum et al. 2012; Miao and Hoppe 2011). Particularly for constructivist learning approaches, the role of learning design and the specification of learning flows have been discussed (Lejeune et al. 2009a; Harrer et al. 2007).

Modeling learning processes as a basis of learning design can be applied to different learning approaches including inquiry-based learning. IBL can be a successful pedagogical approach, if effective support is offered to the students at various levels (de Jong et al. 2013): first, activities are structured in successive phases; second, in each phase scaffolding tools supporting the activities are provided; and third, relevant cues are given to the students when necessary.

2.3.2 Product Analysis

The analysis of products with semantic technologies have so far received less explicit attention from an LA perspective, although computational linguistics techniques have been successfully applied to the analysis of collaborative learning processes (Rosé et al. 2008). The OpenEssayist-System (Whitelock et al. 2013, 2015) uses linguistic approaches to analyze text-based artifacts generated by learners. While the outcome of the system aims to adapt learners to write essays with a specific structure, it under-represents epistemic aspects of learning analytics. Students are forced to think about why the system produces a certain outcome, but it does not provide any help by delivering explanatory models. To support the revision or the evaluation of students' learning outcomes effectively, learning analytics applications might need to process learner-generated content automatically. The processing of text-based representations, e.g., wiki articles, highlights the importance of semantic technologies for LA. Several systems use semantic technologies to represent knowledge. The Xerox Incremental Parser (XIP) Dashboard uses approaches of Natural Language Processing to aggregate salient sentences of scholarly papers, providing a variety of analytics to the students (Simsek et al. 2014). The Debate Dashboard (Quinto et al. 2010) focused on distributed knowledge management. A central aspect is an argument mapping tool with a visual component to support collaborative work and collective sense-making. To effectively provide scaffolds for the students' interaction with learner-generated content, the design of LA interventions as proposed by Wise (2014) can serve as a point of reference. An LA intervention can be defined as a frame in which analytic tools, data, and reports are gathered and used. Wise formulates four design principles for the successful integration of LA tools: The smooth integration of LA results into the learning environment; the agency in interpreting and responding to analytics in terms of supporting and not detracting learners; setting up a frame to give learners a comparison point when interpreting results from analytics; and the chance to discuss and negotiate the LA results in a dialogue. These statements have been extended by Harrer and Göhnert (2015) adding the two principles scope, focusing on contextually relevant information for the learner, and representational consistency, i.e. adapting the interface to the learning environment. Apart from the question of what data to visualize, how to support pedagogical interventions and present them to learners, the nature of knowledge and its implications to learning is relevant. Epistemic aspects are framing research in learning analytics (Knight et al. 2013). Dimensions of epistemic beliefs highlight the learner's view on what is to be learned, particularly classifying knowledge on part of the students (Mason et al. 2010, 2011).

The design of the concept cloud was influenced by Wise's framework for learning analytics interventions and incorporates Mason's epistemic beliefs and aspects to foster reflective behavior on part of the learner and connect the visual items to students' knowledge. Similar to the OpenEssayist system, it uses a tag-cloud-like visualization to display key concepts used in the learning spaces, particularly the learner-generated content. Such representations support the user in monitoring a large number of items and thus provide a medium the learner can manipulate and interact with (Qiu et al. 2007), reducing navigational paths (Rivadeneira et al. 2007).

2.3.3 Learner and Knowledge Models

According to Bull and Kay (2010), "open learner models are learner models that can be viewed or accessed in some way by the learner, or by other users (e.g. teachers, peers, parents)." The user can investigate the learner model and particularly try to understand why the system inferred the model. As a consequence, the user can intervene on certain things. This process is seen to improve metacognitive skills on part of the learners (Bull and Kay 2013). However, such models usually lack in a certain amount of transparency on how the system builds a model. This can lead to a disagreement between user and system, which cannot be resolved.

In the field of (open) learner modeling, the dimensions and aspects of what to analyze and model can differ, and different aspects can be taken into account. For successful scientific inquiry, knowledge has been emphasized as one of the key ingredients. Therefore, we focus on the field of learner modeling with respect to open learner models that encompass learners' knowledge. In the field of intelligent tutoring systems, open learner models have shown positive effects regarding metacognitive skills and learners' self-regulation (Bull et al. 2010; Mitrovic and Martin 2007; Long and Aleven 2013). Mitrovic and Martin (2007) conclude:

> "Students appreciated having access to their models, and they felt this feature contributed to their understanding of the domain. Performance of less able students becomes significantly better than that of their peers of similar abilities without access to their models."

Hoppe and Ploetzner (1999) identified cognitive models of collaborative learning as important to the research in the field of CSCL. The use of such analytical models can provide intelligent run-time support for collaborative learning, for example in the field of automatic group formation and problem solving. They used an open framework for multiple student modeling, which contains the three different modeling methods "overlay", "bug library", and "perturbation method" (Hoppe 1995). These three models can be analyzed and integrated. Overlay models are simple and optimistic additive models. In the case of an overlay model, knowledge can be integrated as the union of individual portions of knowledge. A student model of type overlay is then the subset of all knowledge items. In their modeling approach, Hoppe and Ploetzner (1999) declared a knowledge item as a certain topic a student knows, or a skill the student dominates. For the work of this thesis, we highlight that the integration does not necessarily take individual portions into account, but also works for more heterogeneous settings, for instance, integrating for a single learner knowledge from different artifacts. This is the case for Go-Lab scenarios, where learners usually have different knowledge externalizations or learner-generated artifacts such as texts, concept maps and hypotheses. A buggy model is a model in which rules for errors are encoded. Particularly when dealing with potentially wrong knowledge items, things that are "mistaken" or do not relate to a correct understanding, such models are necessary.

Hoppe and Ploetzner (1999) identified knowledge complementarity as a useful base line for knowledge integration and collaborative scenarios, where learners benefit from each others knowledge. The definition of knowledge complementarity has been emphasized as useful to trigger cooperative situations with the following notion based on the predicates *can_help*, *knows* and *has_difficulties*:

$$can\_help(S_1, S_2, T) \text{ IF } knows(S_1, T) \;\&\; has\_difficulties(S_2, T).$$

In this definition, a relatively clear definition of knowledge complementarity within an overlay model is given. When student S_1 knows about a topic T, and student S_2 does not, they are complementary regarding topic T. In this cooperative setting, this means that student S_1 can help S_2. Following this definition, we can add that S_1 and S_2 are *diverse* regarding T. However, in this simple model, the complementarity does not lead to a quantifiable measure for groups, particularly for group diversity.

Openness, besides the idea of making the model accessible to learners and teachers, has another important connotation. In tutoring systems, learner models are often defined as deviations from a predefined path or in accordance to erroneous behavior to match buggy rules. In *model tracing* approaches by Anderson (1984) the student is basically forced to stick with the model the system accounted the learner to. Such modeling approaches are limiting the set of scenarios. The context of this thesis is situated in the Go-Lab project, which enables open and flexible inquiry-based learning scenarios. We do not want to restrict the knowledge models with reference models or reference frames that need a high level of specification. Therefore, we need open models that can be defined additively, but also optimistically without the need of overspecification. Open overlay models satisfy this condition and will play a role in the knowledge management and integration approach of this work (see section 4).

2.3.4 Architectures for Learning Analytics

Current architectures for learning analytics software systems are being developed in different contexts. This incorporates business analytics and also data mining tools (Kraan et al. 2013). Most of the tools are designed for specific types of learning environments like learning management systems (LMS). LMS platforms such as Blackboard[7] and Desire to Learn[8] offer their own bundled learning analytics software solutions, which are dedicated to the end user exclusively and hence not extendable. Fortenbacher et al. (2013) developed the LeMo tool ("Learning process Monitoring") which is capable of descriptive analysis of resource usage and student activity as well as more complex analysis like the identification of frequent learning traces. This tool offers several connectors to learning management systems from different vendors. However, the connector fetches a snapshot from the target LMS and converts the data

[7] Blackboard Inc., Blackboard LMS: `http://www.blackboard.com`, accessed: 2018-04-02.
[8] D2L Corporation, Desire to Learn LMS: `https://www.d2l.com`, accessed: 2018-04-02.

set into the LeMo data model. This leads to a static analysis, which makes it difficult to place direct interventions. To achieve a near-realtime analysis of the learner data requires a direct forwarding of new learner data, but also a feedback channel to integrate and place the interventions in the learning environment.

PSLC datashop (Koedinger et al. 2010) is a more research oriented platform that enables sharing of large learning data sets. Although the focus is on effective data management, it also offers some analysis and visualization tools. Another platform dedicated to analysts is the CRUNCH infrastructure . It offers an analytics workspace to create analyses and reports based on R scripts[9]. Scripts can be run as cron jobs or released as public web services. Particularly the web services are intended to be reused by other users of CRUNCH. Tools like the PSLC datashop and CRUNCH are more focused on the development and reuse of analytics services and data. They can be used to develop and test analytics services very well, but do not provide direct feedback mechanisms for teachers or students on their own.

More emphasis on analytics systems for intelligent user feedback comes naturally from intelligent tutoring systems research (ITS). Throughout the MiGen project a layered architecture for intelligent feedback is presented (Gutierrez-Santos et al. 2010). Feedback is produced when activity data flows through an analysis layer, where several components analyze different aspects of the learner behavior. A dedicated aggregation layer combines the analysis results to a learner model, whereas a feedback layer presents personalized scaffolds to the learner. All the mentioned systems serve different aspects of learning analytics. The challenge is to fulfill requirements of learning analytics architectures and to integrate different approaches into one open and extendable infrastructure in order to prevent fragmentation.

The Open Learning Analytics project (Siemens et al. 2011) advocates modular systems that allow openness of process, algorithms, and technologies which is an important feature in a heterogeneous field as learning analytics. This should also be the line followed by the analytics architecture in Go-Lab presented in this paper. Two existing learning analytics infrastructures that also go into this direction are the analytics services of the Metafora platform (Harrer 2013) and the ROLE sandbox (Renzel and Klamma 2013). The Metafora platform is a web-based multi-tool environment for complex learning activities in small groups. It uses heterogeneous and decentralized components for action logging, analysis of group behavior across the usage of multiple tools and user feedback. The ROLE sandbox is a platform for Personalized Learning Environments (PLEs). Its analytics system uses widely accepted protocols and standards for action log data and web services in order to achieve interoperability of data sets and services. This system implements a pipeline based processing of action logs in which it is also possible to enrich action logs with context information

[9] The R Foundation, The R Project for Statistical Computing, `https://www.r-project.org/`, accessed: 2019-06-02.

and metadata. The ROLE interoperability framework has been used to track activities through the communication between embedded OpenSocial widgets. A dedicated analytics widget fetches the learner data from other widgets, enriches contextual metadata ("CAM" format) and exports them into the activity database (Govaerts et al. 2011). However, supporting and facilitating client-side activity tracking harbors several risks. First, enabling applications to listen to activity data without a control mechanism can be seen as a privacy concern. Second, with an open repository of widgets, where external developers can add their own applications, there is a need for a rich interface and well defined protocol for the description of activities in a way that analytics applications can make use of it.

2.3.5 Integration of Learning Analytics Applications

The definition of learning analytics by Ferguson (2012) highlights the origin and closeness to business intelligence systems. Business intelligence systems are often characterized through the clear separation of the end user system and the analytics platform. In our case, the end user systems are the connected learning portals, e.g. the Go-Lab portal, while the analytics platform contains the workbench for the creation of workflows and their visual gadget representation. The syndication of different gadgets in a single view or web page corresponds to the idea of analytics dashboards to enable better informed decisions.

This idea of embedding learning analytics gadgets in the context of online-learning has been shown in Malzahn et al. (2013). There, two courses in vocational training in a learning system based on the Liferay portal have been compared. The output produced through the analytical methods was represented as a gadget in the portal. In comparison, the novelty in our approach is based on the flexibility in the choice of different target platforms supporting connectivity of standardized gadgets, e.g. following the OpenSocial specification. Since the rise of integrated personal learning environments, the syndication of social media and learning management systems becomes more important. Gadget platforms like OpenSocial, which plug into social networking platforms, are well suited for this purpose. Examples of collaborative learning systems using OpenSocial are ROLE (Govaerts et al. 2011) and Graasp (Bogdanov et al. 2012). Following the approaches of ROLE, Graasp as the platform for Go-Lab, also uses OpenSocial to provide a pluggable application architecture (Govaerts et al. 2013a; Gillet et al. 2013a). The case study of the learning analytics architecture for Go-Lab is presented in section 3.

In more recent research by the group of Ulrik Schroeder at RWTH Aachen, a rich architecture and an underlying model for learning analytics has been created (Chatti et al. 2012). The Open Learning Analytics Platform was intended to integrate indicators for learning analytics within a target learning management system (Chatti et al.

2017), which enables teachers or instructors of online learning to use analytics for their respective learning environments. The learning analytics platform is specifically designed to interface the learning management system, with an own relational action logging format comparable to the CAM-format.

2.4 Computer Supported Collaborative Learning

Collaborative learning can be seen as "a *situation* in which *two or more* people *learn* or attempt to learn something *together*" (Dillenbourg 1999). Although this definition seems to be broad and general, it lines out some characteristics of the field. However, there is no consensus about a precise definition (Dillenbourg 1999). With the rise of digital technologies, the field of Technology Enhanced Learning (TEL) investigated the impact of technology on learning. This led to new challenges for learning and teaching within the integration of computer-support, but also brought out new opportunities and novel applications. Computer-supported Collaborative learning (CSCL), an emerging branch of learning sciences and computer-supported cooperative work (CSCW) has drawn a lot of attention in research since the 1990s (Stahl et al. 2006). As one of the important land marks in the research, the CSCL conference series has been started in 1995. The research in this area focused on sharing and construction of knowledge through social interaction and processes within a technology-supported learning environment, mainly underpinned by theories of constructivist epistemology and social cognitivism (Resta and Laferrière 2007). "For many educators and researchers, CSCL appears to be one of the most promising ways, not only to promote, but also to achieve desired changes in teaching and learning practices" (Lipponen 1999).

2.4.1 Artifacts in CSCL

As a highly interdisciplinary and emerging field, the community of CSCL proclaims a new paradigm of research on instructional technology, which is in different from earlier approaches (Koschmann 1996). However, as CSCL is based on learning sciences, the field has a particular view on learning as a research subject. The framing conditions of learning, when and how learning takes place, have been discussed throughout the community. The assumptions underlying this research are usually in contrast to traditional learning theories that observe learning as something that "takes place inside the learner and only inside the learner" (Simon 2001). Furthermore, Simon (2001) points out two facets that define a baseline for learning: (a) the learning activity ("active learning period") and (b) the role of the learning environment. The latter relies on the assumption that certain aspects of an environment (in a broad sense) can enhance the abilities and willingness to learn actively. Throughout both facets, different kinds of artifacts incur and exist in the process of learning. This view of CSCL as an artifact-mediated research discipline has been elaborated by Stahl (2002) as a new paradigm of learning research. Koschmann (2002) is often cited in this context:

> "CSCL is a field of study centrally concerned with meaning and the practices of meaning-making in the context of joint activity, and the ways in which these practices are mediated through designed artifacts."

In the research of CSCL, (physical and virtual) artifacts play a central role (Stahl et al. 2014). In his original work about a theoretical framework for CSCL, Stahl (2002) provided a very broad definition of the term *artifact*: "An artifact is a meaningful object created by people for specific uses". In addition to the aforementioned notion of artifacts as learner-generated objects, the term "artifacts" consists of the following specifications in the context of CSCL according to Stahl et al. (2014); Overdijk et al. (2012); Dimitriadis (2012); Ludvigsen et al. (2015); Kienle and Wessner (2006); Suthers (2006):

Technical artifacts are technological components or parts of a technical infrastructure that enables or supports the collaboration. Examples are web applications that support the communication, for example (digital) chat tools that are available in the learning environment. These artifacts are sometimes connected to agents in order to support the inherent communication or the learning by placing interventions or actions throughout the agent-artifact-connection(Overdijk et al. 2012).

Intersubjective artifacts Intersubjectivity is mainly characterized as interactions at the interpersonal level, where cognitive processes and activities may be distributed across members of a social group (Hollan et al. 2000; Suthers 2006). This attributes the interaction between individuals, or between an individual and information objects in the context of learning. According to Koschmann et al. (2005), this involves messages which have been produced in computer-mediated exchanges. Therefore, intersubjective artifacts mediate communication. Suthers (2006) states out that "intersubjective learning further specifies that the process of meaning making is itself constituted of social interactions."

Instructional artifacts According to Stahl et al. (2014), instructional artifacts present "domain topics, lessons, guidance, scaffolding or scripting". They motivate and direct the collaboration. In this sense, they organize work (Sutter 2002) and provide a frame for the learning activity. While providing direct instruction is a common practice in tradition teaching, the field of CSCL aims to make collaboration more effective. Dillenbourg (2002) states out that "free collaboration does not systematically produce learning." In order to improve collaboration processes, scripting emerged as a common approach to mediate instructions and artifacts in CSCL (Fischer et al. 2006). CSCL scripting is based on the concept of social cooperation from educational psychology, which fosters, according to Weinberger et al. (2005), "the processes of collaborative knowledge construction as well as learning outcomes." Dillenbourg and Jermann (2007) define a script in the context of CSCL:

> "A script describes the way students have to collaborate: task distribution or roles, turn taking rules, work phases, deliverables, etc. This contract may be conveyed through initial instructions or encompassed in the learning environment."

The development and facilitation of scripts through run-time and orchestration of collaboration have been investigated by many researchers (Weinberger et al. 2005; Kobbe et al. 2007; Harrer and Malzahn 2006; Tchounikine 2008; Dillenbourg et al. 2009a; Dimitriadis 2012). Efforts to adapt scripting languages from other fields such as IMS LD to collaborative scenarios led to new scripting languages and respective software tools to develop the script instances (Miao et al. 2005). However, in the field of CSCL, scripting posed a debate about the complexity and risks of facilitating and orchestrating such instructional artifacts in regular teaching practices (Dillenbourg 2002). For example, scripting raises the risk to disturb natural interactions or problem solving processes of learners. In addition to the complexity of the interaction itself, this boils down to an important aspect in this context: for a pedagogical setting, in which technology-supported learning takes place, the presentation of instructions plays a crucial role. An explicit scripting is less important than an environment to mediate communication and instruction. The concept of instructional artifacts also appears in the context of instructional technologies, particularly in the field of research on mathematics education (Meira 1998; Evans and Wilkins 2011). Instructional devices play an important role in mathematics education and pose an activity-oriented and a knowledge-oriented view on instructional artifacts, where tool mediation according to Vygotsky's theory of cognitive development plays a central role (Dixon-Krauss 1996). Although we created a distinction to technical artifacts, the facilitation and embedment into a pedagogical setting through instructions cannot be completely decoupled from technical artifacts. However, modern approaches of scripting CSCL contain an explication of the tool mediation and explicit transitions between artifacts through the scripting language (Dillenbourg 2015). In contrast to this, Go-Lab does not use any explicit scripting of collaboration that is facilitated or mediated through technical artifacts (de Jong et al. 2014).

Learning artifacts are the products of learning, the interaction between learners and the learning environment, including intermediate and final representations of these products. While some traditional definitions see learning artifacts as physical objects in order to support the learning of others (Sherin et al. 2004), in the context of technology-enhanced learning this definition can be extended to (technological) tools. In the context of IBL environments such as SCY-Lab or Go-Lab (compare section 2.2), each learning activity is bound to a certain learning object. In this case, the artifacts of learning are the products of scaffolds or (inquiry) apps, and therefore the output of the technical artifacts. In line with the research in the field of IBL we do not limitate learning artifacts to technological or physical artifacts, but also include final and intermediate products of learning, for example scientific arguments (Bell and Linn 2000), experimental data, hypotheses, or documents (Lejeune et al. 2009b). In this sense, learning artifacts are mainly learner-generated, but could be also prepared by teachers or through an instructional design (Lejeune et al. 2009a). An example of a prepared learning artifact in Go-Lab is an experimental design (created through an *experimental design tool*), where the variables are already entered by the teacher. In this example, the ILS already contains a learning artifact (experimental design), which is an intermediate, but not the final product.

Knowledge artifacts While Dillenbourg (1999) argues, that learning is often seen as a side-product of problem-solving, a lot of emphasis has been put in the facet of knowledge construction, particularly framed by group processes (Resta and Laferrière 2007; Stahl et al. 2006). Externalizations of knowledge, such as concept maps, can be seen both as knowledge artifacts and learning artifacts, if they are explicit products of the learning. This is the case for tasks that facilitate such a knowledge artifact construction, which is prototypical for some inquiry-based learning designs (see section 2.2.3).

In summary, CSCL scenarios can be seen as a composition and orchestration of these artifacts, which structure, facilitate and mediate communication and learning. Research in CSCL is investigating the conditions and environments in which collaboration takes place. The definitions above imply that the dependency on interactions in a social and observable context demands a research paradigm that is not only restricted to a quantitative research in controlled (laboratory) conditions. It is necessary to have qualitative and mixed studies of learning practices in order to "explore the understanding of the participants in collaborative learning" (Stahl 2002). To create a rich understanding of the activities and processes, this might involve a variety of evaluation methods and data collection approaches that enhances traditional empirical research methodology. According to Wise and Schwarz (2017), this microgenesis, where fine-grained observations of interactions during the collaborative processes in the moment of meaning-making is specific and now classical to the research field of CSCL. Such methods and research paradigms also intersect with the newer discipline of learning analytics (see section 2.3).

2.4.2 Group Formation

Group formation is a key aspect of CSCL because it can affect the way people work together towards a common goal and eventually the learning outcome itself. Collaborative activities are expected to promote learning through common knowledge building and the social interaction among users (Stahl et al. 2006). However, collaboration alone does not ensure knowledge gain or successful practice (Jermann et al. 2001). Usually the task of group formation is carried out by the teacher who uses his experience on pre-defined criteria that may refer to students' social skills, gender, motivation or knowledge background (Ounnas et al. 2009). This complicated process requires time and does not always lead to success.

Based on the availability of student performance data in computerized learning environments, (semi-) automatic or algorithmic approaches to group formation have been suggested. For example, Balmaceda et al. (2014), define group formation as a weighted constraint satisfaction problem (WCSP) depending on the characteristics of students such as personality traits, team roles, and social relationships. Also network analysis techniques have been employed for analyzing the interaction of users through a learning platform and clustering students based on their similarity (Sadeghi and Kardan 2014). As one of the most sophisticated technical solutions so far, the GroupAL algorithm (Konert et al. 2014) allows for optimizing group composition according to a variety of features, with the option of choosing between homogeneity and heterogeneity for each of these features. The "MoodlePeers" plugin is an adaptation of the GroupAL algorithm to the Moodle LMS (Konert et al. 2016). It integrates the assessment of personal traits, attitude and skills into Moodle, which is then used to calculate an optimal grouping. In order to detect "soft skills" such as the attitude towards learning, the learners have to fill out a questionnaire first, which is implemented as a part of the Moodle plugin.

The role of group homogeneity in collaborative classroom activities has been a subject of various studies. There are indications that heterogeneity of knowledge is beneficial for group performance (Webb et al. 2002; Kizilcec 2013). However a certain baseline of background knowledge appears to be required for the collaboration to be beneficial (Gijlers and de Jong 2005). In prior work it could be observed that the positive effects of diversity had an impact on the performance of learning groups (Chounta et al. 2014). However, it is crucial to define goals for the algorithm of an automatic group formation in order to evaluate the quality of the output. For instance, such goals can be an optimization of the differences among groups in order to guarantee a fairness of the algorithm (Konert et al. 2014), certain conditions to create a specific type of learning scenario, or a desired pedagogical goal (Hoppe and Ploetzner 1999; Mujkanovic et al. 2012).

Former automatic or algorithmic approaches to group formation used skill- or score-based diversity on a variety of performance characteristics (Manske et al. 2015c; Konert et al. 2014). It is debatable, whether the score-based approaches can be used to form heterogeneous groups in terms of knowledge diversity. The Concept Cloud has presented both a technical and pedagogical approach how to incorporate semantic analyses with inquiry-based learning in Go-Lab (Manske and Hoppe 2016). Although the (automatic) formation of learning groups is in the interest of CSCL, less research has been put into knowledge-based groupings. Apart from skills and criteria, former approaches utilized analytic models following learners' knowledge complementarity (Hoppe and Ploetzner 1999). In this sense, computational methods have been used in order to create learners scenarios that benefit from knowledge diversity.

2.4.3 Cognitive Group Awareness

In the research of CSCL, fostering and facilitating different kinds of awareness has always been of interest. The key aspects of group awareness are, according to Bodemer and Dehler (2011) "the knowledge and perception of behavioral, cognitive, and social context information on a group or its members." They further state out that this field encompasses "the development of tools that implicitly guide learners' behavior, communication, and reflection by the presentation of information on a learning partner or a group". A framework by (Fransen et al. 2011) includes behavioral, social and cognitive aspects of group awareness. Cognitive group awareness is proposed as a "suitable means in order to support learners in using their cognitive capacities for meaningful individual and collaborative learning activities" (Bodemer 2011).

Cognitive group awareness tools (CGATs) provide learners with cognitive information on others that is usually not directly observable and that suggests performing specific behaviors (Janssen and Bodemer 2013). Traditionally, research in group awareness was focused on distant, computer-mediated communication situations. However, the benefits of group awareness are not limited to computer-mediated communication. Furthermore, it can enhance situations in which information is exposed that is not even visible in face-to-face situations. Group awareness tools typically target cognitive and social variables. For instance, these tools have the following functions: (i) highlight or list significant aspects of learning materials and resources to organize co-learners' communication, (ii) provide cognitive information on learning partners to facilitate grounding and partner modeling processes (Dillenbourg and Betrancourt 2006; Bodemer and Scholvien 2014). While the first function supports information filtering and coordination, the second one enables co-learners to easily compare said cognitive information. This draws their attention to specific constellations such as conflicting opinions or knowledge differences, which can be enhance the planning of tasks or initiate negotiation of a shared understanding. Particularly for visualizing knowledge differences, it was shown that the visualization of knowledge distributions in such tools can lead to a significant improvement of learning processes (Erkens et al. 2016a). One reason is that the visualization of knowledge distributions causes cognitive regulation as co-learners adapt their questioning behavior in help seeking based on the visualized levels of knowledge. They prefer to ask questions on topics with own missing knowledge visualized, and also take into account that their learning partner is knowledgeable on the topic, if this information is given (Dehler et al. 2011; Erkens et al. 2016b; Erkens and Bodemer 2017).

Another function of CGATs is that visualized knowledge distributions cause cognitive elaboration. Making learners aware of deficits in learning partners' knowledge makes them not only give more or longer explanations (Dehler et al. 2011; Erkens and Bodemer 2017), but also more elaborated explanations (Dehler Zufferey et al. 2010; Erkens and Bodemer 2018). Taken together with the findings of complementary group formation, we can conclude that particularly co-learners in groups with complementary knowledge distribution can profit from the additional visualization of this distribution. First, knowledge acquisition might be optimized, since the visualization allows learners to regulate the requests of explanations in terms of asking targeted questions on missing knowledge and receiving explanations on it. Second, cognitive elaboration might be better, since learners explain more elaborated, if they are aware of their learning partners' knowledge gaps. Third, they might better prioritize topics to be discussed and thereby better sequence their communication, since knowing about shared and unshared knowledge resources can trigger discussions about topics, with which only one learner in a group is familiar (Schittekatte and van Hiel 1996). According to Wise and Schwarz (2017), group awareness tools help to prioritize the learners' agency:

> "By making people aware of the qualities of their peers, characteristics of the contributions made thus far, or the knowledge development of the group collectively, these tools put the learners' agency to the front of CSCL focus and afford desirable actions among willful learners."

To exploit the potential of these three benefits, we formed groups of learners with complementary knowledge. To assess the learners' knowledge, methods from information mining and learning analytics provide efficient solutions that can be applied to educational data such as essays or homework (Erkens et al. 2016a). Especially text mining techniques can used to convert semi- or unstructured text data into a structured, numerical format (Miner et al. 2012). The structured data can, in turn, be used for grouping learners based on complementary knowledge. For instance, Erkens et al. (2016a) used Euclidian distances and grouped learners starting from the highest difference downwards to form learning groups. Although they found a relation between distance and knowledge acquisition and a suitability of text mining values to illustrate degrees to which learners wrote on specific topics (Erkens and Bodemer 2018), this procedure brought in small classrooms the problem that the average distance of these groups was not higher than it was in randomly assigned groups. We tackle this problem by creating the semantic group formation algorithm that assigns a diversity score to each grouping and selects the set of groups that satisfies two conditions: total coverage of knowledge items is maximized in the set, and overlap is minimized (Manske and Hoppe 2017). This approach is described in detail in section 4.4.2.

In summary, semantic technologies (e.g., text analysis methods, computer linguistics or AI-driven approaches) and group awareness tools have a high potential to enrich learning and teaching in the field of CSCL. These tools can be used to form groups of real complementary knowledge and to visualize the cognitive information resulting from learner-generated artifacts or even from complementary learning material. In this regard, such mechanisms are likely to facilitate or improve cognitive group awareness. Particularly the exploration of the combination of both (knowledge-based) approaches is still underexplored and thus part of this thesis.

3 Technical Architecture

Derived from the literature about learning analytics, we define an architecture for Go-Lab, that enables learning analytics to support learners and teachers. Go-Lab serves as a pedagogical middleware for promoting inquiry-based science education using online laboratories. Using the Go-Lab system, learners create artifacts, which express and externalize their knowledge to some extent. To enable knowledge integration and to define scaffolds that use knowledge representations and open learner models, we design an architectural layer for learning analytics in Go-Lab (see figure 3.1).

This chapter is mainly built upon three publications, which document my work of creating an architecture for learning analytics in Go-Lab (Hecking et al. 2014; Manske et al. 2014). This architecture has been employed to analyze the use of ILS in Go-Lab with the goal to get a (desired) sequence of phases in an ILS and the deviations in the creation process (Manske et al. 2015a).

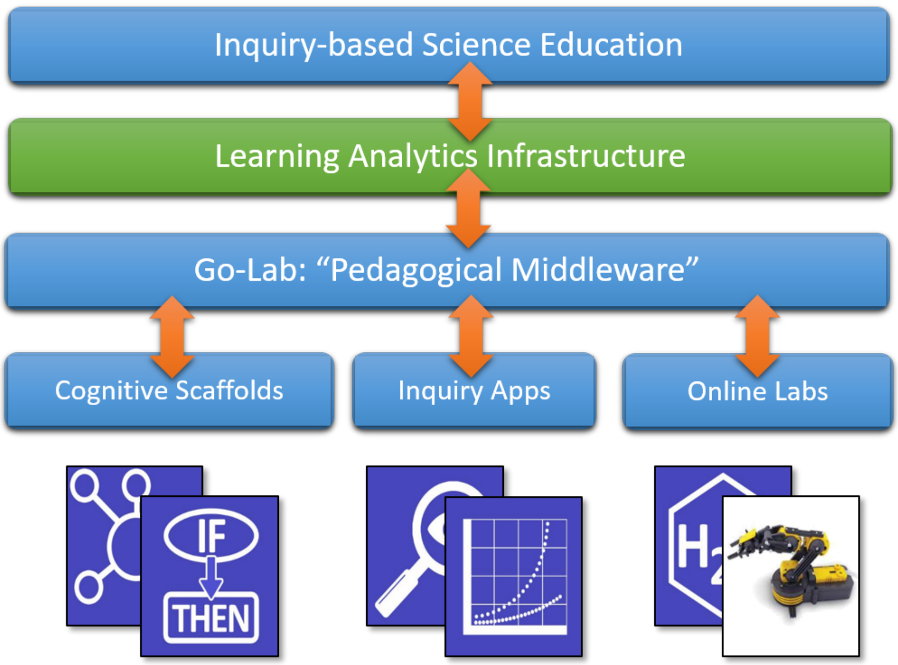


Figure 3.1: The conceptual model of this architectural approach.

3.1 Architecture for Learning Analytics

The analysis of the increasing amount of educational data at large scale in order to improve learning processes has become a growing research topic in the recent years (Drachsler and Greller 2012). The emerging field of learning analytics brings together different fields, i.e., business intelligence, web analytics, educational data mining and recommender systems (Ferguson 2012). Apart from that, there has also been research focused on the pedagogical and epistemological aspects of learning analytics (Knight et al. 2013). However, solutions to support web-based learning environments as a whole with analytics services on the technical level are still underrepresented in the field. There exist learning analytics systems tailored for special use cases. Especially in web-based learning environments with flexible authoring facilities, that are not bound to a single domain, the set of different learning scenarios, which can be supported by analytics features, is unpredictable. Hence, instead of presenting a closed software system for a limited set of analytics tasks, the aim of this work is to design an analytics infrastructure for web-based learning environments, which functions as a general framework for several aspects of learning analytics. This comprises logging mechanisms for student actions, data storage and retrieval as well as intelligent user feedback. Algorithms for data analysis are implemented as independent software agents which makes the infrastructure flexible and extendable. The work is based on current achievements in the ongoing EU project Go-Lab on personalized online experiments with virtual and remote labs for usage in school. To achieve this, Go-Lab offers a web-based platform (Govaerts et al. 2013a), which allows teachers to set up reusable inquiry learning scenarios for students in an easy way. Consequently the descriptions in this work concentrate on analytics for this platform. The following sections describe general characteristics and aspects of architectures to enable learning analytics.

3.1.1 Functional Characteristics of a General Learning Analytics Infrastructure

There are various opportunities to use the Go-Lab environment to create inquiry scenarios with virtual and remote labs. This requires the possibility to create custom analytics solutions as well as the offering of general services by integrating existing systems. While many systems meet the demand of modularity, they dismiss the chance to tailor learning analytics to multiple stakeholders. Analytics services can be used for ex-post analysis by researchers to get insights into learning processes or to design new guidance mechanisms. In contrast to the perspective of ex-post analyses, the learners can also immediately benefit from such systems, typically through interventions.

Action Logging

Before an analysis can be performed, the user activities need to be captured through the system, which can be achieved through action logging. Action logs must consistently reflect the users' actions in the system. This comprises user access to resources as well as specific actions when using web apps. The analytics system observes the learners' actions and thus interpolates their trajectories, which are then persisted in the form of an action log file. Each action that is captured is called an action log, whereas the whole log file can be used to reenact the interaction of learners within the system. Such actions are differentiated from logs that capture the current state of a system, learner or any other entity related to such a system.

The logs have to be in a common and well-defined format, an action log protocol, so that analysis methods can be developed independently. Typical aspects of actions that exist in most of the action logging formats are subject ("who"), verb ("does what"), object ("with which entity") and time ("when"). These basic characteristics are needed to describe an action. Particularly for larger environments or sandboxes, it is necessary to encode the learning context, for example parameters of the learning environment or the learning design, into the log protocol. The context might also consist of activities bound to the context, which is of interest in the arise of social network platforms in education. Due to the simple nature of action logs, there was not a need to create one unified protocol. However, the rise of learning analytics, social network platforms and the approach to provide cross-platform analytics led to the design of unified protocols. The "Common Format" (De Groot et al. 2007) was an XML-based representation schema to import log files from several learning environments into one repository. The activity streams protocol aimed to integrate activity protocols from different social web applications similar to Facebook (Snell and Prodromou 2017). The format is JSON-based and contains in addition to the common fields flexible placeholders for extensions.

During the development of the Go-Lab system, another architectural approach to a storage system with a custom JSON-based logging format for action logs has been advanced, the concept of a learning record store (LRS) and the xAPI format [1]. The LRS can be seen as a generalization of action logging that includes a data store that validates and accepts logs in the xAPI format (Bakharia et al. 2016). The xAPI format ("experience API"), captures the above mentioned typical aspects of actions and can be extended through so-called recipes that describe the action format. One of the goals is to create micro formats that contain detailed descriptions of the particular aspect in the action log. This can be achieved by providing URIs for each entity that point to a web resource with a description of the entity. Following this approach, it is possible

[1] Advanced Distributed Learning (ADL) Initiative, xAPI standard. `https://github.com/adlnet/xAPI-Spec`. Retrieved: 2020-02-10.

to get a detailed picture of the learning activity from an external perspective, which is beneficial for open data or developers of learning analytics applications. However, there are weaknesses in the technical approach such as key names that are not valid to a pure JSON standard and the lack of a reference implementation during the design of the Go-Lab architecture. With respect to the flexibility, current implementations of learning record stores do not foresee extendable learning analytics APIs.

Notifications and User Feedback

According to the target and scope of learning analytics, this can be either seen "institutionally" in a way that it provides general insights in learning, or "directly" that individually and immediately addresses the users that are involved in the learning scenario through the learning environment or the learning analytics platform. In this direct approach, learning analytics can be conceived as a cyclic process in which analysis and feedback steps are interleaved with learning. Referring to the learning analytics cycle, Clow (2012) describes the key to the successful application of learning analytics as "Closing the loop by feeding back this product to learners through one or more interventions". Therefore, appropriate channels for the feedback need to be established. To produce immediate results in the form of interventions, analysis components should be triggered in such way, that notifications can be generated on time to be fed back to the learners. Tools that operate with notifications, such as scaffolds that adapt or (re-)present the interventions, have to be able to handle different kinds of notifications ranging from prompts to reconfiguration of tools to provide tailored guidance mechanisms. An example of such an environment that was capable of adapting tools according to the specific guidance was the SCY-Lab (de Jong et al. 2010). The multi-agent architecture analyzed learners emerging learning objects and according to the configured level of guidance, the concept mapping tool has been reconfigured in order to propose specific concepts through agent-based notifications. Other examples are monitoring tools that provide feedback to the user and thus provide interventions more implicitly by rendering data and dragging the attention of the user to specific aspects through analytics. Thus, better informed decisions can be made by the user.

Ex-post Analysis

In contrast to immediate interventions, a collection of data over a certain period of time is required for many analytics tasks. The scope, target and stakeholders of the ex-post analytics can vary. An example of a learning dashboard that is facilitated in an ex-post analytics setting is the learning analytics dashboard (LAD) by Kim et al. (2016). In this case, learners use the dashboard after the learning phase to review

```
{
    type : "prompt",  // other possible types are "configuration" or "resource."
    importance : "8",  // importance level with range [1, ..., 10].
    target : {
        type : "app",
        id : "provider_id-actor_id-generator_id"   // unique id to address a particular app.
    },
    content : {
        text : "This is an example message"  // message content if notification type is "
            prompt".
        url : "http:\\..."  // url if notification type is "resource".
        configuration : {
            specific_config_parameter : "config_value"
        }
    }
}
```

Figure 3.2: An example of a notification message for a prompt that contains arbitrary or tool-specific configuration parameters.

their online learning behavior patterns. Other examples comprise of ex-post analysis tools that can be used retrospectively by the end user of the learning environment, for example by teachers who want to acquire more insights about the learning scenario. This is particularly the case, when a teacher wants to improve the own teaching practices through systematic observation. In contrast to these individual cases that directly affect the two main stakeholders of learning scenarios, such analytics can be placed at a larger scale. In order to improve a learning environment as a whole, the retrospective analysis of large data sets can be used for providing decision support to educational designers. Additionally, they are also very important as research and validation instruments. Learning analytics and educational data mining can be used in such cases to acquire knowledge about the learners at a larger scale. The intervention does not immediately affect the same learners that produce the data, but following generations of learners. Another reason for long time storage of data is to use real data sets for the data driven development of new analytics and guidance components and the comparison of algorithms on different data sets (Verbert et al. 2012a). These tasks require an adequate data management where data from different sources can be aggregated for analysis purposes. In order to be open, the gathered data must be accessible by various analytics technologies that might already exist outside the infrastructure. However, such openness might have implications on the level of privacy, particularly for the provision of benchmark data or comparative analytics when sensitive information such as grades are presented to teachers.

3.1.2 Go-Lab Learning Analytics Architecture

The Go-Lab Learning Analytics Server provides four interfacing components for the different aspects of data acquisition, the provision of an analysis infrastructure, and mechanisms to provide a technical infrastructure for feedback and interventions in the Go-Lab portal. These interfaces are the *action logging service*, the *notification broker*, the *analytics service* interfaces, and the *web server* interface for deploying learning analytics apps (see figure 3.3).

Logs of learners' activities are the main data source for learning analytics as stated out in the previous chapter. According to the concrete implementation of action logging on the tool level, the action logs approximate the learning activities, for example by tracking the construction of concept maps or hypotheses. The action logging service establishes an endpoint for clients to push event logs of user activities to the server. In the Go-Lab portal, user tracking is handled by the ILS tracking agent. This agent collects logs that are generated when a learner interacts with apps or learning resources and sends it to the mentioned logging service, if the privacy setting allows for it. Action logs are encoded in the well-defined Activity Streams protocol (cf. figure 3.4). In order to keep the client server communication transparent, the action logging client API encapsulates the complexity of sending logs to the server in the right format and can be used by every client component as a JavaScript library. This library handles the injection of metadata based on the context the particular app occurs in.

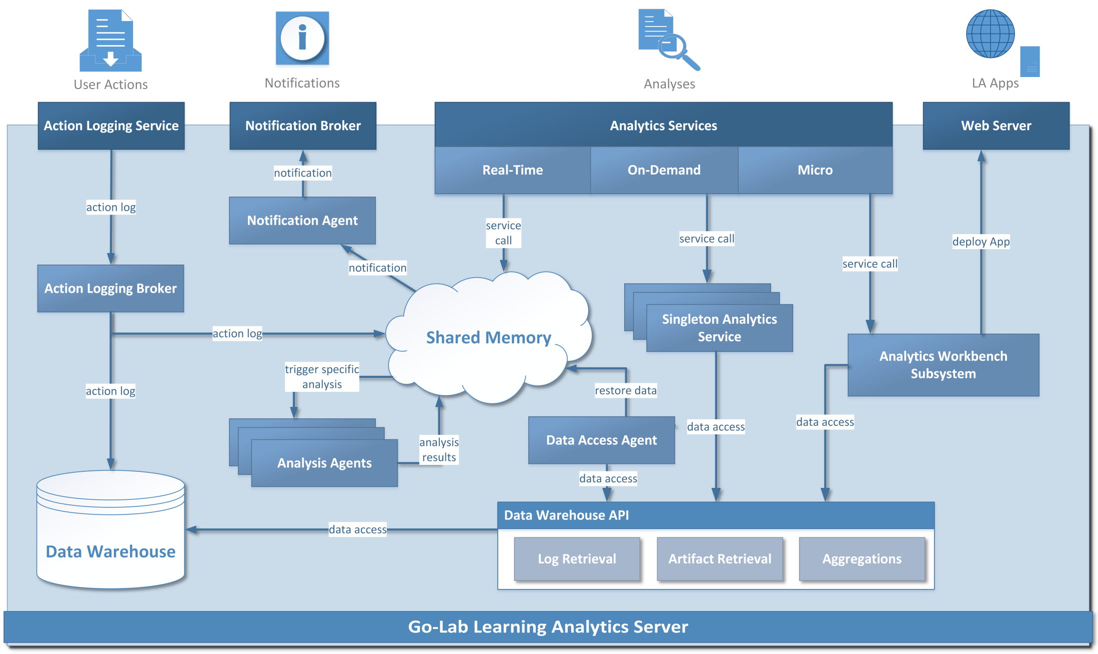


Figure 3.3: The Go-Lab Learning Analytics Server.

The second interface enables the ability to feed analysis results back to the client side for intervention. For this purpose, the notification broker is a dedicated endpoint to establish a channel back to the Go-Lab portal. Clients, i.e., guidance apps in the portal, can register for certain message types by establishing a connection with the notification broker by using the notification client API. This API uses the WebSockets technology based on socket.io[2] to enable a bi-directional communication. Displaying a message that has been created by the backend is completely handled on the client then. However, the notification mechanism relies on a multi-agent subsystem around a shared memory, as outlined in the next section. This agent architecture allows agents to register trigger on certain message patterns (i.e., represented as tuples) which immediately notify them in order to process the data and write responses back into the shared space.

For enabling different kind of learning analytics applications, the analytics services provide interfaces that facilitate different subsystems depending on the certain modality. For real-time access, the agent-based subsystem is used. However, processing data such as real-time updates of learner models is challenging in terms of scalability, as the agents have to hold the data structures during the sessions. Therefore, singleton analytics services have been established that have the ability to aggregate large data sets. For applications such as the usage statistics for online labs over time, it is not necessary to keep the (intermediate) results or states updated, therefore, such services operate on-demand. Finally, micro-services contribute to the analytics infrastructure in Go-Lab. The analytics workbench subsystem is able to execute workflows in the external format (cf. section 3.2).

The fourth interfacing component is a web server, that is necessary to deploy learning analytics apps that can be embedded into the Go-Lab learning environment. This holds for the applications created with the analytics workbench subsystem (cf. section 3.2) as well. However, this subsystem is retrieving data necessary for the execution of the workflows through the data warehouse API, which provides access to action logs, artifacts and aggregated data of both sources.

Another component for the acquisition of data is the artifact retrieval. This service is not directly exposed externally, it is accessed internally through the data warehouse API. This service can be considered as an adapter to different external data sources which allows the internal analytics components to gather artifacts from databases, e.g. metadata repositories. A typical application of this service is to retrieve a list of topics for a specific resource such as an online laboratory from the Go-Lab lab repository (Govaerts et al. 2013a). Another application is the reconstruction of artifacts based on action logs. The retrieved topics in the particular domain can be facilitated to adapt

[2]Socket.IO enables real-time bidirectional event-based communication using WebSockets: `https://socket.io/`. Retrieved: 2020-02-10.

scaffolds or to create a better understanding of the analytics or the corresponding contexts.


```
{
  "published": "2014-03-28T15:28:36.646Z",
  "actor": {
    "objectType": "person",
    "id": "e1b8948f-321e-78ca-d883-80500aae71b5",
    "displayName": "sven"
  },
  "verb": "update",
  "object": {
    "objectType": "concept",
    "id": "a1ad6ace-c722-ffa9-f58e-b4169acdb4e3",
    "content": "osmosis"
  },
  "target": {
    "objectType": "conceptMap",
    "id": "4b8f69e3-2914-3a1a-454e-f4c157734bd1",
    "displayName": "my first concept map"
  },
  "generator": {
    "objectType": "application",
    "url": "http://www.golabz.eu/content/go-lab-concept-mapper",
    "id": "c9933ad6-dd4a-6f71-ce84-fb1676ea3aac",
    "displayName": "ut.tools.conceptmapper"
  },
  "provider": {
    "objectType": "ils",
    "url": "http://graasp.epfl.ch/metawidget/1/b387b6f",
    "id": "10548c30-72bd-0bb3-33d1-9c748266de45",
    "inquiryPhase" : "Conceptualization",
    "inquiryPhaseName" : "conceptualization phase",
    "inquiryPhaseId" : "c7723ad6-dd4a-6f71-ce84-fb1676ea3bbd",
    "displayName": "EnergyCity - Group Phase"
  }
}
```


Figure 3.4: An example of an action log in the Activity Streams format.

Agent-Based Analytics Infrastructure

The service interfaces for action logging, notifications, analytics and artifact retrieval are connected internally in the learning analytics infrastructure. The internal components are depicted in figure 3.3. The architectural approach is based on a multi-agent system with a distributed shared memory following the Tuple Spaces concept. The implementation uses the SQLSpaces framework (Weinbrenner 2012), which provides a shared memory for agent coordination, communication, and a workspace for analyses. Basically it can be seen as a blackboard through which agents exchange messages in the form of tuples as flat ordered collections of data. The basic operations are *read* for reading tuples of specific type, *write* for writing a concrete tuple and *take* for reading and removing a specific tuple in an atomic transaction. Software agents, for example an agent that analyzes artifacts produced in inquiry learning spaces, can register listeners by specifying certain tuple templates. Each template tuple can be defined by the length of the tuple, concrete values or abstract data types on each positional argument. Whenever a tuple that matches such a template is added to the space, the SQLSpaces will actively notify the agent that subscribed by registering the callback on the template. This enables a loose coupling of components because data exchange and communication is completely mediated by the shared memory, manifesting an implicit protocol for agent communication. Agents can be designed to perform analyses and data acquisition autonomously or on demand. This approach has been used successfully in other inquiry learning environments (Giemza et al. 2007). Although the shared memory is persistent, for Go-Lab it is intended as a temporary storage of tuples. For persistent data storage we rely on a (light-weight) data warehouse approach (Inmon 2005). This is a common way to aggregate heterogeneous data from different sources for analytics purposes. The action logging broker (figure 3.3) writes incoming activity logs to the shared memory for direct analysis and agent communication, but also into the data warehouse for a long-term storage. In the data warehouse these activity logs can be enriched by resource content gathered by the artifact retrieval service, for example contextual metadata such as inquiry learning space information. The data warehouse adds a layer for accessing the long-term memory of Go-Lab action logs and artifacts by providing methods for data aggregation and a NoSQL-oriented query language. The data in the data warehouse can then be used for long-term ex-post learning analytics and is available for specialized analysis tools and apps.

Feedback Mechanisms and Example Case

The previously described feedback loop enables direct and immediate interventions through the Go-Lab system. In order to implement such a feedback system, an interplay between logging and notifying is necessary. It incorporates a cycle of logging user actions to the infrastructure, analyzing these events through agents, and notifying end points in the application context of the user. This section outlines the typical information flow when feedback should be given to a student directly by scaffolding apps. Figure 3.5 depicts the complete data flow of a cycle logging and notifying an app in the portal. This feedback loop can be outlined in the following example: a student uses a concept mapping tool and receives guidance in form of a prompt that recommends a specific concept that has not been used yet. The concept mapping app uses the notification API to subscribe to the notification broker by providing a unique client ID. Thus, the notification API registers a listener for messages from the learning analytics server, which is initialized on the startup of the application. Whenever the student modifies the concept map, the action is logged by the corresponding app. The user tracking agent ("AngeLA") takes these logs (1.2) and sends them to the action logging service (2), which itself delegates the log to the action logging broker (3). The logging is prevented by the tracking agent, if it is configured to preserve full privacy with a "do not track" mode. The action logging broker stores the received logs in the data warehouse for long-term storage but also writes the logs as tuples into the shared memory (SQLSpaces). The action logs contain a unique ID for the app that sends the logs. A dedicated concept mapping analysis agent listens for tuples that have been sent by corresponding apps, and hence is triggered whenever action logs from these apps are written into the SQLSpaces (5). When the agent detects that the student constructs a concept map in an inappropriate way, e.g. the user only adds a few sparsely connected concepts, it sends a concept recommendation message back to the app by inserting a notification tuple into the SQLSpaces (6). Therefore, it uses the unique client ID, which can be extracted from the action logs. The notification agent will be triggered by the SQLSpaces when a notification tuple is written into the space (7). The notification broker holds socket connections to all the connected clients. This agent uses the notification broker to send the message to the right client (8). Because the client app is registered with its unique ID as a listener, the broker can choose the right socket connection to emit the message (9). The final handling and displaying of the concept recommendation is under the responsibility of each particular app. However, the notification API comprises methods for internationalization and displaying of agent prompts.

Integration of an External Analytics Framework

The data warehouse layer and the analytics services enable other external analytics tools to be integrated into or to be connected to the infrastructure. To allow for a visual specification of complex analysis workflows, our analytics infrastructure is integrated with an analytics workbench that has been developed in the finished EU project SiSOB (2011-2013). The SiSOB project was devoted to assess the social impact of science using network models and techniques from social network analysis (SNA)

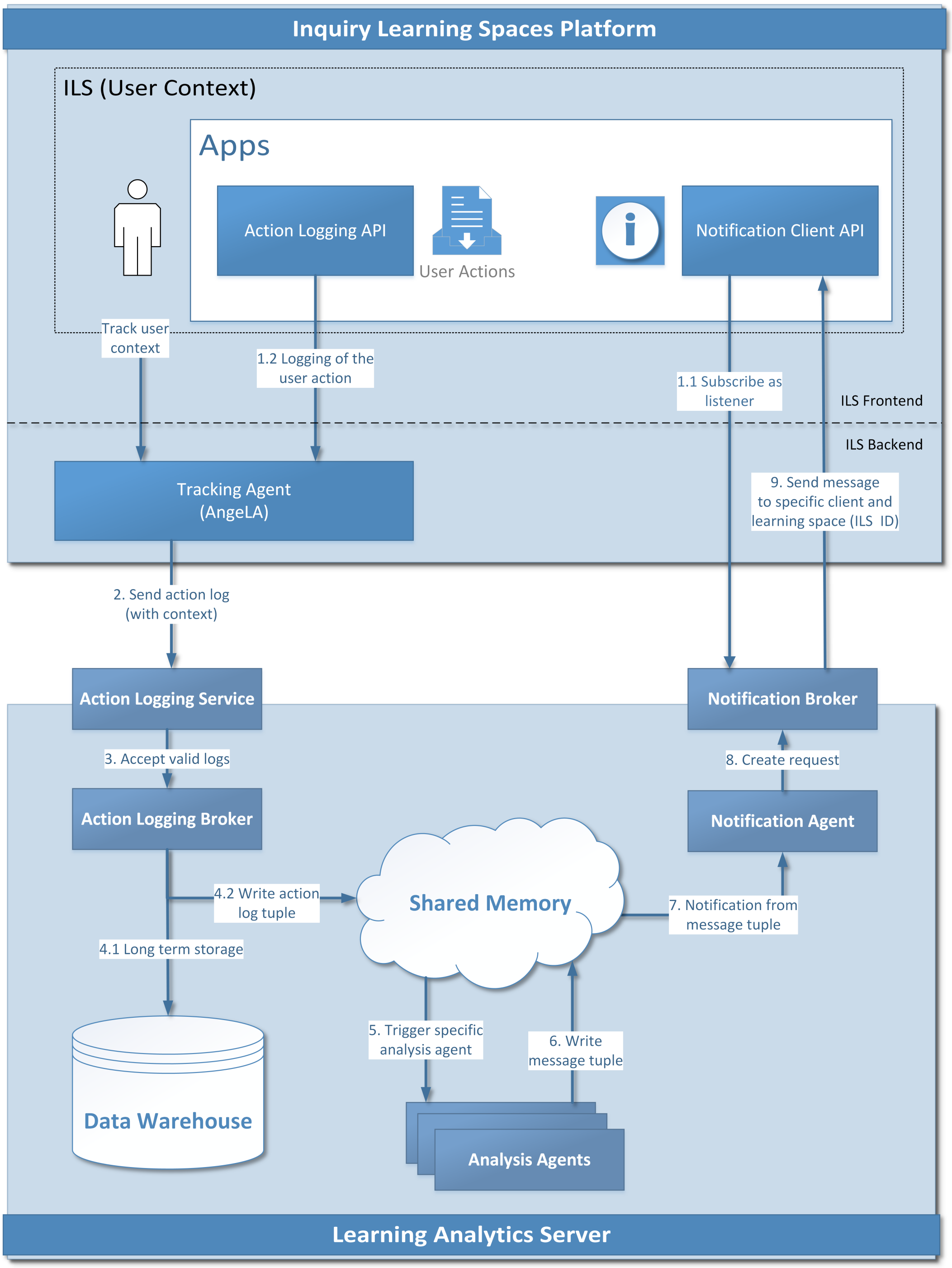


Figure 3.5: Information flow in the feedback loop of the Go-Lab learning analytics infrastructure. Not all the components are displayed in this figure.

that go beyond classical bibliometric methods. A technical outcome of the project was a web-based visual environment for the composition and execution of analysis workflows, including a variety of visualization techniques (Göhnert et al. 2013). The analytics workbench from the SiSOB project has been extended in order to execute analysis workflows based on external specifications and integrated the system that it uses the data warehouse layer of the learning analytics infrastructure in Go-Lab as data sources. In addition, the workflows can be automatically converted into portable and embeddable applications, which makes it a rapid prototyping platform for learning analytics micro services and apps in Go-Lab. In this sense, each app will be deployed along with a specific micro-service, which contains a representation of the analytics workflow. Enabling the visual creation of analytics workflows supports both novices and experts, which extends the target groups of learning analytics in Go-Lab, enforcing a multi-stakeholder perspective. A separation of analysis (the authoring of workflows) and target platform (displaying the results) helps to address different target groups such as students, teachers, researchers and lab owners. This modification and integration is described in more detail in section 3.2 of this work.

3.1.3 Privacy

To support the analysis of learning activities, Go-Lab captures action logs through a well-defined logging format and defines interfaces in its architecture (see section 3.1.2). However, the logging of user activities, particularly in a learning context, exploits data of high sensitivity and inhibits the risk of privacy violations. Two design decisions affect the Go-Lab architecture in order to ensure privacy. First, the artifact storage has been separated from the Go-Lab learning analytics infrastructure. Therefore, the retrieval of user actions and artifacts are separated from each other. The log retrieval can be performed server-side using the data warehouse API of the learning analytics infrastructure. The allocated methods allow analysts to retrieve action logs from different and multiple spaces and time windows. Thus, action log analysis can be facilitated to acquire a global view on usage and behavioral characteristics of learners. The artifact retrieval is always contextualized in an inquiry learning space, which separates this clearly from the action log analysis. Client-side apps are able to retrieve artifacts and process them client-side or server-side through the learning analytics infrastructure under consideration of action logs. However, combining action logs and artifacts requires a learning analytics app for the retrieval to be added (explicitly) to an ILS.

The second mechanism to ensure privacy is an interface to directly control the flow of action logs from the inquiry learning space to the server. Go-Lab employs ways to opt out the logging of learning activities explicitly (Vozniuk et al. 2014). The creator of an ILS, usually the teacher, has the freedom to turn off the logging of user activities ("tracking"). To make this choice of tracking transparent, a tracking agent, namely "AngeLA", is characterized as a member of the ILS, which can be removed from or added to the list of members in the ILS (see figure 3.6). This realizes the user interface metaphor of locking someone out of the classroom and it facilitates common and well-known mechanisms of managing the members of an inquiry learning space.

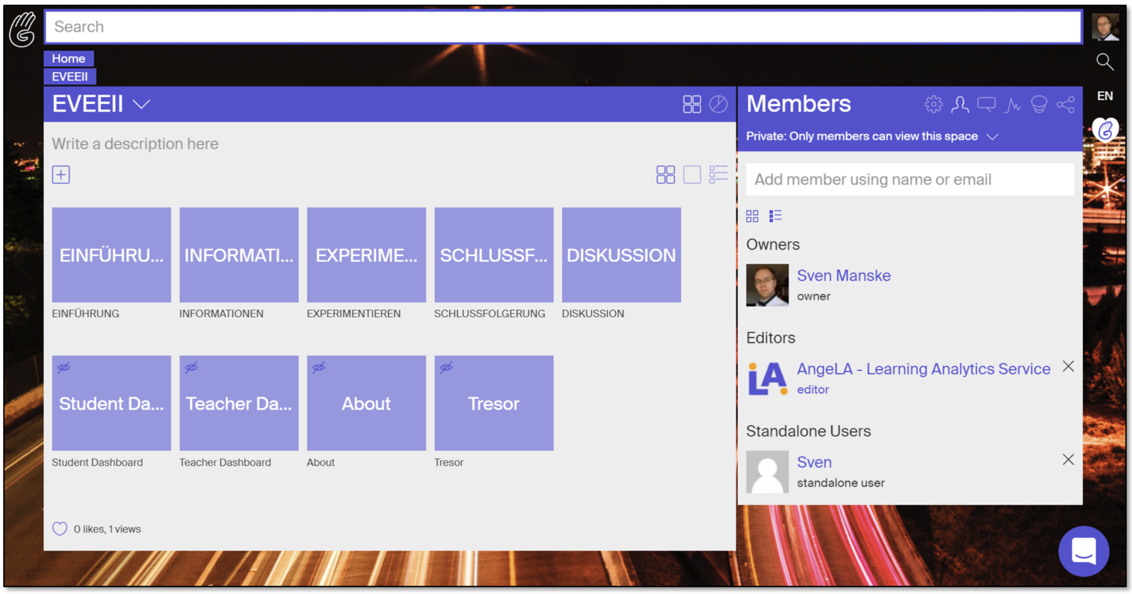


Figure 3.6: The tracking agent AngeLA is a member of an ILS.

The Go-Lab action logging API defines AngeLA as a logging target. In this architecture, AngeLA intercepts the action logging and serves as a single point of logging. Each action log is enriched through the API with contextual information, for example the name of the actor or the ILS ID. Therefore, it employs the OpenSocial API to redirect logs through the backend of Graasp to AngeLA. If the teacher decided to have a "do not track" mode, where the logging is disabled, all action logs are withdrawn. If not, the action logs are then forwarded to the learning analytics infrastructure.

However, such a mechanism does not only provide transparency in handling privacy issues. The analytics system needs to have information about the disabling of logging itself. If the logging gets disabled during a run, this leads to inconsistent traces and false observations. Therefore, the traces need to be flagged as "logging: false". Otherwise it is not possible from the perspective of the learning analytics infrastructure to differentiate the data from experiments that regularly ended. However, indicators as the session time might help to identify incorrect or inconsistent data but there won't be any clear evidence.

3.2 Prototyping and Embedding of Learning Analytics Applications

Modern learning environments such as the Go-Lab portal are nowadays often facilitating mechanisms of learning analytics to support multiple stakeholders such as learners, teachers and researchers. This creates specific challenges to employ the stakeholders with appropriate mechanisms for data analysis and the embedding into the learning environment and possibly into different target platforms (cf. section 3.1.2). It is obvious that teachers and students, being the main stakeholders of the Go-Lab portal, are not able to perform complex analysis tasks in specialists environments such as RStudio or SPSS. However, visual languages or simplified interfaces (e.g. using wizards) are helpful in order to facilitate analytics and guide users through the formulation of such tasks.

In this section, a framework to create reusable learning analytics components that are portable to different target platforms is proposed. The logic of each analysis component is specified in a separate web-based visual environment (or "workbench") from where it is later exported to the target environments in form of a widget-based dashboard displaying embeddable applications. Such embeddable applications are called "apps", "gadgets" or "widgets" in this context and are usually embedded through a specific container such as provided by OpenSocial. Although the approach is more generic, this mechanism is demonstrated in the context of the Go-Lab portal. An example shows how such analytics apps can be created and used to support collaborative learning while the Go-Lab environment itself is non-collaborative.

3.2.1 Analytics Workbench

Initially, the analytics workbench (Göhnert et al. 2013) has been developed during the SiSOB project with the goal to observe science activities through analytical methods. It offers an analysis framework that integrates a wide range of analysis tools and libraries with a user interface that also allows non computer experts to use the full power of the workbench. This is achieved by a pipes-and-filters user interface metaphor, where users create explicit representations of analysis pipelines in a visual language. These workflows usually end in visualizations of the corresponding results, for example a force-based network visualization. However, the original mechanisms are mainly focused on (social) network analysis, for example, to create and visualize citation networks. Therefore, this work also explains the extensions in order to provide Go-Lab specific data processing and analytics methods.

Architecture

As a result of these requirements of the SiSOB project, the workbench combines a multi-agent system as computational backend with a web-based user interface. This user interface comprises explicit representation of analysis workflows as the main building blocks. Analysis workflows are composed graphically using a visual language based on a pipes-and-filters metaphor. In this language the modules ("filters") represent individual analysis steps and the connections ("pipes") between the modules represent the data flow. Such modules can filter, transform or process data and pass it through the pipes to the next filter. See figure 3.10 for an example of this workflow representation. In this example, centrality measures are filters that are used to enrich the data with certain measures. On the technical level, the workbench is divided into two parts: (1) the computational backend, which is realized as a multi-agent system, and (2) the frontend, which is implemented as a web-based user interface using modern web technologies.

In the computational backend, each of the filters in the analysis workflow corresponds to a single agent in the multi-agent system. The multi-agent system in the framework of the analytics workbench uses a SQLSpaces server (Weinbrenner 2012). The SQLSpaces framework is an open source implementation of the Tuple Space concept (Gelernter 1985), which especially focuses on heterogeneous multi-agent systems in terms of language heterogeneity. The communication between the components of the system (agents) is based upon a simple protocol that consists of tuples that are written into a certain space as medium of communication. In this sense, tuples define an implicit and application-specific protocol for the communication in the multi-agent system. Elements in the tuples can be used to characterize the function of a tuple or how it can be interpreted. Agents itself register callbacks at the tuple space using certain patterns in order to get notified when a particular tuple matches the request. In the protocol of the analytics workbench, the two main elements are *command tuples* to control the execution of a workflow, and *data tuples*, which are used to transfer data between agents. Since all steps of the analysis processes are encapsulated in individual agents, the functionality of the workbench can be easily extended by adding new agents. With this approach, it is easy to implement new filters for data input, data transformation or visualization of workflow results. All the components are loosely coupled through the system. The only requirement for an agent is a connection to the SQLSpaces server and the accompanying tuple-based communication protocol. This is, for the analytics workbench, encapsulated in a dedicated agent API.

For the frontend, a custom web server based on Node.js is used to provide the user interface and to connect the user interface to the computational backend. This is achieved by transforming the external format of the workflow representation, which has been created in the user interface, into command tuples to trigger and control the workflow execution in the agent system of the computational backend. Additionally, it presents the results of the analysis processes to the users. Most of the filters consist of a Java agent connected to the workbench framework, and a description to display the specific parameters of filters in the frontend. In case of output filters that visualize results, it incorporates interfaces for displaying the visualization in a web browser and for offering interactive data exploration to the users.

Data Exchange Formats and Standards

The workbench uses its own internal formats to represent graph and data table information in each phase of the analysis. Both are based on a JSON structure with two main sections: metadata and data. The metadata section contains supplementary information to enrich the available data. This is the case for calculations such as centrality measures, which are added to the data as annotations to enrich the original data set rather than removing the previous data. Therefore, as an internal format it was necessary to allow for annotations. However, for internal data flow and as an external exchange format, the analytics workbench supports widely used standards such as Graph Modelling Language (GML) for network data and comma separated value (CSV) files for data tables.

3.2.2 Go-Lab-specific Extensions to the Analytics Workbench

The existing analytics workbench has been extended in this work to allow the visual creation of learning analytics apps that can be embedded and contextualized in target learning environments such as the Go-Lab inquiry learning spaces using portable formats. This section presents a short overview of the extensions to the analytics workbench. First, it is required to adapt the data formats to meet the requirements of the Go-Lab project (action log import). Second, a mechanism to create apps from analytics workflows needs to be implemented. In contrast to the analytics workbench, which allows to import data from static contexts, the aim of this work is to contextualize the learning analytics apps inside the learning environment. Therefore, dynamic import facilities need to be established, which demands for a certain flexibility in the filter description. Thus, it is necessary to enable parametric input in order to take context (such as the current ID of the inquiry learning space) into account. Finally, the backend of the workbench needs to be modified in order to trigger the external execution of workflows through the apps, because the apps themselves are integrated into a widget container (e.g. OpenSocial) without a direct connection to the backend of the analytics workbench.

Data Format Extensions

The use cases of the analytics workbench of the SiSOB project are typically restricted to the analysis of network structures, for example by creating citation networks. However, in Go-Lab, network structures are of less interest by the nature of the (non-collaborative) learning scenarios. The typical entities to be analyzed in the context of Go-Lab are traces of learners. The traces are typically captured through event logging in the Go-Lab portal. For analyzing log data, the workbench has been extended to support the JSON-based ActivityStreams format, which is used by OpenSocial compliant platforms. ActivityStreams follow an "actor - verb - object - target" metaphor, which represents information about who did what with which object on which artifact (cf. section 3.1.1).

Learning Analytics App Creation

The process of creating a learning analytics application can be summarized through the following steps (cf. figure 3.7):

1. **Workflow creation** This consists of connecting specific data sources, filters, format converters, analytical methods and algorithms in order to create a functioning analysis workflow. The workflow is created in the user interface of the analytics workbench. A typical definition of a workflow comprises a source and a sink. The data source can be an import from common, public examples, but for the Go-Lab case it connects to the learning analytics infrastructure to import log data. In many cases, the sink is a visualization of the analysis results, for example a force-based graph layout highlighting centrality measures in an artifact network of a learning platform. However, in order to continue with the next step, the workflow needs to be correct in terms of consistency regarding the data flow (i.e., corresponding formats) and thus executable.

2. **App export** After the creation of the workflow, it can be exported as an app. A templating engine creates a file that can be rendered or embedded in the target platform, for example as an OpenSocial gadget. The gadget file is hosted in the integrated web server of the workbench and can be accessed through a web link shown to the user. Figure 3.8 shows the technical implementation of the app creation for the case of OpenSocial (cf. section 3.2.2).

3. **Embedding** The embedding of the app through the accessible link depends on the target platform. Gadget containers like Apache Shindig[3] provide the possibility to embed the gadget file directly through a link. This is the common way

[3] The Apache Software Foundation, Apache Shindig: `https://shindig.apache.org/`. Retrieved: 2020-03-16.

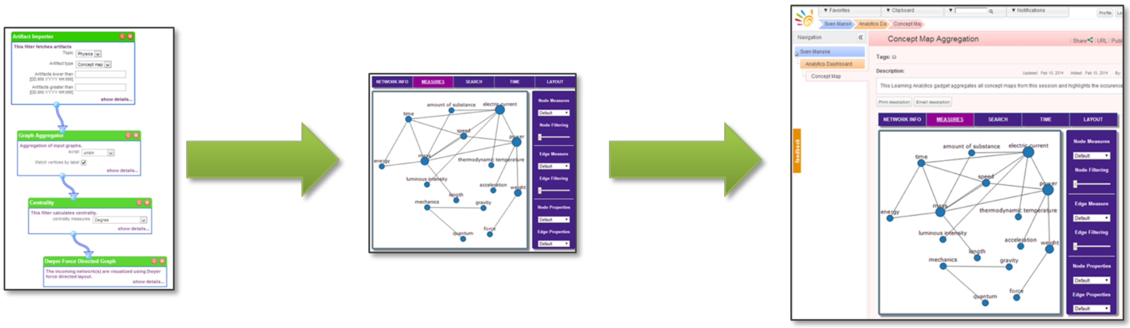


Figure 3.7: The process of widget creation: a predefined workflow (1), which creates a visualization through an output filter (2), can be transformed into a widget. This widget can be embedded into the target learning environment such as Go-Lab (3).

for platforms based on OpenSocial like Graasp, which is the technical basis for the Go-Lab learning environment. The Graasp platform integrates an Apache Shindig to render the gadget files. However, the mechanisms to embed applications require a certain degree of integration, particularly for accessing the contextual parameters such as the identifier of the specific inquiry learning space the app is running in (cf. section 3.2.2).

Multi-Contextuality and Parameterization

Apart from the changes to the mode of execution, specifically the additional user interface controls to create an app, the contextualization of workflows into the target learning environment is crucial. While workflows in the analytics workbench operate on static data sets, the integrated analysis of learning spaces from within those spaces requires the access to contextual parameters and the corresponding data. When creating a workflow, data sources are explicitly connected through filters in the workbench. The requirement of such a flexible system for portable and reusable apps calls for multi-contextuality of learning analytics applications and workflows. As an example, consider a workflow which merges concept maps and visualizes the aggregated graph. The data source might be either all concept maps in the data warehouse or all concept maps from the context of the app, which is the specific learning space in a portal. To achieve this, dynamic context variables (e.g. "$session.id") have been introduced to parameterize workflows independent of static values. The context wrapper of the app engine will inject the particular context and replace all dynamic context parameters at run-time, for instance, when the app triggers the workflow execution. However, the injection needs interfaces to access environment-specific functions such as retrieving a list of users or IDs. Therefore, each context wrapper contains specific libraries to access these functions and provides interfaces to the learning environment. As the data might change during the run-time, this needs to be fetched dynamically.

Backend Services

The modifications to the analytics workbench comprise two major changes to the backend services: (1) a hosting of apps, and (2) services for external workflow execution. Apps, created by the templating mechanism (section 3.2.2) are persisted in specific file formats and kept on the workbench server. Standards like OpenSocial use servers in order to read and render file formats such as OpenSocial gadget XML files that represent embeddable apps in a standardized format. Therefore, the workbench server needs to act as a web server to make those files accessible by external containers, for example the gadget container from the Go-Lab portal. Therefore, the analytics workbench has been modified in order to host and deploy gadget XML files for the corresponding containers that can fetch the specific app code and render it inside the target system. Additionally, to provide learning analytics as a service, workflows need to be executed directly through the backend (without directing the frontend of the workbench). Each app has a persistent representation of the corresponding workflow in it. The workbench has been extended with a REST web service interface to accept the external format of workflows. Their external and parameterized representation will be rendered to a concrete JSON-based format and sent as a parameter to the web service interface. This service interface takes the workflow and executes it directly in the backend. The response contains the necessary information for the app to display the result, i.e., the visualization output. In summary, the gadget within the OpenSocial container bridges the two decoupled platforms Go-Lab and the analytics workbench. Thus, it is possible for novices or non-experts in analytics to create and deploy analytics services for such external platforms without the need to modify the backend of the learning environment.

Visual Framework

When initialized, an app triggers the execution of its embedded workflow through the workbench web service interface, which results in a particular data visualization. The visualization framework of the analytics workbench is web-based and contains resources consisting of HTML code, JavaScript libraries and the relevant analytics data sets to be visualized. The app picks up the created resources, and injects them into its rendering context. The visualizations are implemented using the D3.js[4] library, which delivers dynamic web-based visualizations using the SVG standard, HTML5 and CSS.

[4] D3.js, Mike Bostock (2019): `https://d3js.org/`. Retrieved: 2020-03-16.

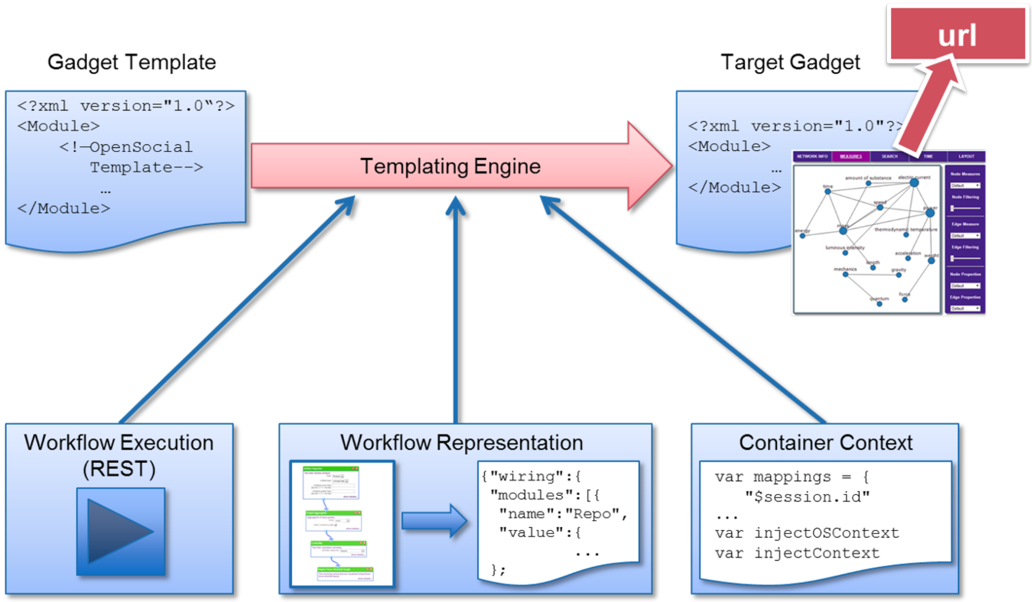


Figure 3.8: The technical implementation of the app creation through the analytics workbench.

Following this approach, the visualization is created dynamically by triggering a workflow execution, instead of presenting static, pre-calculated results. Therefore, it is possible to refresh the results simply by repeating the analytics workflow execution on the workbench server and updating the visualization. This enables updating visualizations dynamically when the data inside the context changes without limitations to the analytics platform.

App Templates

When a user of the analytics workbench creates a workflow that can be visualized, the "export app" button in the user interface of the workbench triggers a templating engine, which creates a concrete instance of the app by using template files and libraries. Figure 3.8 shows this process for OpenSocial gadgets that are used for the Go-Lab learning environment. The templating engine includes three elements in the app: (1) the code to execute a workflow via REST interface of the workbench, (2) a persisted external format of the workflow, (3) libraries to provide access to the context of the container (cf. section 3.2.2). When the templating engine finishes the app creation, the file can be accessed through an URL. The app itself is hosted in an web server extension of the workbench. In Go-Lab, such URLs can be added to an ILS easily.

3.2.3 Applications in Go-Lab Using the Learning Analytics Workbench

In the context of the Go-Lab environment, teachers embed predefined analysis workflows in the form of OpenSocial gadgets into their personal learning space, in order to compose a set of analysis tools to support monitoring and supervision. The following section outlines two use cases of the proposed system. The intention of these example cases is not to introduce novel or advanced analysis algorithms, but to show the flexibility of such a system. It covers different areas of learning analytics by connecting different methods such as artifact and social network analysis in predefined workflows that can be handed over to novices. This will be demonstrated by an analysis of concept maps created by students. The first example (Concept Map Aggregation) has been used in a study about group formation presented in chapter 5. The second application has been provided to teachers but not been used in empirical evaluations.

Analyzing Individual and Aggregated Concept Maps

Concept maps have been used as an instructional tool in inquiry-based learning contexts that supports the students to structure, conceptualize and externalize their scientific knowledge by explicating concepts and their relations (compare section 2.1.2). The Go-Lab learning environment provides a concept mapping tool as one of the existing scaffolds in order to support inquiry activities. The concept mapper makes use of the existing APIs to connect to the learning analytics infrastructure, particularly for logging all actions. The traces consist of actions like adding, removing and labeling of concepts and relations. From such a sequence of actions, the states of concept maps can be reproduced at any point in time on the server side. These log protocols and reconstructed concept maps are a valuable source of data to gain insights on how the individuals structured their own knowledge, but also to assess the group knowledge of a whole learning group (i.e., a class using Go-Lab) for example by applying methods of social network analysis (Clariana et al. 2013; Hoppe et al. 2012). In this approach, an aggregation of concept maps are used to make statements about the knowledge state of a whole learning group. Formally, a concept map can be seen as a multi-graph with labeled nodes and edges. Aggregating concept maps can be operationalized as the union of such graphs. Thus, it means to overlay all individual concept maps. The resulting multi-graph of this operation contains all concepts (nodes) and relations (edges between nodes) that occur in at least one individual concept map. As concept maps can be seen as an externalization or explication of a mental model, this operation presents the group knowledge in one representation. Figure 3.9 shows an example of the embedded app in an inquiry learning space in Go-Lab displaying the aggregated concept map. Section 4.2 shows a more general approach of shared group knowledge models.

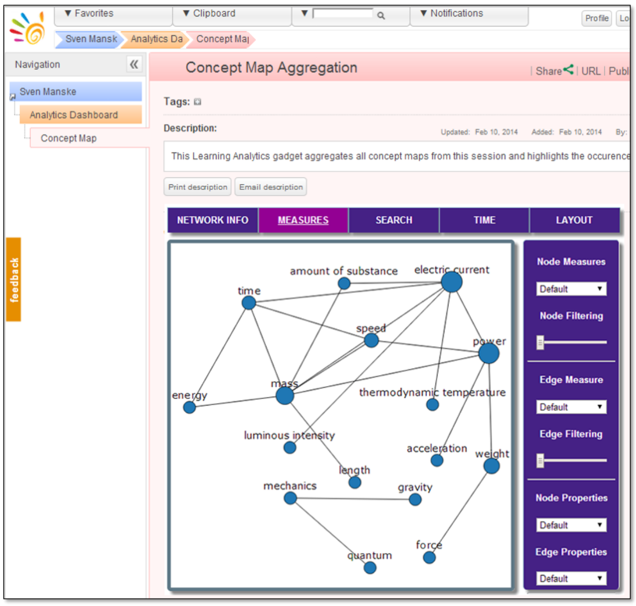


Figure 3.9: Network derived by aggregating individual concept maps. Node size corresponds to betweenness centrality.

To quantify the group knowledge about certain concepts, the number of occurrences of concepts and relations in the set of individual maps are counted and assigned as node attributes in the aggregated map. Such a quantification can be interpreted as a measure of consensus or - as an inverted measure - to identify singularities in the knowledge model. Additionally, centrality measures known from the field of social network analysis like degree, betweenness, closeness, and eigenvector centrality (Wasserman 1994), among others, can be calculated to get additional weighting parameters of concepts (Clariana et al. 2013; Hoppe et al. 2012).

The whole process involves (1) data import, (2) aggregation of concepts, (3) calculating centrality measures, and (4) visualizing the aggregated graph in a force-based visualization. Figure 3.10 shows the workflow in the user interface of the analytics workbench, where the user has several degrees of freedom to parametrize the workflow, for example, by setting the artifact type to "concept map". The first step in this workflow is the data acquisition. The "artifact importer" is one of the Go-Lab specific extensions to import artifacts directly from the Go-Lab learning analytics server. The workflow shown includes the context variable "$session.id" in order to contextualize the workflow to retrieve the ID from the inquiry learning space the app runs in.

The second step is to aggregate all the individual concept maps by performing the union operation as described above. The aggregation operator can be either the union or the intersection of the graphs. The matching of edge labels can be omitted, which is a practical relaxation due to differences in creating concept maps. Learners might

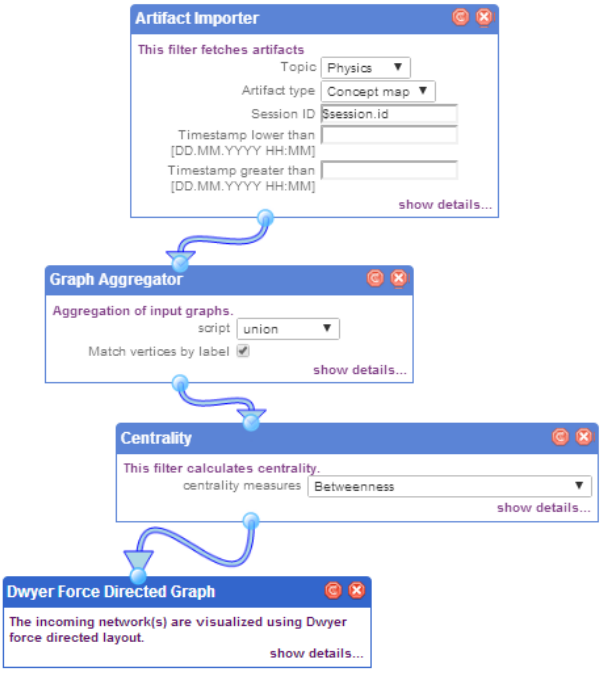


Figure 3.10: Centrality calculation of aggregated concept maps as a pipes-and-filters workflow.

use inverse relation, for example, "increase" versus "decrease", which leads to inverse directions of vertices. To automate this, it needs additional knowledge to be encoded into the system in order to perform a semantic matching of the labels. Therefore, this can be eased by omitting the edge labels for the aggregation. As a third step, a centrality calculation is added as a single component that has been part of the regular workbench filters, which takes the aggregated graph as an input. At the end of the workflow, the results are visualized using the Dwyer force-directed graph visualization technique (Dwyer 2009). It arranges nodes automatically on the screen based on physical models of attractive and repulsive forces that are assigned to each node. The size of each node is either determined by the number of occurrences or any of the specified node measures such as the centrality. The visual interpretation of the results is dedicated to the teacher and (apart from the interpretation of centrality) quite intuitive. However, the assembly of the particular workflow requires some expert knowledge. This lines out the use of separating the two steps of creating the workflow in an expert platform (analytic workbench) and delivering the app through the Go-Lab ecosystem for novices. Once the workflow has been constructed, its internal representation can be stored and reused to execute the process several times without consulting the user interface of the workbench again. With the export as OpenSocial gadgets, predefined workflows can be embedded and executed from any widget container like ROLE or Graasp if the context injection for parameters like session IDs is present in libraries of the app. Thus, the app itself only displays the visual results while hiding the complexity of the analysis workflow behind it.

Comparing a Concept Map with a Reference Map

A common task in the analysis of concept maps is the comparison of concept maps created by students with an expert concept map. Such an expert map can be a concept map created by an expert (Conlon 2004) or a domain ontology (Hoppe et al. 2012). The comparison to the expert map first needs to retrieve the aggregated concept map. For this task, most of the filters from the previous example can be reused. The only change affects the centrality filter, which has been replaced with a "Graph Comparison" filter. This filter takes two graphs as input and returns the union, similar to the previously described "Graph Aggregator". Unlike the Graph Aggregator it can only handle two graphs as input. It checks which relations between concepts occur exclusively in one of the two graphs and which relations can be observed in both graphs. Thus, it highlights the differences between the first (a student map or an aggregated concept map) and the second input (expert concept map). The output of this filter is an aggregated graph with edges decorated with attributes indicating whether the edge exists only in graph 1, only in graph 2, or in both graphs. The visualization component uses these attributes to color the edges accordingly. An example of the output of this workflow is shown in figure 3.11, which compares a student map with an expert map. In analogy to the first example, the visual results of the comparison of the aggregated concept map and a reference map can be very useful for a teacher to uncover potential shortcomings in the group knowledge of the learning group.

3.2.4 Limitations

The approach of creating portable apps from analysis workflows as presented in this work is particularly helpful for the rapid prototyping of learning analytics apps. It helps to kick-start the delivery of analytics to novices in this field, particularly teachers and students, who can benefit from insights. Additionally, analysts are supported in developing workflows using existing analytical methods. The framework automatically creates apps that are embeddable in portal systems, such as the Go-Lab learning environment. The Go-Lab environment is based on the Graasp system, which uses Apache Shindig as an OpenSocial container to render the apps and provide interoperability. However, several limitations exist.

As a first limitation, privacy issues might exist - depending on the widget container of the target platform the app needs to be embedded into. Privacy and rights management on the level of data access is not intended to provide a fine-grained differ-

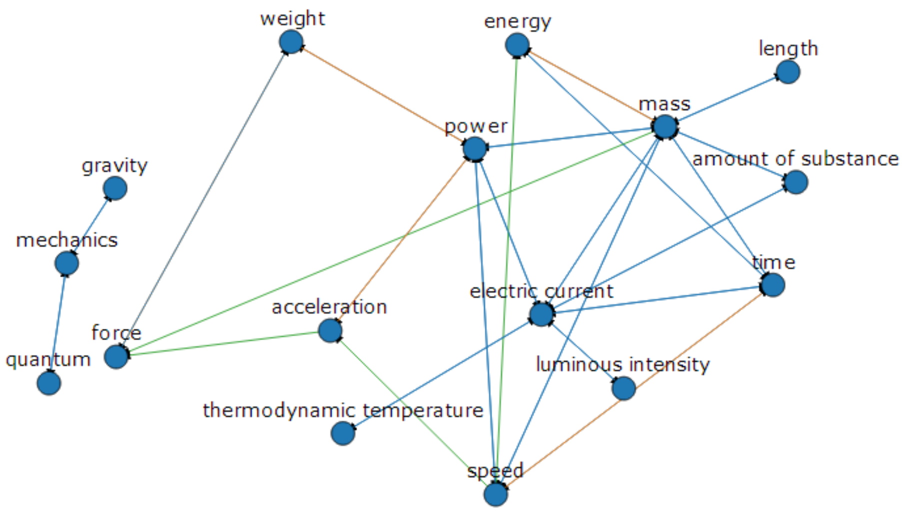


Figure 3.11: Comparison of two concept maps. Edges that only occur in the expert map are green. Edges exclusively in the student map are blue. Brown connections are created both by students and the expert.

entiation of roles in OpenSocial (Apache Shindig). Thus, an app is either allowed to retrieve all data in a space or none. As a second limitation, the execution of workflows creates a performance bottleneck on the workbench server. The workbench is a single server to execute workflows, which makes use of a multi-agent system in its backend (SQLSpaces). However, for each filter in a workflow, an agent gets instantiated, which might limit the capacities of the system, particularly of the shared memory, with many users accessing and using the web service to execute workflows. This can be prevented by further decoupling the system and creating micro services for specific reoccurring workflows. As a side note, this system is intended for the rapid prototyping of learning analytics app and not for a large scale implementation.

The presented framework system provides lots of possibilities to embed applications of learning analytics into portal systems to support different stakeholders. The example applications demonstrate the possibilities of such open and flexible systems. It is easy to extend the templating system with little programming efforts to support and target more platforms. In addition to many already existing analytical methods in form of filters and data converters, the workbench can be easily extended to further push the boundaries to open learning analytics platforms. Besides the obvious stakeholders of learning analytics in Go-Lab, namely students and teachers, it gives researchers, institutions and authorities the opportunity to explore their data sets and to conduct analyses with the means of conserving the freedom of decision.

3.3 Analysis of the Go-Lab Environment

Go-Lab provides teachers and instructors with a specific inquiry model to structure their classroom activities according to a set of phases (Pedaste et al. 2015) rather than a cyclic model (see section 2.2.3). This sequence is not fixed to the prescribed phases, it can be customized and modified by teachers. This affects the number of phases, the names and corresponding functions regarding the inquiry activity. For example, a teacher might split the *conceptualization* phase into two phases *question* and *hypothesis*. Additionally, resources, apps and labs can be integrated in these phases. From the theoretical point of view, the resulting pedagogical structure enriched by apps, online labs and learning resources constitutes the *teacher model*. In practice, the teacher model is represented by an inquiry learning space in the Go-Lab environment. Then, the students are expected to go through the different phases and their content, either sequentially or moving back and forth between them. This choice for a pathway or trajectory within an ILS might be dependent on their personal preference, learning flow, or the specification of the task in the learning environment.

Two years after the Go-Lab project has been started, many learning spaces have been created and used for learning, teaching, development and testing. The purpose of this work is to analyze how teachers and students adapt and follow the inquiry-based learning approach that has been proposed in the project. Therefore, the 102 most frequently used ILS have been selected and evaluated in an exploratory study. Consequently, a processing chain has been defined and the corresponding analytical model for this work has been applied to different, heterogeneous data sources available in Go-Lab. A general architecture to integrate, filter and analyze contextual, activity-, and artifact-related data has been utilized to generate higher abstractions such as the learning process models (LPM). Furthermore, different metrics have been applied to determine deviations from the intended models ("out-of-order" behaviors) of students and teachers. The results from the data analyses have implications for teachers, researchers and pedagogical instructors in the field of inquiry learning. The work presented in this chapter is based on a publication for the International Conference on Computers in Education (ICCE) in 2015 (Manske et al. 2015a).

3.3.1 Data Analysis and Processing Chain

This section describes the different data sources and analysis procedures used to obtain the results presented in this work. An overview of the data collection and processing is provided in figure 3.12. The architecture of Go-Lab consists of a flexible backend with services for the collection and aggregation of action logs (cf. section 3.1.2). This can be facilitated to capture traces of learners in the Go-Lab learning environment. The analytics architecture consists of an abstraction to retrieve aggregated log data

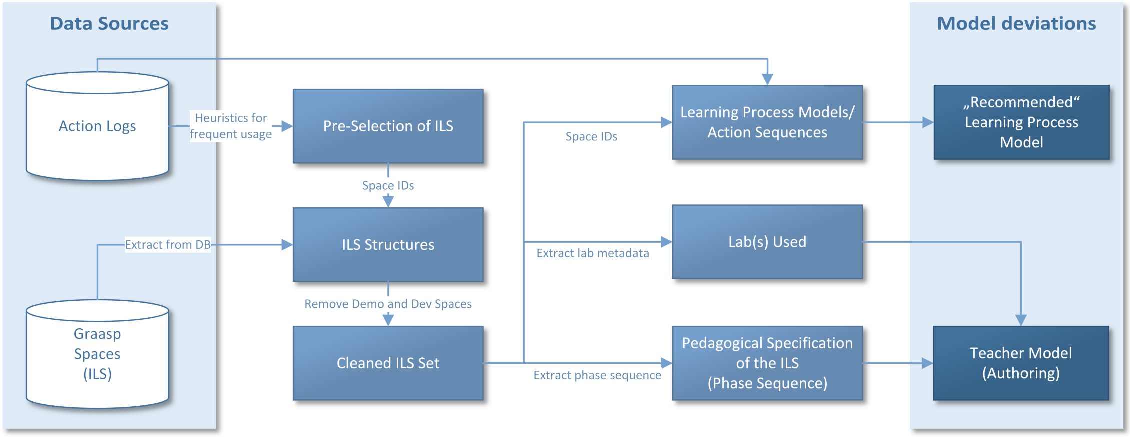


Figure 3.12: The data processing is based on action logs and contextual data from ILS in Go-Lab.

("data warehouse API"). In addition, contextual data from the Go-Lab learning spaces have been used to complement the behavioral aspects of learners' data. The composition of such combined analysis workflows as micro services and the embedding of visual results for learners, teachers and researchers in the context of Go-Lab has been described in section 3.2. For our analysis, we utilize this architecture for the collection and processing of data, as well as the embedment of analytics apps for the proposed prototypes (cf. section 3.4).

Data Sources

The analysis is mainly built upon two different data sources: action logs (behavioral) and Go-Lab inquiry learning spaces from Graasp (contextual). The contextual data from learning spaces contain information about the authoring of each space, particularly which apps, resources and online labs have been used in which phase. In summary, the log data capture the learners' behavior, while the contextual data represent the teachers' specification as a pedagogical and instructional reference frame for the learners.

In our statistics we count online labs as a special type of app – in most cases there is exactly one lab per ILS. In the Go-Lab portal, teachers usually start to define an experiment as a central component for the ILS. Thus, they are adding a single online lab that corresponds to the experiment to the space, which has been communicated as the foreseen process in creating ILS from the Go-Lab portal. However, the difference between resources and apps is quite important: while apps and online labs are predefined by Go-Lab or third-party app and lab providers, the resources represent learning materials. Such materials are usually selected or even created by teachers and added to the space explicitly, which represents a different degree of customization on the part of the teachers. Although the action logs captured by the system are contextualized within the ILS (e.g., identifier), the combination with the contextual data from the spaces completes the picture that is drawn throughout this analysis.

Data Set and Preprocessing

The raw data set consists of 2826 existing inquiry learning spaces in Graasp. First, these spaces have been ranked according to the user activity registered, i.e., based on the amount of action logs. Spaces that do not seem to be related to "real" Go-Lab classroom activities (development, demo, testing spaces) need to be filtered out. To determine which learning spaces have been used frequently, a threshold based on typical values for classroom size and a minimum amount of action logs per user has been defined. This restricts the set to those learning spaces, in which at least every ILS is visited and an app has been used. The product leads to a threshold of 500 actions and 10 users per space as a minimal requirement to be included in the filtered set of learning spaces. This indicates a minimum of activity to be useful for further processing and to draw meaningful conclusions from the results. Examples of actions that have been captured through the learning analytics infrastructure are either space-related activities such as *logging in to an ILS*, *starting an app*, *changing a phase*, or tool-specific activities such as *adding a concept to a concept map* or *creating a hypothesis*. After the filtering, which was based on these criteria, has been performed, the *Cleaned ILS Set* has been made up by 102 ILS that hold a certain degree of activity.

Data Processing and Analytics

For the exploratory analysis, generic metrics that describe the ILS have been used. This comprises the amount of user activity, the type of learning phases and the tools that have been used. To indicate the volume of activity in an ILS, the number of logged user actions per ILS has been employed. Thus, a high number of logs indicates a high volume of user activity, which points to a more active ILS. The number and the particular sequence of phases has been used to describe and characterize the learning spaces. The Go-Lab platform provides and thus recommends five typical inquiry phases (orientation, conceptualization, investigation, conclusion and discussion), but the teacher may customize the structure of the ILS. This can be achieved by introducing new phases, by removing or by renaming existing ones according to the planning of the inquiry-based learning activity. A coding scheme for the inquiry phases has been created using this typical inquiry model, while additional categories for non-standard or non-default phases have been added. Apart from this, Graasp as the technical platform of Go-Lab provides an exceptional space that does not correspond to an inquiry phase. The so-called "Vault" is a subspace used to store learner-generated artifacts (e.g. concept maps) that are created from the different apps inside an ILS. The vault subspace is an interface (particularly for teachers) to access and manage all the learner-generated content that belongs to an inquiry activity. This can be understood as a graphical and technical interface to the artifact storage. For the analysis, the existence of a vault space was used as a binary variable to describe the structure of the ILS, along with the number of inquiry phases and their respective sequence. Furthermore, the teachers can add learning resources, apps and online labs to an ILS in order to design and structure the activity. The number of resources, apps and labs is used as metrics for the description of the ILS.

Based on the aforementioned metrics, a descriptive analysis of the data set has been conducted. Moreover, the ILS of the data set have been clustered using the metrics as attributes following a k-means clustering. The results of the descriptive analysis are presented in the following section. Finally, by analyzing the tools and resources added to the different phases, a collection of design trends and lessons learned can be derived from the results. The Go-Lab approach induces a specific, recommended inquiry model for which we have investigated, whether teachers adapt or customize their spaces. This analysis puts a global view on the contextualized observation of learners' behavior in inquiry activities within the reference frame of teachers' specifications for ILS. This combined study of the two perspectives -the teacher and the learning process model- may provide insights on deviations of students' practice. From the run-time perspective, these deviations from the actual learning processes or sequences of phases can be measured to determine "out-of-order" behavior relative to the pedagogical specification by the teacher, which has been created during the design of the ILS.

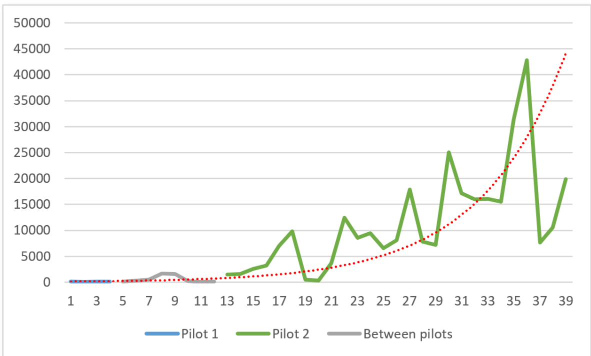


Figure 3.13: Activity during the pilot phases, based on the aggregated number of action logs. The domain axis indicates the week, ranging from 2014/04/14 to 2015/05/17

3.3.2 Results

The cleaned data set that has been used for this analysis consists of 102 ILS built by teachers who have used the Go-Lab platform for inquiry activities. The aforementioned metrics have been used to filter out inactive spaces or spaces with a low number of participants. The teachers were able to plan the activities in various phases as well as to choose and distribute tools and resources in the ILS. The activity volume of the learning spaces ranged from 500 to 16426 logs (on average 2672 logs per space). Figure 3.13 shows the distribution of action logs over time.

Number of Inquiry Phases

The majority of the learning spaces (60.78%) have been created using the recommended scheme that consists of the five inquiry phases promoted in Go-Lab as a default: Orientation, Conceptualization, Investigation, Conclusion and Discussion. However, there were cases where the teachers deviated, customized or enriched the original model, either using less (three phases as a minimum) or more phases (eight phases as a maximum) for the planned activity. In those cases, additional phases such as "Data interpretation" have been created. Such a phase can be seen as a subphase of the investigation, which is still compliant to the Go-Lab inquiry model. The recommended model is a synthesis of various inquiry models from the literature (Pedaste et al. 2015). In some of the cases teachers split the original phases into multiple subphases either because each phase was too long or to give more emphasis on certain processes. Due to the user interface of the inquiry learning platform, adding many resources to a single phase forces learners to scroll a lot as the content (resources and apps) is aligned vertically inside each phase. In one of the examples, a lecture about electronic circuits was planned as a three-phase activity (orientation, conceptualization, and investigation). In another example, a teacher organized a lecture on Foucault's proof of Earth Rotation as a five-phase activity (orientation, conceptualization, investigation, conclusion and discussion) that was further divided into subphases, which resulted in a total number of eight phases in the ILS. The teacher introduced the additional phases of exploration, experimentation and data interpretation as subphases of the investigation phase, which led to the following sequence of phases: orientation, conceptualization, investigation, exploration, experimentation, data interpretation, conclusion, and discussion. From the descriptive statistics, overall, 3.92% of the ILS were planned with less than 5 phases and 35.29% involved more than 5 inquiry phases.

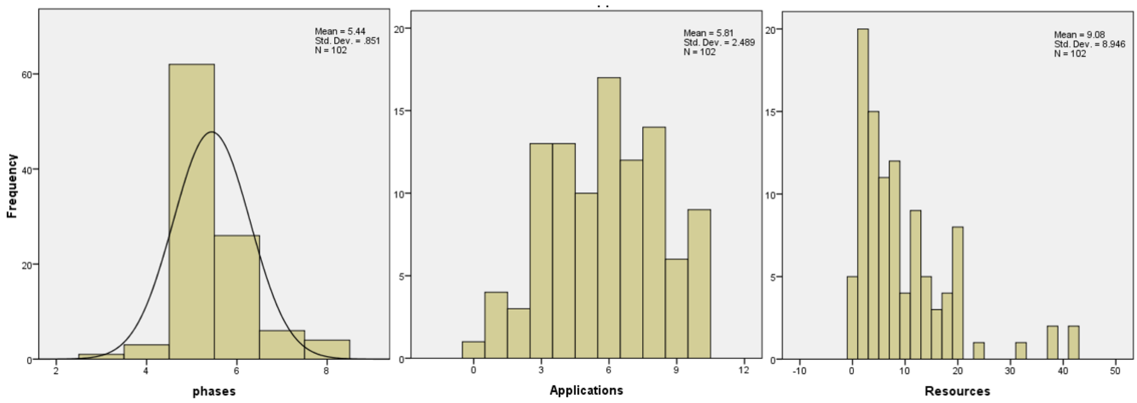


Figure 3.14: Histograms showing the distribution of phases, applications and resources used in the design over the learning spaces.

Descriptive Statistics of Inquiry Learning Spaces

In the design of Go-Lab learning spaces, the teachers are able to choose the applications and resources for their lectures freely. On average, each ILS made use of 15 items: 6 ($Mean = 5.81$, $\sigma = 2.489$, $N = 102$) of them were applications while 9 items ($Mean = 9.08$, $\sigma = 8.946$, $N = 102$) were learning resources of various types (pictures, videos etc.). Figure 3.14 displays the distribution of the number of phases, applications and resources, used over the ILS. The 57% of the ILS integrated more than 5 applications while the 55% of the spaces used more than 5 resources. Out of the 102 learning spaces that we studied, only 27 allowed the use of the vault. The vault allowed the permanent and visible contribution of students to the learning space but it was not widely used as means of promoting reflection or participation.

The descriptive statistics for the ILS are displayed in table 3.1. The analysis of the results shows there is a statistically significant, but weak, correlation between the number of applications used in an ILS and the number of logs recorded during the activity ($\sigma = 0.215$, $p < 0.05$). This indicates that student activity in an ILS increases with the number of available applications. Furthermore, the number of resources correlates significantly but in a negative way with the number of phases ($\sigma = -0.233$, $p < 0.05$). This indicates that teachers tend to distribute the available resources over the various phases.

Table 3.1: Descriptive statistics for the ILS of the study.

	logs	**phases**	**items**	**apps**	**resources**
Average	2761.58	5.44	14.892	5.81	9.08
Min	500	3	1	0	0
Max	16426	8	48	10	42

Clustering

The aforementioned metrics have been used as attributes of the learning spaces in order to cluster the ILS of the data set using a k-means clustering approach. The number of extracted clusters was set to 3 as estimated by a plot-based method (Everitt and Hothorn 2009). The main objective was to find related groups in the data set and discover potential dependencies between factors that describe ILS. The results of the cluster analysis provided one dominating cluster of learning spaces (cluster 2) and two smaller but nonetheless distinctive ones (clusters 1 and 3) shown in figure 3.15. The multivariate clustering presented in this figure displays the first two principal components that explain 73.92% of the point variability. Cluster 1 consisted of 8 out of 102 spaces and cluster 3 consisted of 5 out of 102 spaces. The spaces of these two clusters integrated a vault in their structure and made use of a great number of resources and applications. In particular, Cluster 3 consisted of the learning spaces with the biggest number of resources. The vault is a special space that stores learner-generated artifacts, which are also counted as resources in the calculations. Therefore, this cluster consists of the spaces which follow a more active learning approach that make use of apps that let learners create artifacts. Cluster 2 contained 89 learning spaces. The majority of these spaces did not include a vault in their structure and the number of resources was similar to the number of the applications used. This can be characterized as the normal use of Go-Lab ILS. The cluster analysis does not provide any further indication with respect to the inquiry phases in the learning spaces. However, it can be observed that the existence of a vault has an effect that is depicted on the statistics.

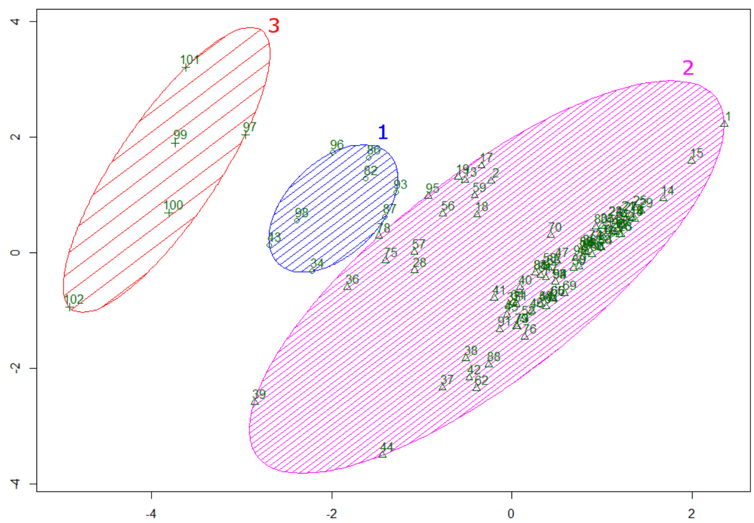


Figure 3.15: Results of the cluster analysis based on the metrics from the descriptive statistics of the learning spaces as presented in table 3.1.

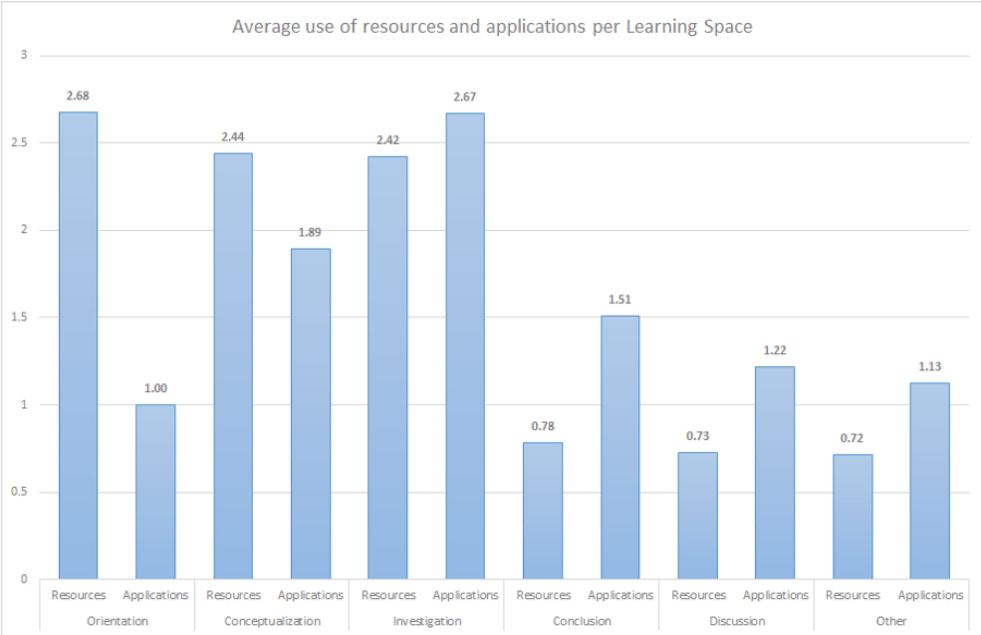


Figure 3.16: Average use of resources and applications per phase in the learning spaces.

In order to gain an insight with respect to the design of inquiry phases, the use of resources and applications within the various phases has been studied. This analysis is based on the hypothesis that different inquiry phases serve different purposes. Therefore, the use of resources and applications should vary depending on the objective of the inquiry phase. Figure 3.16 presents the average number of resources and applications used per inquiry phase. Overall, resources are mostly used in the orientation and conceptualization phases. In all other phases of the recommended IBL model (investigation, conclusion and discussion) as well as in other phases (i.e. phases introduced by teachers) the applications have a higher usage. This is particularly interesting since it indicates a shift into more active learning processes, where the students are encouraged to participate. The teachers do not focus on distributing their own resources around the classroom, but promote the active involvement of students through the use of applications.

Learning Process Sequences

The previous analyses provided an overview of the teachers' specifications of the learning spaces, particularly of how teachers created and edited the ILS. The teacher might or might not change the ordered sequence of recommended phases. Complementary to this, the following examinations focus more on the perspective of the learners and how the spaces have been used actually. This consists of the analysis of the deviations from the designated learning process, which has been specified as subsequent phases in the Go-Lab ILS. Then, a learner might or might not follow this specifications, which has been operationalized for this analysis. Therefore the static number of phases and their respective sequence have been coded for each ILS. The learner might then follow this sequence or deviate from this specification. To measure and quantify this, different parameters that can be linked to deviations ("out-of-order behavior") have been extracted and calculated. We characterize the run-time behavior or trajectory of the learner as a sequence of actions within a model predefined by the teacher. The learning process sequence ("LPS") is then a coding of the (ordered) phase sequences each learner visited subsequently. Considered as an example, a teacher creates an ILS that consists of the three phases: orientation, conceptualization and investigation. For each learner who runs the ILS, a single LPS is created from the action log protocol. All the logs will be aggregated and folded in order to represent the transition between the phases rather than a fine-grained log sequence of tool actions. Within this example, the LPS [0,1,2,1] encodes for a single learner the visiting of the phases in the presented order: orientation, conceptualization, investigation, conceptualization. In this example, the learner jumped back from the investigation to the conceptualization phase. Such sequences deviate from the actual specification by the teacher and might indicate such an out-of-order behavior.

The following measures have been considered: (1) the number of inversions in the learning process sequence, (2) length of the LPS with repetitions, (3) length of the LPS without repetitions, and (4) number of phase omissions. In this context, an inversion is the jump to a non-successive phase. This quantifies the moves through the ILS that are against the natural order of the phases in a teacher's specification, for example, from an experiment back to the conceptualization. Complementary to this, if the length of the LPS without repetitions is smaller than the number of phases specified, at least one phase has been omitted by the learner. Figure 3.17 shows the distribution of LPS lengths (without repetitions) across the static number of phases in the ILS specified by the teacher. Only few learners followed strictly the recommended sequence and omitted at least one phase. With a lower number of phases, it seems to be more likely that learners follow the sequence defined by the teacher. For instance, with a three-phase specification of the ILS, roughly 80% of the learners have a LPS length of 3. Even with the recommended model (5 phases), less than half of the learners follow this specification.

Figure 3.18 is dedicated to the distribution of omissions of inquiry phases on the part of the learners. The right chart relates the number of visited phases and the number of omitted phases -both by the learners- to the number of phases specified by the teacher. It demonstrates a clear trend: with an increase in phases specified goes an increase in the chance of having deviations in the learning process sequences from the design. This is particularly the case for omissions of phases by the learners. The

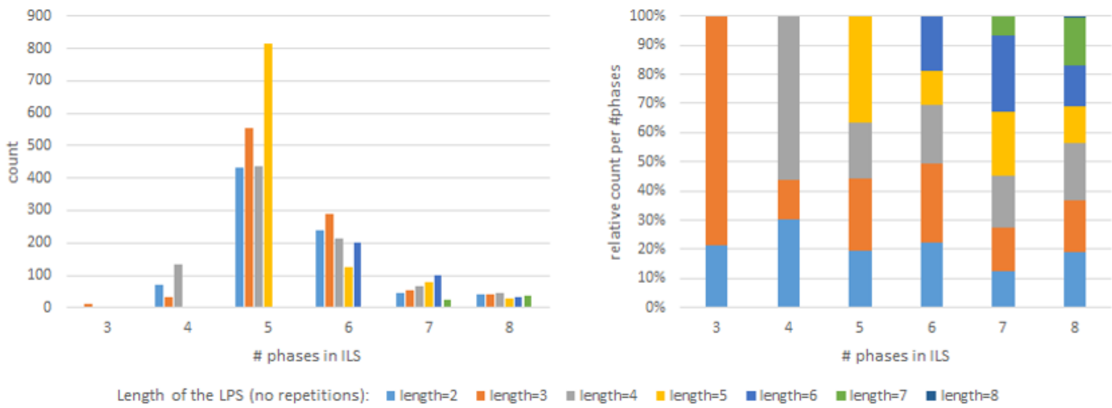


Figure 3.17: Absolute (left) and relative (right) number of LPS of a specific length without repetitions (color coded), aggregated by the number of phases per ILS.

left chart shows a more detailed view about the phases that have been omitted, aggregated by the number of phases specified in the ILS. The phase "other" is a special coding for spaces that do not directly refer to the inquiry cycle. In practice, this has been used for additional information or monitoring. Spaces are named, for example, "dashboard" or "reflection" and are usually for self-reflection, self-monitoring or metacognitive activities such as the planning of the learning.

Contrary to our expectations, the orientation phase is one of the more frequently skipped phases. This might indicate that the learners focus more on the tasks in other phases that are either more motivating or activating – orientation phases usually are typically very general descriptions for the students or a collection of motivational resources. Videos, for example, might have been skipped as a consequence of bad internet connectivity. Conceptualization and investigation phases are rarely omitted, which seems to be obvious with regards to the approach of inquiry-based learning and the fact that both phases usually contain inquiry apps and scaffolds that produce learner-generated content. This can be, for example, an online lab for the experimentation and inquiry apps such as the concept mapper or the hypothesis scratchpad for the conceptualization phase. Another peak is the discussion phase that has been present in spaces with more than four phases on the one hand, but then frequently skipped on the other hand. A possible reason might be that such phases could potentially be carried out to (pure) classroom activities and teachers might not really see the need for ICT support in this phase. Due to typical time or room constraints when dealing with ICT in classrooms, such activities might take place outside the Go-Lab environment and are therefore not reflected in the data captured through the learning analytics infrastructure. In some self-reports from teachers it has been declared that due to bad internet connections the introductory video has been projected and shown in classroom. Practical workarounds like those for weak ICT infrastructure might yield to such singularities and possibly explain some of the examples.

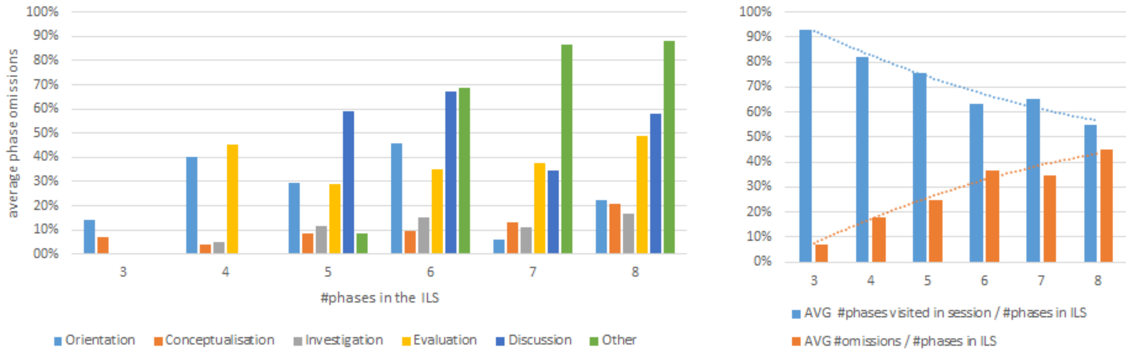


Figure 3.18: Left: distribution of average phase omissions aggregated by the number of phases in ILS. Right: number of phases visited and number of omissions related to the number of phases in the ILS design.

The charts displayed in figure 3.19 emphasize on the inversions in the LPS. The left diagram in this figure presents a distribution of the average number of inversions across different lengths of LPS. The LPS is counted with repetitions in this case. This diagram shows that a proportionality between the learning process sequence length and the average number of inversions is likely. The relative proportionality shows that the pedagogical design, for instance, the number of phases in an ILS plays a subordinate role - in contrast to the actual learning sequence. In this sense, this relation expresses that with the length of the LPS there is a nearly linear increase in the number of inversions, independent of the teachers' specification. A reason for this might be found in a lack of process awareness on the part of the learners. However, such awareness components might play an important role in guiding learners through inquiry activities as stated out in section 2.1.1.

Furthermore, the results obtained from the pedagogical specification of ILS have been compared to the actual run-time behavior. Thus, the comparison served to find deviations between the ILS model (teacher-defined) and the recommendations (by Go-Lab) provided to the teachers regarding the use of the apps and labs. This provides some insights on deviations from the default inquiry model in the authoring perspective. From the run-time perspective, the deviations from the actual learning process sequences of phases can be measured to determine "out-of-order" behavior relative to the pedagogical specification of the teacher. The chart on the right-hand side of figure 3.15 shows the distribution of phase sequences lengths (i.e. the number of sequential phases) that are extracted from the actual learning process models of the learners. It points out that most of the scenarios contain 5 or 6 phases. Additionally, a large number of students have LPS with a length much longer than the recommended model.

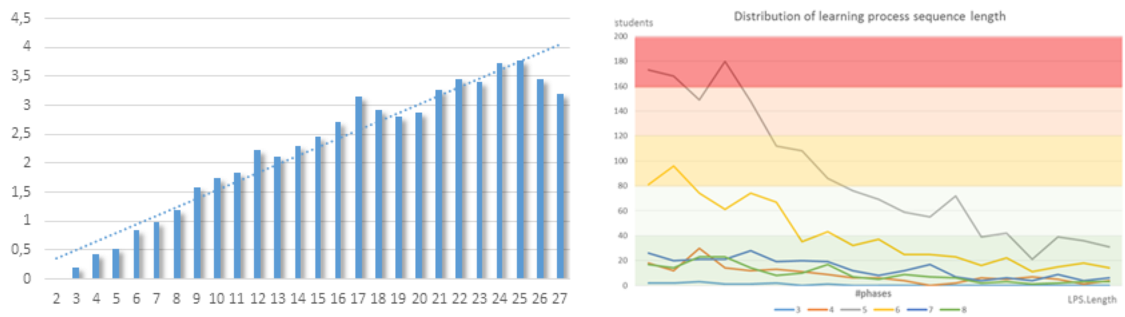


Figure 3.19: Left: Average number of inversions depending on the length of the LPS (domain axis). Right: Number of students depending on the LPS length aggregated by number of phases specified.

The line color in the chart corresponds to the number of phases in the ILS specification, while the color code of the graph areas indicates the number of students.

3.3.3 Discussion

The analytic results of this work point out that (structured) inquiry-based learning induces a shift to learning activities where students are encouraged to actively participate. In 39.22% of the ILS, teachers took advantage of the possibility of customizing the inquiry sequence to their needs. In these spaces, there is a high number of deviations of the learning sequence from the specification. Therefore, appropriate apps would be necessary in order to support teachers and students to regulate and intervene in case of out-of-order behaviors during the learning process. For example, apps for monitoring the learning process might help teachers to be aware of the students' activity and to identify out-of-order behavior based on the metrics used for this work.

Figure 3.20 shows two prototypes of supportive apps for process awareness of teachers and students. Both applications have been developed based on the findings of this work using the learning analytics infrastructure from the Go-Lab project. The visualizations are based on the action logs of the learners that are aggregated using the data warehouse API (cf. section 3.1.2). The app on the left-hand side is a monitoring tool for teachers which provides an overview of the phase sequences of each student. The app allows the teacher to uncover deviations in the sequence of phases as well as in the time spent per phase (e.g., a short time devoted to an experimentation phase in comparison to the recommended value). The statistics presented in the analysis have shown that a lot of students omitted phases during their run of an ILS, which can be monitored with such an application. As the app visualizes real-time data that

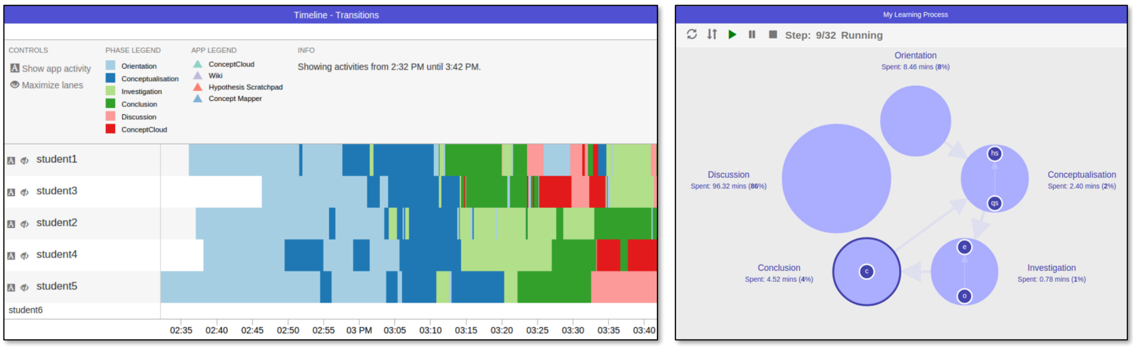


Figure 3.20: Go-Lab apps to provide process awareness for teachers (left) and learners (right).

has been logged into the system, it can be used for on-site monitoring during a classroom activity, which helps the teacher to spot sequences of interest and thus to place a pedagogical intervention in case it is needed.

Complementary to the teachers, students need to be aware of their own learning processes in order to adopt inquiry models in a useful way. To achieve this, cognitive scaffolds to foster process awareness and self-reflection are necessary. The right-hand side of figure 3.20 shows an app suitable for learners to display a detailed learning process sequence. The app displays an interactive learning process visualization, which provides information about the process sequence on two levels (phases and apps) as well as the time spent by the student in each phase. The outer circles (light violet) display the phases and the corresponding phase transitions as connecting arrows. The size of the outer circles indicates the relative time spent in each phase, which is also displayed next to the caption. Inside each phase, the transitions between applications are displayed as the enclosed circles (dark violet). With this visualization, a student can evaluate the deviations of the own learning process from the suggested model and compare the times spent in the different phases. For the purpose of reenactment and reflection, the visualization contains a "replay" functionality to animate the history of transitions.

Guiding learners can be seen as one of the (beneficial) main ingredients of the Go-Lab approach for inquiry-based learning. Inquiry apps to support learners' activities are promoted throughout the Go-Lab portal to be used in inquiry learning spaces. However, there was an upcoming debate about the necessary amount of instructions and guidance mechanisms in inquiry-based learning, pointing out the danger of going for less is better (Kirschner et al. 2006). As a consequence, both the learning design and the learning process need to be aligned to each other. Even though, support tools for monitoring processes can help to facilitate process awareness mechanisms to improve learning, metacognitive or reflective skills. From the analysis of the learning spaces and the distribution of artifacts (apps and resources) that were used in the ILS creation, it was evident that teachers consider the use of more apps while they keep the use of resources limited. Although the average number of applications and resources added to an ILS is similar, the majority of spaces use a minimum number of resources and they tend to increase the number of applications. Furthermore, the use of resources is negatively correlated with the number of inquiry phases. This implies that teachers tend to create ILS without adding a balanced amount of resources (such as external learning materials) to the phases, but to use more apps inside the phases. However, as a limitation to this work, it cannot be observed or captured through this large data analysis whether the teachers used learning materials that are not stored in the platform, for example, printed materials or text books. Maybe, this is also due to the fact that applications lead to an increase in students' activity and therefore applications are perceived as a way to support and encourage students to act and take the initiative. Of course, interactive experiments, inquiry apps or scaffolds provided by Go-Lab are promoted to teachers and might act as a selling point. Differences in the design of the learning spaces might have many reasons. Starting from the attractiveness of the provided applications, experiments and public examples of ILS in Go-Lab, this might also be rooted the expert level of the teachers in using ICT or facilitating technology-enhanced learning. Such aspects need deeper investigations that take the user's perspective into account.

3.4 Conclusion

The work presented in this chapter is one of the first attempts to describe a general learning analytics infrastructure that can be adapted to a wide range of scenarios. In other analytics fields such as business analytics those infrastructures are already quite elaborated. In the relatively novel and emerging field of learning analytics such general architectures are starting out. With this work we aimed to draw attention to the challenges and requirements of general approaches for analytics infrastructures in web-based learning environments and proposed our solution as part of the Go-Lab environment. The backend components of the infrastructure are implemented as a multi-agent system, in which agents communicate implicitly through a shared workspace.

Additionally, it features a flexible architecture for creating and deploying micro services and portable analytics apps. The platform used for the creation of analytics apps offers a visual language to specify the workflows. This allows for a flexible integration of new functionality, for example new analysis algorithms or visualizations. Well-defined data formats, protocols and interfaces enable communication channels for action logging, feedback mechanisms and data access from analytics tools that use the Go-Lab learning analytics server. In Go-Lab, the analytics workbench has been used to create the first prototypes of learning analytics apps that are presented in section 3.2.3. The architecture has been extended to build static code bundles based on the multi-agent that could be integrated on a deeper level into the Go-Lab learning analytics infrastructure. The presented infrastructure has been employed to analyze the activity of learners and teachers in the context of inquiry-based learning with online experimentation in Go-Lab. Our study is based on more than 100 inquiry learning spaces and combines heterogeneous data sources with various filters, metrics and indicators. The results show trends in the design of teachers' ILS, e.g., in terms of number of phases, apps and resources per phase. The models induced by teachers can be further evaluated taking into consideration parameters such as the functional type of apps regarding the inquiry process (e.g., apps for reflection, metacognition, etc.) and the concrete types of resources a learning space is composed of.

Additionally, the results point out that although most of the teachers adopt the recommended inquiry model, a significant number adapts it according to the needs imposed by the learning context. Noteworthy that, students often do not follow the teachers' model. These deviations might originate from a lack of process awareness that could be overcome through appropriate scaffolds. The detection of "out-of-order" behavior is a complex task and possibly includes a variety of indicators. To support process awareness, we have proposed prototype applications for students and teachers. In the future work we will validate our prototypes regarding the interpretability of the rich representations of learning processes. Other metrics such as dynamic time warping could be useful to measure the costs to match an actual learning sequence into a sequence specification. Such a metric will involve different kinds of deviations (inversions, insertions, repetitions, etc.). As a continuation, first participatory design studies with teachers will show their usefulness and provide some further input on the indicators. This involves particularly a more integrated usage of the metrics and indicators in rich representations that go further than a simple mirroring of values. From the perspective of learners, this can be useful to support self-reflection and metacognition, fostering 21st century learning skills as well as helping teachers to support such competencies.

In summary, the architecture fulfills the requirements for creating and serving flexible and portable learning analytics applications. However, the first data-driven observation of the portal-usage (learning and teaching) using the architecture show that there is a need for process and cognitive awareness. Knowledge management approaches have the potential to support both aspects. We employ specific learning scenarios (individual and group work) in the context of inquiry-based learning in Go-Lab that can make use of learners knowledge. Therefore, the next chapter describes the conceptualization and implementation of knowledge management approaches, particularly based on knowledge diversity and a shared group knowledge model.

4 Knowledge Management Approaches

> "Diversity may appear to be a straightforward concept which can be quickly and painlessly measured. This is because most people have a ready intuitive grasp of what is meant by diversity and have little difficulty in accepting, say, that tropical rain forests are more diverse than temperate woodlands or that there is a high diversity of organisms in coral reefs. Yet diversity is rather like an optical illusion." (Magurran 1988)

The construction of knowledge, both in an individual and a social context, has shown its importance from the history of inquiry-based learning up to modern approaches (compare section 2.2). Therefore, we identified knowledge (co-)construction as one of the key aspects of the foundation of this work. The field of computer supported collaborative learning defines one of the key challenges as the definition of preconditions for the successful orchestration of collaborative learning scenarios or knowledge building. Particularly for the formation of learning groups, there has been a debate about different characteristics of learners in each group and how to benefit from differences. The suggestion to form heterogeneous learning groups became quite popular in the line of CSCL research. The apparent consensus about the benefits of heterogeneous groupings influenced the area of automatic group formation in a way that many systems define heterogeneity as a criterion or goal for the groupings. However, an evaluation of the group performance is highly dependent on the quality measure of the learning outcome. For example, when using learning gain per group for an evaluation, studies suggest that heterogeneous groups perform better (see section 5). But on an individual level, particularly weaker students benefit from such group constellations. The idea of heterogeneous groupings leads to a political discussion about inclusiveness and fairness of learning as well. Particularly ideals of internal differentiation of learner groups foresee that everyone should benefit from learning, including students with higher skills. We argue from an inclusive point of view that there should be possible groupings of learners that lead to a learning gain for everyone. A key towards achieving this goal is to advance from heterogeneity in skills to a successful managing of knowledge diversity, where the group dynamics and learning processes are initialized through utilizing and facilitating the differences in knowledge positively. We move from stigmatizing learners ("low achievers") to exploit diversity. This work focuses on characterizing the diversity of cognitive aspects on the part of the learners, namely through measuring the knowledge diversity of a cohort. The measurement of diversity, as highlighted in the introductory quote by Magurran (1988), is a challenging task as the term diversity has a weak operationalization. This work presents case studies and exploratory studies where knowledge diversity of learning groups has been exploited in order to support learning through "semantic group formation" or through visualizations of group knowledge models ("concept cloud"). Stoyanov et al. (2017) used a group concept map "as a group's common cognitive construct can consolidate individual differences and serves as a tool for managing diversity in groups of participants." This thesis presents a more generic approach, as it defines an operationalized model of knowledge and diversity in conjunction with an open learner model that is suitable for inquiry-based science education, where learners express their knowledge through different and heterogeneous representations and artifacts.

Section 4.1 summarizes the use of diversity across different disciplines. Traditionally, there has been a seamless transition and sometimes a synonymous usage of *heterogeneity* and *diversity*. The next section spots the differences and integrates it into a working definition of diversity. Section 4.1.2 reflects on different knowledge representations that can be used in order to operationalize the term knowledge. The subsequent section, 4.1.3, focuses on an operationalization of knowledge and knowledge diversity. This can be used to define computational models for knowledge, which can be integrated into inquiry-based learning scenarios.

4.1 Facilitating Knowledge Diversity

There is a current movement in different disciplines towards an increased diversity as it seems to be promising in many applications which range from social over workplace to educational settings. While social and educational sciences investigated in cultural and ethnic aspects of diversity, the research of this work is focused on knowledge diversity. Knowledge (co-)construction has been highlighted as one of the key aspects in CSCL research. In order to create learning scenarios that benefit from knowledge diversity, exploratory work is needed that is driven by the question of how we can operationalize and facilitate knowledge diversity.

4.1.1 Heterogeneity and Diversity of Learning Groups

The meaning and use of the term heterogeneity is manifold and varies in different disciplines and domains such as chemistry or computer science. The work of this dissertation is situated in the field of technology enhanced learning, therefore it focuses on learning contexts. There is no consensus in the research of TEL and CSCL about the term *knowledge diversity*. However, in these communities there is a (traditional) debate about heterogeneity in learning, and the terms *diversity* and *heterogeneity* are often used interchangeably. This section aims to bring clarity into the debate by deriving a working definition of knowledge diversity from the literature in the contexts of diversity management and learning.

The constructs of heterogeneity and diversity also appear in pedagogical areas such as the field of group learning. Particularly to characterize the goal of the grouping, heterogeneity and homogeneity can be defined. Such groupings can have advantages or disadvantages, depending on the pedagogical goal or specification (Barkley et al. 2014). When Dillenbourg (1999) defined conditions for collaborative learning in his work, he suggested to take heterogeneity into account. As a caveat, he distinguished symmetry and heterogeneity: "two learners may have a similar degree of expertise but different viewpoints of the task". This view on heterogeneity shows how collaborative learning is connected to discourse. Trimbur (1989) pinpoints heterogeneity to consensus and difference in group discussions as important characteristics of collaborative learning. However, Weinberger and Fischer (2006) criticize the concept of heterogeneity regarding group learning conditions: "At best, only some learners may benefit from knowledge co-construction scenarios while others are left behind. CSCL may contribute to a more homogeneous participation, e.g., by representing the discourse history on a discussion board." This emphasizes the importance of finding more inclusive characteristics to group learners. Furthermore, it shows the potential of open learner models in the context of CSCL and knowledge co-construction, which frame the research on inquiry-based learning.

Diversity is a term that has been used a lot in different fields, encompassing research, political discussions, organizations and economy. One would expect a consensus about the definitions and meanings of diversity or diversity management. However, Harrison and Klein (2007a) state out that "diversity has often been studied in an indeterminate manner; (...) the substantive or constitutive definition of diversity often is not clearly specified." Although the introductory quote by Magurran (1988) is from another discipline, it shows the difficulty of finding a definition. Depending on the variable, the extent of how diverse two individuals are, can even have more subordinate aspects, that span their own spaces over the set of characteristics.

Diversity with respect to learning is usually referring to a social construct. Social psychologist Aronson designed the jigsaw technique originally as a teaching method in order to force integration of racially and ethnically diverse groups and to overcome social barriers that are a consequence of this diversity (Aronson 2002; Ziegler 1981). A situation in which learning should take place usually underlies certain conditions. In this case, ethnicity as one possibility of diversity can influence the learning. With the jigsaw method, Aronson tried to manage the diversity and use it in a way to design learning which overcomes the barriers in order to create a beneficial learning situation for all the participants.

In the field of cognitive and administrative sciences, there has been a discourse about cognitive versus demographic diversity. Research has shown the greater impact of cognitive diversity on the decision-making processes and on the team performance (Jehn et al. 1999; Williams and O'Reilly III 1998; Miller et al. 1998; Schilpzand 2010). Knowledge can be seen as one of the possible manifestations of cognition. Therefore, we investigate in one facet of cognitive diversity, the knowledge diversity.

Ryan (2006) highlights the differences in dimension of the term diversity, and that all categories of differences have to be considered: "race/ethnicity constitutes only one dimension of difference. Many other kinds of diversity pervade our schools and communities." Besides obvious or demographic aspects, this can be transferred to the context of (creative) problem solving or to the level of decision making. Diversity can also mean having "diverse viewpoints on the problem" (Falk and Johnson 1977). Diverse viewpoints can exist, for example, because of differences in knowledge. The field of knowledge management has investigated in the processes of knowledge building and emphasizes the importance of diversity. Close to the idea of the aforementioned Jigsaw approach is the idea of having dedicated experts on specific topics. "The term 'expertise diversity' refers to differences in the knowledge and skill domains in which members of a group are specialized as a result of their work experience and education" (van der Vegt and Bunderson 2005). To compose learning scenarios that use expertise in a certain topic in a way that learners are grouped regarding knowledge complementarity has shown success (see section 2.3.3). Therefore, we state out that an operationalization of diversity needs to take the idea of knowledge complementarity into account. Learners should be maximally diverse regarding their knowledge, if their knowledge is complementary. Figure 4.1 shows a comparison of knowledge distributions of homogeneous (high and low), heterogeneous and complementary groups regarding the knowledge of learning partners in dyadic groups. This can be seen as a baseline to distinguish the different connotations of heterogeneity and diversity in order to create a terminology of the work in this thesis.

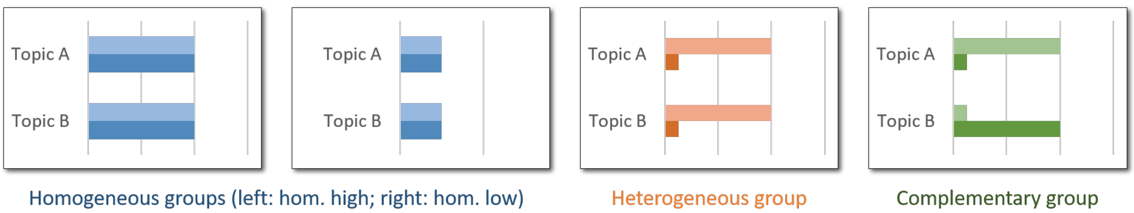


Figure 4.1: Comparison of knowledge distributions in homogeneous, heterogeneous and complementary groups regarding two topics A and B of two learning partners (dyadic groups).

In this sense, diversity can be seen as a property that is given for a certain learning group. It cannot be seen as an output parameter of learning, but it defines the input for groupings. Therefore, this work proposes the use of diversity in a given setting and facilitates it in order to improve learning. Particularly the field of diversity management does not only take individual differences as given and accepts them, but goes further towards utilizing these differences in a constructive way to contribute to the success of a business. Cox (1994) emphasizes the benefits of managing diversity:

> „Planning and implementing organizational systems and practices to manage people so that the potential advantages of diversity are maximized while its potential disadvantages are minimized."

This work by Cox is contextualized in the research of cultural diversity in workplaces. Of course, the aspects to be included in diversity do not only exceed cultural aspects. This concept can be applied to knowledge diversity as well. Cox highlighted the dynamics of diversity: rather than attempting to have a separation of concerns, he described an integrated approach of the different aspects and dimensions (Cox 1994). In research of diversity in the organizational and academic fields, knowledge was often mentioned as one of the diversity dimensions, but has not been discussed and researched in depth so far. Andresen (2007) described a first step in using knowledge as a dimension of diversity. However, social diversity in classrooms is a different issue and not part of the main work presented in this thesis. While the traditional work in the field of diversity in organizations claims that an increased knowledge is the consequence of utilizing diversity (Olsen and Martins 2012), we facilitate knowledge as a component or dimension of diversity. It is important to spot the weaknesses in handling heterogeneity in educational contexts as it has been done, and furthermore move forward to a notion of diversity, which makes use of the differences in order to improve a certain outcome. Diversity management is the art of optimizing situatively the heterogeneity and homogeneity to achieve set goals. Therefore we manage knowledge diversity by creating learning scenarios, which will benefit from the diversity. This is in line with the notion of knowledge complementarity, which can be a baseline for collaborative learning (compare section 2.3.3). We define knowledge diversity as a quantified difference in knowledge, a measure for the degree of complementarity of the knowledge of two learners.

4.1.2 Knowledge Representations

Anderson (2013) wrote in his general theory of cognition ACT-R about two different types of knowledge: declarative and procedural. Declarative knowledge (e.g., learning the rules of how to play football) comes first, and procedural knowledge comes after (e.g., putting those facts and rules in practice to gain football skills). One of the key

Figure 4.2: Left: "If it thunders, it lightens.", right: "it thunders without lightning". Compare Sowa (2011).

aspects in dealing with knowledge is to find appropriate representations. Depending on the distinctive type of knowledge, it can be modeled in different ways: declarative knowledge can be represented as a network of concepts, while procedural knowledge can be better represented as a set of rules (Ramirez and Valdes 2012). Focusing on IBL applications like Go-Lab, where the learners actively create artifacts such as concept maps, hypotheses, or texts, they externalize their knowledge about a certain domain (Erkens et al. 2016a; O'Donnell et al. 2002). Such representations can be reduced to a textual representation and used as a model to express declarative knowledge without major loss of information in the context of inquiry-based learning (Manske and Hoppe 2016).

However, textual representations of knowledge are unstructured data in the sense that they do not have a predefined underlying model or are not organized in a predefined manner. To make use of knowledge through automated methods, irregularities and ambiguities need to be eliminated. Apart from the creation of a model for the (declarative) knowledge externalized in learner-generated artifacts, one of the key challenges for the field of Knowledge Representation (KR) is to perform automated reasoning (to make inferences, for example) on knowledge. This work focuses on the modeling of knowledge, and not on automated reasoning. Although certain operations on the final model (measures of diversity) were performed, this approach should not be considered as "reasoning" as it would be the case for knowledge representation models in artificial intelligence for building intelligent systems (Russell and Norvig 2016). There are a variety of knowledge representation models such as semantic networks, rules, logic, frames, and ontologies.

Semantic networks are a starting point to build a "network of concepts and topics", which is a common knowledge representation method for declarative and object-oriented knowledge (Helbig 2006). Charles Peirce proposed in 1869 existential graphs, a simple system and syntax for first-order logics, which also models simple semantic structures, e.g. relations between concepts. Figure 4.2 shows relations between the concepts thunder and lightning modeled in a visual form of existential graphs (Sowa 2011). Existential graphs share many concepts with semantic networks, but are slightly different. Semantic networks are diagrams with nodes and links between them, representing logical sentences (Russell and Norvig 2016). The idea behind se-

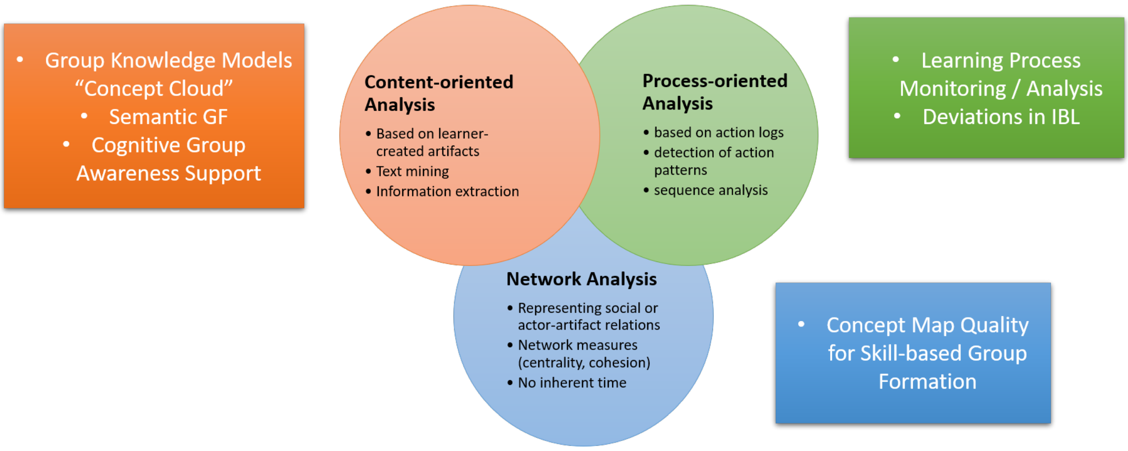


Figure 4.3: The trinity of methods for Learning Analytics in the context of this thesis.

mantic networks is that we can express with them the "taxonomic structure of categories of concepts and the relations between them" (Grimm et al. 2007). In this notion, semantic networks are useful to represent the network consisting of concepts and topics. The topics and subtopics can be seen as categories which encode a type of taxonomy with concepts as leaf nodes, and the "is-a" relationship between the nodes. Text networks are networks extracted from texts using a "network text analysis" approach (Hecking and Hoppe 2015). Approaches like this demonstrate the possibility to create semantic representations of knowledge by using automated, computational models. This is further explained in the context of semantic extraction in section 4.2.3. In summary, this can be seen as a technological and methodological foundation of this work.

4.1.3 Operationalizing Knowledge Diversity

Miel (1952) and Thelen (1954) point out the potential of cooperative learning in heterogeneous conditions – assuming that in a cooperative classroom everyone takes part, and that everyone's contribution to the group is valuable. Taking into account that diversity is a property of cohorts and the individual differences (based on the diversity dimensions) can be utilized in order to contribute to the positive output of the group, namely learning success, we argue that knowledge diversity plays an important role in CSCL. While the notion of diversity comes from the field of inclusive pedagogy or cultural, social or organizational diversity, the term is neither well-defined nor operationalized in the field of CSCL. However, as part of this thesis, a working definition of knowledge diversity is given, and an operationalization in the form of measures of knowledge diversity is proposed. Using these measures, knowledge diversity is quantified in order to specify desired output settings that can serve as useful constraints or as criteria for automatic group formation. In addition to rendering such a model to group characteristics, this can be used for making knowledge of learners visible (Manske and Hoppe 2016), which is one of the key challenges in CSCL (Lipponen 1999). The term "diversity" is usually referred by many authors under different but related nouns such as "heterogeneity", "dissimilarity", and "dispersion". This makes it difficult to understand its precise meaning. Taking a broad definition of diversity, we aim to clarify the "meaning of differences within objects of a unit" (Harrison and Klein 2007a). Therefore, we define diversity as "the distribution of differences among the members [learners] of a unit [learning group] with respect to a common attribute [knowledge]..." (Harrison and Klein 2007a). Tsui et al. (1992) tried to operationalize the calculation of work group diversity by computing individual distances for each variable or dimension, regardless of the underlying metric. More formalized, let S denote a collection of objects. In this work, the elements of S are learners, but in other contexts they can be as diverse as required. Taking two members of S, i and j (learners), we assume that a distance or dissimilarity measure between them $d(i,j)$ is given (Weitzman 1992). This distance measure satisfies:

$d(i,j) >= 0$, $d(i,i) = 0$, $d(i,j) = d(j,i)$.

The choice of concrete distance measure is tied to a particular knowledge model (for instance, distance-based measures on ontologies). For any pair of elements (there are $n \cdot (n-1)/2$ of such pair elements) that belong to S, we have a distance measure (non-negative and symmetric) that can serve as a metric to express the dissimilarities between the pair i and j. This distance measure forms the primitive for quantifying the diversity of elements (learners) of a given set (learning group) S. It will allow us to quantify the diversity between any pair of learners in terms of their knowledge about a certain topic. A distance Matrix contains all pairwise distances of the whole group of learners.

Depending on the knowledge representation the implementation of a distance measure $d(i,j)$ can vary. The vast majority of the literature talks about diversity in terms of dissimilarity alluding the intuitive inverse relationship between diversity and similarity: the more dissimilar two objects are, the more diverse they are. Certainly, this is only a valid approach if the notion of (dis)similarity is more easily accessible than the notion of diversity (Nehring and Puppe 2002). Many similarity measures have been proposed, such as those based on information content (Resnik 1999), distance-based measurements (Lee et al. 1993; Rada et al. 1989), Dice and cosine coefficients (Frakes and Baeza-Yates 1992). Each of them is tied to a particular model and thus renders certain requirements. For instance, distance-based measures of concept similarity assume that the domain is represented as a network, while Dice and cosine coefficients are applicable only when the objects are represented as numerical feature vectors (Lin 1998). Based on this operationalization of diversity, we can calculate and assess a diversity score to a cohort or grouping using a certain diversity measure. Figure 4.4 visu-

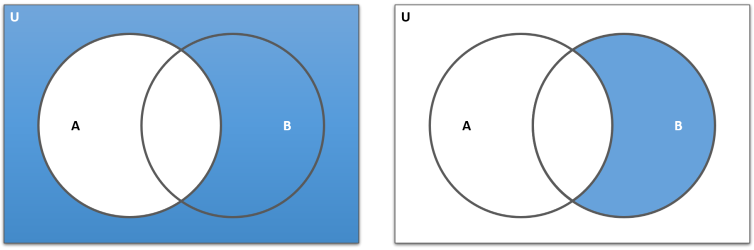


Figure 4.4: Left: absolute complement of A. Right: relative complement of A in B.

alizes the absolute and the relative complement of sets. This outlines, how the notion of complementarity can be used within a set-based knowledge representation, which has been facilitated in section 4.4.2 for the diversity measure.

4.1.4 Learner Models for Representing Knowledge Diversity

We operationalized the diversity regardless of the specific representation of knowledge. We derived a formalization of diversity by calculating distances between individuals in a given space. According to the knowledge modeling approaches mentioned before (see section 4.1.2), we set this to the *knowledge space*. The knowledge space is -in an overlay model- the set of all knowledge items. The knowledge of a learner is then defined as a subset of this space. We also chose an optimistic approach, in which we utilize the paradigms and work by Hoppe and Ploetzner (1999), where we set *knowing* as a binary predicate, which models if a student knows something. We assume that when a learner writes about a specific key concept, that s/he has a certain confidence about the topic and that this topic is part of his or her knowledge. We are aware, that this assumption has limitations and is not representative for all real cases, where learners write about concepts that they misunderstood. However, for the challenging task of inquiry, we assume that learners' confidence is 'high enough' and that they only manage to proceed when they have a certain degree of knowledge in the mentioned topics. Chapter 7 presents the investigation of text mining quality and describes how this approximates the knowledge of the learners.

Learners and their respective knowledge representations can be also modeled in different ways, as mentioned in section 2.3.3. However, we aim for scalability in different ways. The field of intelligent tutoring systems has shown how to model cognitive states of learners in a very accurate way. Usually, these approaches involve a lot of efforts to predefine all possible states. In open inquiry-learning scenarios like in Go-Lab, it is impossible to determine all possible outcomes. Therefore, we employ such an optimistic method that renders knowledge as knowledge items mentioned by learners. We relate that to inquiry-based learning as these knowledge items are key concepts in the science domains for the specific scenarios. Therefore, we can define a learner model as an overlay, similar to the aforementioned approach. Simplified, a learner in this model is then represented by a set of key concepts, s/he has used in the learning scenario. These key concepts will be extracted automatically through semantic methods, which is described in the following section. The knowledge space is then relative to the science domain and the knowledge of all learners in the cohort. It can be defined as the union of all learner models. Reference models can be integrated as well, but serve more as a filter in order to restrict the knowledge space, rendering all other items outside the space as irrelevant.

4.2 Shared Group Knowledge Models

The previous sections in this thesis outlined an operationalization of knowledge diversity. In inquiry-based science education as promoted by Go-Lab, learners actively participate in scientific experiments and create artifacts throughout the process as externalization of their knowledge structures (O'Donnell et al. 2002). In this section, a conceptual framework will be created to transfer the different learner-generated artifacts into knowledge items and thus into measurable knowledge diversity. The conceptual framework aims to promote shared group knowledge models as a specific type of learner models for a whole learning group. Such models can be used to support learners and teachers in (self-)reflective monitoring processes or in collaborative learning.

While it is not a novelty in learning sciences to create conditions for learning which facilitate knowledge, the use of computational methods from learning analytics and semantic technologies is relatively new and advances the knowledge approaches known to this field. While traditional approaches demand the detailed assessment of artifacts, observation of individual work and performance, this work makes use of content analysis techniques, particularly using semantic extraction of knowledge from learner-generated content in order to create knowledge models. Specific for inquiry-based learning is the use of inquiry apps and scaffolds (cf. section 2.1.2) which help learners to externalize and structure mental models, for example in concept maps. Those artifacts carry information that can be automatically processed using computational methods of learning analytics that help extracting semantic information. Thus, knowledge structures of the learners can be explicitly processed and transformed into computational models that support learners and teachers. This computational approach is contrary to early learning analytics, which focused on performance- and activity-oriented data that has been processed using methods of descriptive statistics. One of the examples of such an application is the student activity meter (Govaerts et al. 2012) which visualizes the level of activity of learners over time. Although such applications provide valuable insights, such information is difficult to transform into actionable results or interventions if they are not directly coupled to learning design. As a caveat, learners are usually novices in the interpretation of such complex visualizations.

Besides the challenges that come with the difficulty of interpreting, such statistics mainly focus on supporting the ex-post analysis phase on the part of the teachers. Many guidance mechanisms targeting students are often scaffolds on a micro-level, for example, prompts inside a specific tool (cf. section 2.1.1). Such mechanisms demand a high degree of adaptation and encode very narrow domain knowledge - or match simple and hard-coded interaction patterns. In this sense, it takes a lot of efforts to adapt systems to effectively provide scaffolding. However, the Go-Lab approach calls for more general mechanisms to support students and teachers in order to achieve a large-scale implementation in schools. In contrast to the use of micro-scaffolds, this workflow aims to enforce critical thinking, to initialize reflective processes and to use representations of knowledge structures that are connected to the actual learner productions and thus to the learning outcomes. The importance of these scaffolds has been pointed out for the engagement (Leeman-Munk et al. 2014) and the support (de Jong et al. 2014) of the learner as well as for fostering critical thinking and the development of 21st century skills (Wheeler et al. 2008). In addition to this framework, applications that make use of the shared group knowledge model will be demonstrated in the next sections. In the following chapters 6 and 7, the evaluations of these applications are presented.

4.2.1 Conceptual Framework for Group Knowledge Models

The underlying conceptual framework for the creation of shared group knowledge models is presented in this section. The framework is organized in layers as shown in figure 4.5. The layer on top of this framework represents the pedagogical model of inquiry-based science education that is orchestrated in classrooms using ICT. This layered model bridges from the pedagogy over to technology, particularly using semantic extraction and integration. The arrows show the information flow that connects the learner productions with analytics leading to the shared group knowledge model ("Concept Cloud Data Model"). In such inquiry-based learning scenarios promoted by Go-Lab, students actively create artifacts in the subsequent inquiry phases. The phases prestructure the inquiry activities in terms of the inquiry process and by providing specific inquiry apps and cognitive scaffolds to support learning activities and processes. Concept mapping is a typical activity in the conceptualization phase, while the hypothesis scratchpad supports learners in asking questions (cf. section 2.1.2). Text editors and wiki tools are used across all phases for note taking, documenting of observations or writing of conclusions. Most of the learner-generated content that has been created using these apps is text-based or in a format that can be easily reduced to a textual representation without a major loss of information. Within the previously described optimistic approach for the open learner model, concepts contribute most to this model, although, edges in concept maps show important relations between (scientific) concepts. However, as a rule of a common denominator of the different artifact types, structural information such as the relations is withdrawn. The third layer facilitates format- and artifact-type specific concept extractors. Each artifact consists of a set of knowledge items that refer to specific concepts in the domain of the inquiry learning space. These concepts are extracted from the artifacts and added to each learner model, which is a collection of all concepts used in all phases by a student. Thus, such a learner model contains all knowledge items of a

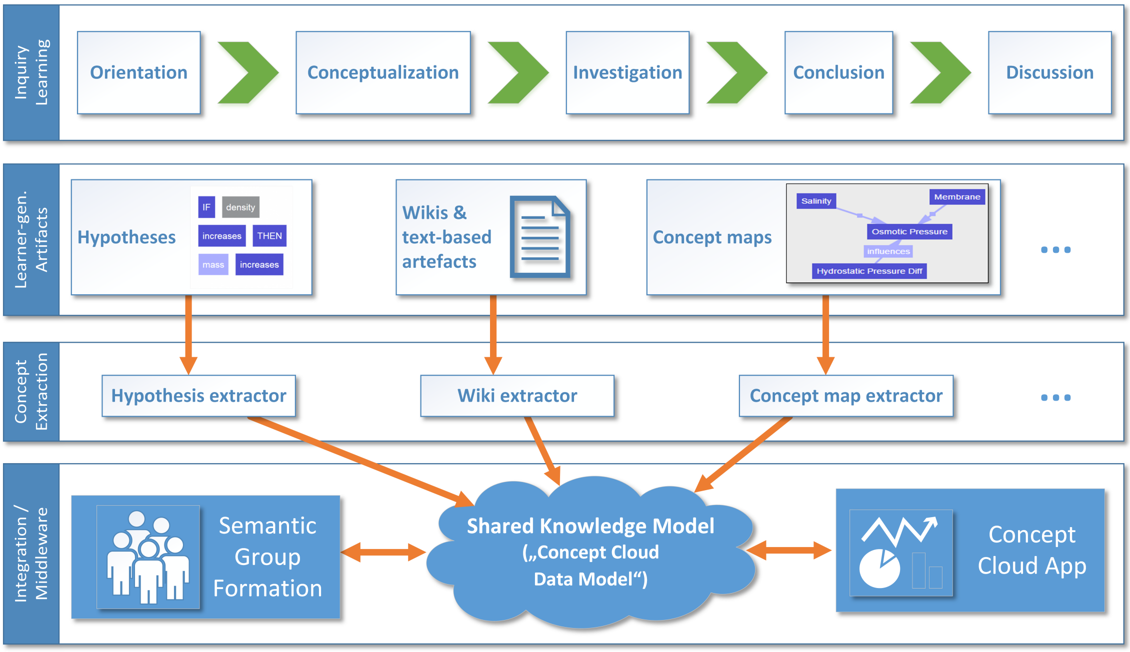


Figure 4.5: Conceptual framework for shared group knowledge models.

student that are relevant to the respective science domain. The aggregation of learners' knowledge models is called a shared group knowledge model, which serves as a basis for further applications, for example to form learning groups based on (diverse) group knowledge ("semantic group formation") or to display the group knowledge as a cognitive scaffold ("concept cloud app").

4.2.2 Architectural Approach for Shared Knowledge Models

The concept cloud framework comprises two main parts: (a) the client framework to create applications, and (b) the concept cloud server, which provides a REST endpoint to request the concept cloud data model for given resources (cf. figure 4.6).

The server provides a simple REST API for all client applications. The main usage and semantics of the web service interface is to send learner-generated content, the artifacts form the learning environment, as well as (optional) configuration parameters to the REST endpoint. The simple semantics of the API (send contextualized artifacts, retrieve the concept cloud data model) masks a complex and distributed backend as a composition of different text analytics services following the principle convenience over configuration. For different applications and artifact types preconfigured extractors and service connectors exist. Each artifact will be processed by a specific extractor

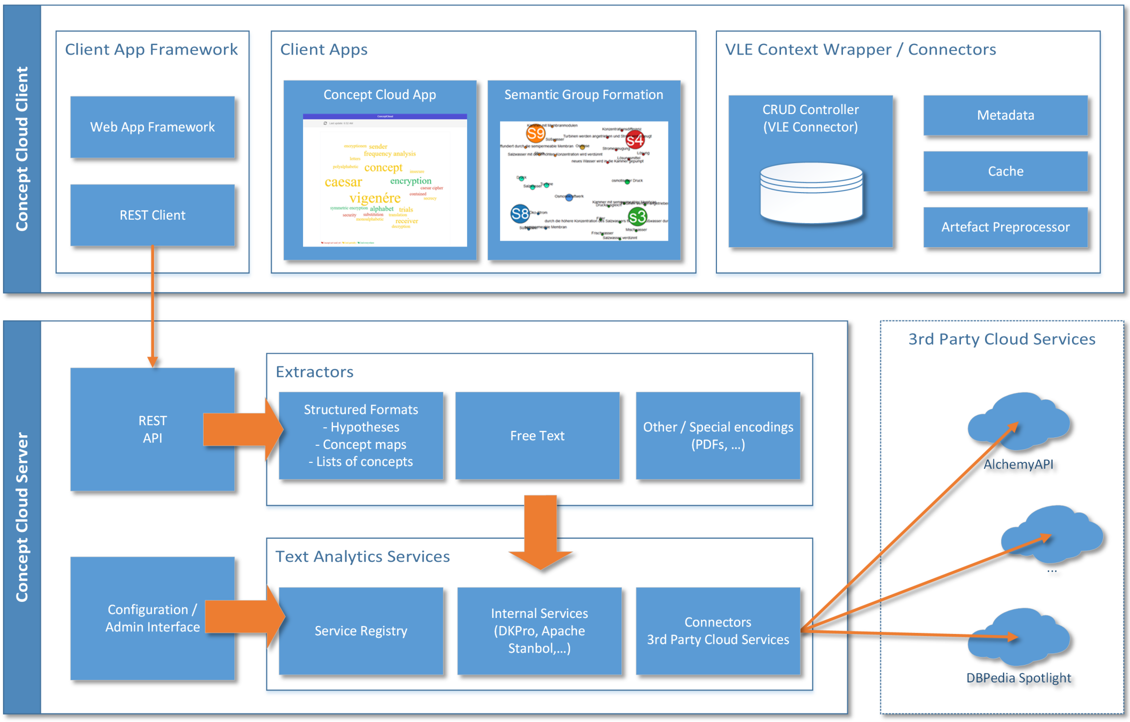


Figure 4.6: Concept cloud architecture. The main parts are (a) the concept cloud client, and (b) concept cloud server with a connector for third party cloud services.

which facilitates a set of text analytics services, to produce a list of key concepts. All concepts are aggregated into the data model in the response of the REST call.

The client framework consists of necessary adapters to be embedded into an already existing infrastructure such as the Go-Lab learning analytics server. The VLE connector is a set of client APIs to integrate within a virtual learning environment ("VLE"). The connector is responsible for fetching artifacts from the corresponding backend, for example Graasp in the case of Go-Lab. The main functions of the VLE connector are to (1) connect to the artifact storage, (2) to retrieve and process metadata, (3) to preprocess artifacts, and (4) to provide caching. This is mapped to a CRUD controller, which performs create, read, update and delete operations on the artifact storage. In addition to the CRUD operations to interact with the artifact storage, parsers are needed to understand metadata and to extract content in the form of basic content types (hypotheses, concept maps, lists of terms, free texts). This encompasses, for example, a mapping from a hypothesis to a list of terms. This content is then delivered to the server, which further processes the content and extracts the semantic information using integrated extractors or third-party services.

4.2.3 Data Model and Extractors

The conceptual model defines a processing chain in order to create a shared group knowledge model that can be further used by applications to facilitate and manage knowledge diversity, such as the semantic group formation or the concept cloud app that visualizes this shared group knowledge model. To achieve this, each learner-generated artifact is processed using semantic technologies. However, different artifact types (e.g., concept map, text) have different technical formats and need their own mechanisms for the extraction of relevant concepts. The general mechanism of the concept extraction can be outlined as the following: each artifact will be mapped from the own technical format to an intermediate representation that is readable for the extractors, particularly a simple text (without markup or metadata) or a token list. Artifact types with a low structure (texts, PDFs) are mapped to plain text that can be directly submitted to the semantic extractors. To make use of the content, markup and metadata is withdrawn, particularly from PDF files. Highly structured artifacts that usually contain a predefined set of operators or terms (concept maps, hypotheses, experimental designs) are transformed into the token lists after removing the structural aspects (i.e., edges, relations, quantifiers, reasoning). For example, after the removal of those structures, in concept maps the list of nodes will be kept, and in hypotheses the list of parameters will remain. Then, the intermediate format will be submitted to the semantic extractor that further reduces the input to a list of concepts. In summary, this is a mapping from the artifact type to a concept list that represent the artifact's inherent knowledge items. In a next step, all artifacts from all students are aggregated into a unified data model that contains all concepts and acts as the shared group knowledge model. Although the mapping leads to a sacrifice of structural aspects as the relations in concept maps, the aggregation can be considered as a common denominator that represents and encodes a majority of declarative knowledge. This unification is consistent with the premise of an optimistic learner modeling approach that has been presented as a foundation of this operationalization. In addition, relations can't be displayed or represented in the loose structure of a tag cloud. For the semantic extraction of concepts from the text-based artifacts (i.e., the intermediate format), the following semantic technologies have been employed: (1) DBpedia Spotlight (Mendes et al. 2011), (2) AlchemyAPI (Turian 2013), and (3) Text2Network (Hecking and Hoppe 2015). The characteristics of the different extractors are described in the following subsections.

DBpedia Spotlight

The concept extraction approach using DBpedia Spotlight targets to identify domain vocabulary and specific terminology. DBpedia Spotlight uses an ontology as a knowledge source. Unstructured texts are matched with the underlying ontology using natural language processing technologies in order to annotate and extract concepts with entities from the ontology. This knowledge source enables a context-aware extraction of terms to an extent which it enables disambiguation of terms from different specific contexts. The ontology has been created within the DBpedia project, which uses multilingual data extracted from Wikipedia:

> "The DBpedia project extracts structured information from Wikipedia editions in 97 different languages and combines this information into a large multilingual knowledge base covering many specific domains and general world knowledge. The knowledge base contains textual descriptions (titles and abstracts) of concepts in up to 97 languages." (Mendes et al. 2012)

For every Wikipedia page, DBpedia creates a URI for the correspondence between entity and Wikipedia page. The URIs are enriched by properties that are extracted through DBpedia and stored as RDF triples. Such triples are used to model semantic data using subject–predicate–object expressions, for example "Konrad Zuse is a Person". As RDF in DBpedia this would be modeled as the following, whereas each element of the triple refers to a URI:

```
<http://dbpedia.org/resource/Konrad_Zuse>
<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
<http://xmlns.com/foaf/0.1/Person> .
```

This example clarifies the heavy use of URIs in order to link and disambiguate data correctly. All of the URIs corresponding to the three used entities have different name spaces. The first element points to the resource "Konrad Zuse" as the subject, where the predicate classifies this as a "type"-relation with the object "Person". The URIs can be accessed via API or the web in order to browse the linked data such as other persons or other attributes of "Konrad Zuse". Data set version 3.8 of DBpedia consists of 1.89 billion RDF triples in 111 languages[1].

DBpedia Spotlight is a semantic technology to analyze a text and to extract concepts from it. This is based on the DBpedia ontology and facilitates disambiguation using the context of the linked data. In natural languages, concepts can be misinterpreted as their surface form is not unique. For example, the surface form *keystone* might refer to an architecture, to a town in Colorado or several other meanings. The surface form of a URI represents the concept in natural language. When we refer to concept extraction and respectively to the results of the extraction as concepts, we refer to the surface form of the corresponding URI. Therefore, we define concept extraction in the context of this work as the retrieval of relevant URIs which characterize a certain source text. Then, the URIs are projected to their surface forms for the output of concepts for our

[1] DBpedia, Data Set 3.8 (2015). `https://wiki.dbpedia.org/data-set-38`. Retrieved: 2020-04-06.

algorithms. However, the results of this extraction are then limited to concepts that have a corresponding Wikipedia (respectively DBpedia) entry. As a consequence, the concepts that have been extracted typically represent domain knowledge or declarative knowledge. Procedural knowledge items can hardly be represented through this. The (internal) concept extraction of DBpedia Spotlight works in four stages:

1. The *spotting* stage uses a lexicalization data set prepared from DBpedia. It uses the LingPipe Exact Dictionary-Based Chunker[2] in order to identify potential mentionings of DBpedia resources. This NLP pipe is based on the Aho–Corasick algorithm, which uses a finite-state machine to match multiple elements from a dictionary (Commentz-Walter 1979). The algorithm outputs all matches including substrings. This is necessary for spotting compound terms, which is not possible through greedy matching algorithms.

2. During the *candidate selection*, possible resources (URIs) are linked to possible surface forms. This step provides a preranking of the candidates in order to employ a "default sense" in a similar way as Wikipedia does. This serves as a direct preprocessing for the next stage.

3. During the *disambiguation* stage, concrete resources are selected and linked to the surface form. To select the correct URI, the disambiguation uses a vector space model of the resource occurrences in the DBpedia corpus. It is a reduction to a ranking problem with a TF-IDF-like scoring.

4. The *configuration* stage employs the configuration by the user in order to define metrics and filters to generate the concrete annotations. This encompasses, a confidence for the candidate selection, a specific entity type, SPARQL queries, and more program parameters.

AlchemyAPI

AlchemyAPI is a service API based on a deep learning approach for text analysis and natural language processing (Turian 2013). The goal of AlchemyAPI was to provide NLP-as-a-service. In this sense, it contained several APIs for different text analysis tasks. However, most of the APIs were not multilingual or did not support the targeted languages of Go-Lab. In addition to this, there are only a few technical details known that explain the concrete mechanisms of the NLP in AlchemyAPI, as this is a proprietary interface. This work employed AlchemyAPI for the extraction of keyphrases from a given text. AlchemyAPI has been acquired by IBM[3] and moved into parts of Watson's

[2] alias-i (2011), LingPipe API. `http://alias-i.com/lingpipe/docs/api/com/aliasi/dict/ExactDictionaryChunker.html`, retrieved: 2019-06-01.

[3] IBM, AlchemyAPI anouncement: `https://www.ibm.com/cloud/blog/announcements/bye-bye-alchemyapi`, retrieved: 2019-06-01.

cognitive service API, in particular into the natural language processing and the discovery API.

In contrast to the DBpedia spotlight approach, AlchemyAPI has the capabilities to extract key phrases from a text. On the one hand this leads to a more diverse aggregation. On the other hand it displays knowledge items that cannot be expressed in a single domain concept. Such phrases might be connected to procedural and metacognitive knowledge items.

Text2Network

Text2Network is web service that employs network-text-analysis (NTA) for the extraction of key concepts (Hecking and Hoppe 2015). The service uses a technical implementation of the NTA approach using the DKPro framework (Gurevych et al. 2007) and Apache UIMA [4]. NTA is a method that is intended to extract conceptual networks that represent mental models from texts. Such networks are characterized by a set of concepts and pairwise relations (Carley 1997). In the implemented approach, concepts are extracted using the Stanford part-of-speech tagger (Manning et al. 2014) and chunking[5] in order to create meaningful noun phrases from the output of the tagger. An entity resolution step helps to identify similar concepts, based on text similarity. For a threshold of 0.7, similar nodes are merged. A relation between two concepts is established if both words co-occur in a sliding window of a certain size.

However, this approach has some limitations. First, a plain NTA approach is not able to identify compound terms without the existence of an external knowledge source for co-occurrences, e.g., in a specific corpus. The sliding window approach splits, for example, *Caesar cipher* into both terms *Caesar* and *cipher*, because the part-of-speech tagger does not have the knowledge about this unique compound concept. Second, the relations between nodes are based on proximity in a sentence, which is error-prone to using relative and demonstrative pronouns. A disambiguation is not part of the process chain and even competing to the entity resolution, which is based on string similarity. To achieve the desirable result of spotting compound terms, a dictionary-based approach has been implemented. In the application (compare chapter 7), an ontology has been used to create the dictionary for the application domain automatically. Such a method helps to bridge science-related concepts that are externalized in knowledge sources like DBpedia to text networks. Following this approach, the dictionary has been enriched by synonyms in order to improve the precision.

[4] Apache, Apache UIMA (2013). `https://uima.apache.org/`, retrieved 2019-06-01.

[5] Apache, OpenNLP Chunker (2019). `https://opennlp.apache.org/docs/1.7.1/apidocs/opennlp-tools/opennlp/tools/chunker/Chunker.html`, retrieved 2019-06-01.

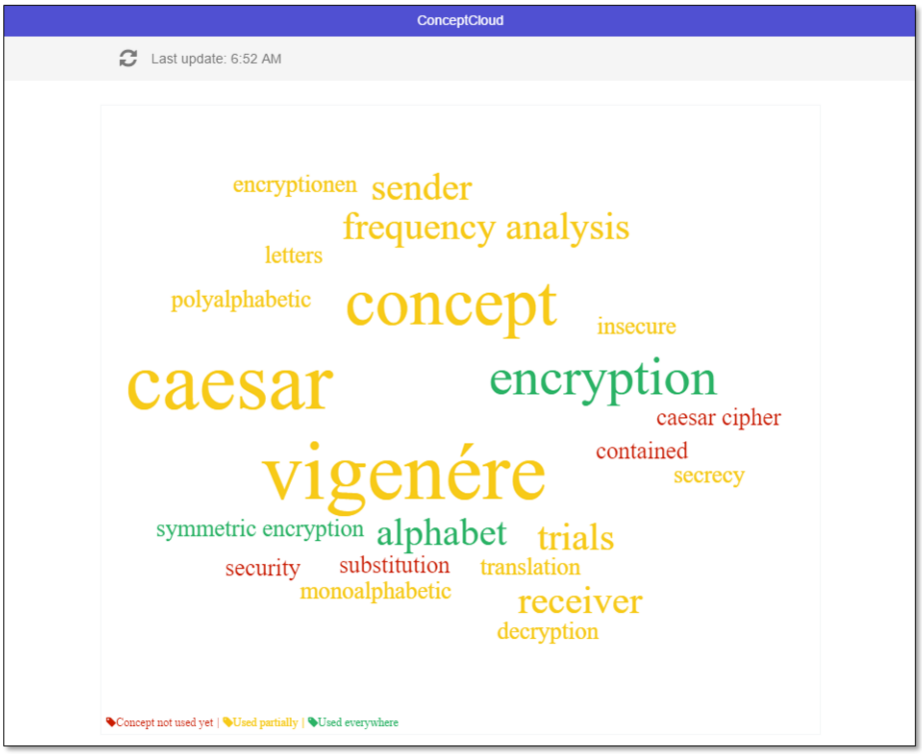


Figure 4.7: Concept cloud app: the data is (translated) from a student in experimental condition C2.

4.3 Concept Cloud

The concept cloud app is an application of the shared group model introduced in this chapter. It serves as an open learner model as mentioned in section 2.3.3. It represents the knowledge of learners, in particular the knowledge of the whole cohort of learners in an aggregated form. The visualization of the knowledge can be used for diagnostic functions on the part of the learners, but it can also support the teacher by providing a summary-like functionality for the aggregation of all learner-generated content. The approach has been presented at the International Conference on Advanced Learning Technologies (ICALT) in the year 2016 and published in the proceedings (Manske and Hoppe 2016). An evaluation of the concept cloud app within a classroom experiment is presented in chapter 6.

The aim of this app is to present a visualization of the shared group knowledge model to support learners and teachers. Following the conceptual model and the processing approach presented in this work, this application can be conceived as a learning analytics tool open to be used by learners and teachers. The app makes use of the technical infrastructure presented in chapter 3.1.2 and is embedded into the Go-Lab learning environment. Depending on the specification and preference by the teacher, this can be integrated into the learning design explicitly as a cognitive scaffold, or implicitly as a reflection tool. Additionally, the monitoring of the group knowledge model has the potential to support the supervision on the part of the teacher. In the evaluated scenario, the app has been used as a cognitive scaffold by the students during the conceptualization phase in a Go-Lab activity. Figure 4.7 shows the concept cloud from a student's perspective with data from a learning space. In the following section, some of the design principles and aspects of the approach of the concept cloud are given.

Visualization of cognitive information The concept cloud app is a tool designed to directly present the information from the data model created by the concept cloud framework (cf. section 4.2.2). It contains a knowledge model of the whole group of learners using a Go-Lab ILS. In such an ILS, each phase might have a dedicated number of production tools, in which the learner creates artifacts such as concept maps, wiki texts or hypotheses. These artifacts are analyzed using semantic technologies as described in section 4.2.3. The concept cloud displays an aggregation of the concepts of all learners. It uses the known visual metaphor of a tag cloud, where more frequent tags have a bigger font size. To render the tag cloud from the aggregated data, a third-party tag cloud library is used. The *Word cloud layout*[6] is based on a layouting algorithm by Jonathan Feinberg (Steele and Iliinsky 2010).

Besides the size of the concepts, the colors also carry information. This enables the learner to have cognitive information about the whole group, and to get additional information about the use or absence of certain concepts in the ILS. Although the concept cloud does not judge about the correctness of the content, the usefulness as a cognitive scaffold to support inquiry processes and knowledge construction has been demonstrated. The contextualized information includes the phase in which the concept occurred, as well as questions for the learners to reflect on. Figure 4.8 shows the information displayed as a tooltip. These questions are selected randomly from a catalog based on literature research. However, first trials have shown that the students did not use the on-demand features such as reflective questions. Therefore, this feature has been removed from the productive version that has been used in the evaluation presented in chapter 6.

Color scheme One of the main design principles of the concept cloud app was to have a learning analytics tool that supports learners through content analysis. Rather than visualizing descriptive characteristics and judging about the correctness or performance of the learners, which gives an illusion of understanding learning processes,

[6] Word cloud layout, Jason Davies. `http://www.jasondavies.com/word-cloud/`, retrieved: 2016.

Figure 4.8: The concept cloud displaying contextual information.

this application is supposed to support knowledge construction and provide useful cognitive information. To make this distinction explicit, we use a certain color scheme exclusively for learners and not for teachers. Although we apply a traffic light color scheme, the information displayed does not correspond to performance characteristics. The colors refer to the usage of concepts across the different inquiry phases by the student. Red means that the student has not used the concept while others did so. A concept is marked green when the learner has used the concept consistently in all inquiry phases and production tools. Yellow color is used when the extracted concept has been used by the learner in at least one production tool, but not in all. As a limitation of this work, the color scheme that has been chosen restricts the target group to people without red–green color blindness. However, the study that has been conducted was based on the concept cloud with this scheme (see 6). In more recent work, another color scheme has been user in order to eliminate these shortcomings. An example of the other color scheme is shown in the outlook of this work (chapter 8.4).

User roles and multiple stakeholders The information displayed is dependent on the role of the user. The concept cloud distinguishes between two roles in the Graasp ecosystem, namely teacher and learner. The view in the role of the learner shows the terms the student used from an "ego-perspective". The terms used (1) throughout

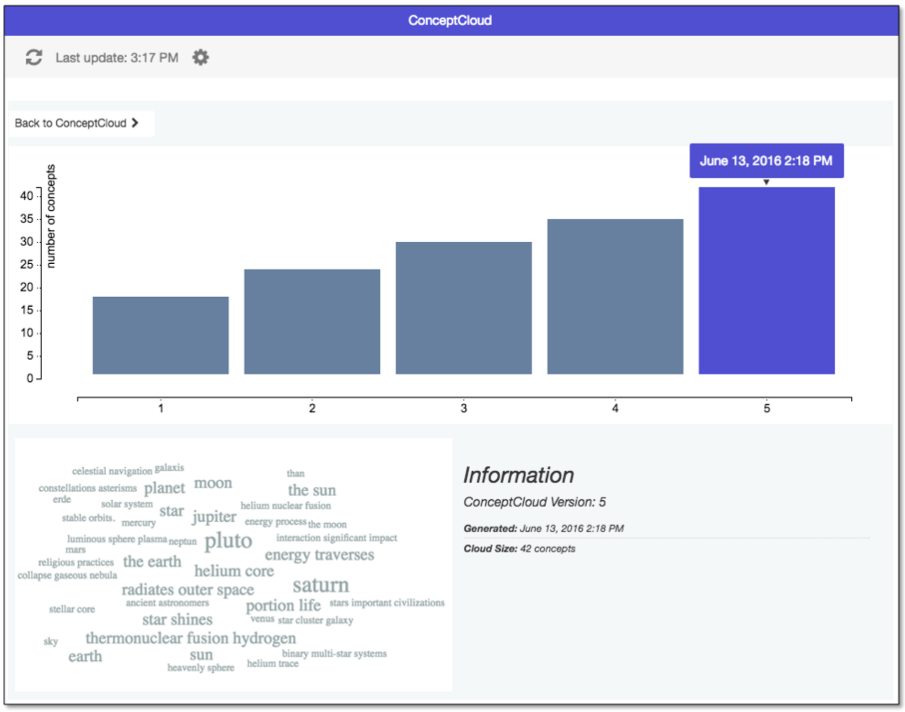


Figure 4.9: The teacher's view displays the concept cloud evolution.

all production tools are displayed green, (2) in at least one tool is shown in yellow, and (3) in no tool at all is highlighted in red. The multiplicities regarding the occurrences of terms in the whole class are displayed, but not personalized. In contrast, the teacher is able to see the names corresponding to the mentions. A support function is embedded to highlight students that deviate from the average. However, this mechanism is not intended to assess learning outcome or other performance characteristics of the learners. This is also achieved by omitting a color scheme for the teacher. The colors displayed in the teacher's view are in a neutral gray, still showing the occurrences through the font size and an explicit display. In addition to the momentary concept cloud, the teacher has the opportunity to review previous versions of the concept cloud. Whether a new version of the concept cloud model is created, depends on the settings of the caching. For every time slice (default: 1 minute) during the editing, one snapshot is created. Figure 4.9 shows the evolution of the concept cloud. A simple bar chart shows the number of concepts in the cloud, but also provides means to select a previous snapshot to be displayed.

Flexibility, interoperability, and transparency Go-Lab provides a rich, technical infrastructure to interact with. As presented in section 3.1.2, the Graasp architecture includes a widget container (Apache Shindig) to render and embed external apps. It embeds a layer for the persistence of application data from within the container. In order to retrieve the learners' artifacts from the Graasp data storage, the concept cloud needs to be connected through client libraries (VLE connector, cf. section 4.2.2). Each concept cloud app retrieves all artifacts of the learners, and normalizes the format or the data structures of the content. For example, different apps produce textual content, in particular the so-called "Input Box", but also the Wiki app. Afterwards, the concept cloud uses the REST interface of the concept cloud server for the extraction of concepts and aggregates the results. The resulting data is then visualized using the aforementioned layouting algorithm.

Contrary to the ROLE sandbox, Go-Lab does not provide any general mechanism or layer for interoperability as part of the system. The openness to third party apps that can be integrated into the Go-Lab ecosystem, and the general nature of data exchange in Apache Shindig lead to a weak specification of application data. Although Go-Lab defined general libraries for storage (CRUD) and a certain metadata format, the content format of each app is specified by the app developer. The retrieval of artifacts is implemented in a library specific for the concept cloud system, for instance the concept cloud app and the semantic group formation app. This library contains content extractors in the sense of data normalization in order to communicate with the text analysis engines on the server.

Privacy The concept cloud performs an analysis of all learner-generated artifacts using the concept cloud server. Due to the design of privacy mechanisms, the concept cloud does not store the artifacts on the server. After each request, the original artifact is discarded and only the aggregated data model is kept on the server (which can be deactivated as well). This approach was compliant with the design of privacy in Go-Lab (Vozniuk et al. 2014). However, storing the model has a huge impact on the performance and the technical scalability of the concept cloud app, which is described in the next paragraph.

Scalability The concept cloud is scalable in different dimensions. First, it does not rely on a domain model or a reference solution, because it is creating a dynamic overlay model based on the extracted concepts. The idea of this approach is to display cognitive information in form of a shared group knowledge model. The optimistic approach (cf. section 4.1.4) approximates and infers knowledge by the mentioning of a particular concept. With the instructions, we prompt learners to critically reflect the occurrence of particular concepts. For example, displaying a red concept in big letters draws a learner's attention to a specific concept that has not been used solving a task such as writing a wiki text. Therefore, we exclude the correctness by design and improve the scalability, as we exclude the authoring efforts of writing reference solutions. However, the design of the group awareness tool presented in section 7 uses a domain model. The results of the extraction have been augmented using an ontology that assigns topics (to be accurate: labeled topics with a URI) to a particular concept. That was necessary to prepare the targeted visualization. However, such a domain model does not claim to act as a filter for correctness.

Second, this approach has built-in multi-language support. Go-Lab is used in more than 1000 schools across Europe. To facilitate ICT and to orchestrate inquiry-based science education in an international context, multi-language support is a requirement per se. Therefore, the concept cloud automatically detects the language of the text in order to use an appropriate extraction mechanism. With DBpedia Spotlight and AlchemyAPI (cf. section 4.2.3), internationalization and text input in multiple languages is supported. Although DBpedia is using a linked data approach in order to provide a multilingual mapping between URIs (Auer et al. 2007), the quality of the underlying ontology based on Wikipedia data varies depending on the language (Lewoniewski et al. 2016). This might have an effect on the text extraction quality.

Third, technical scalability and caching matter for the usage in class. In a prototypical classroom activity, the learners are all working at the same task at one point in time. Although this is more of an optimistic view and not necessarily the case in classroom, most of the learners are in a same activity in a certain time slice. Considering the data processing approach of the concept cloud this has implications for the performance of the concept cloud. In the first instance, this produces a high load of the production system (Graasp) and the school network. If every concept cloud instance fetches all artifacts by all learners, this will cause many requests. For example, if each of the 30 students creates 5 artifacts, this will lead to a total number of 4500 objects to be retrieved from the storage of the Graasp server. There are 150 artifacts in total ($30 \cdot 5 = 150$ artifacts) and for 30 concept clouds to be displayed, there are 4500 objects ($30 \cdot 150 = 4500$) requested from the storage API. With the assumption that most of these requests occur in the same time slice, this might produce a bottleneck in the network. However, processing the texts in order to perform semantic analysis will also produce a high load for the third party cloud services. Therefore, the concept cloud introduces a caching. It only creates one concept cloud per time slice and only if it is requested by a user who opened the concept cloud app (client-side). Then, the whole retrieval mechanism and processing is only initiated, if the client cloud does not find a recent version. The time slice is not a fixed parameter, it can be set flexibly according to the teacher. A smaller value creates shorter time windows, resulting in more concept cloud versions, more traffic, and less caching. The default value is set to one minute. Also bigger values of the time window should not drastically influence the results. During the interaction with the concept cloud, the learners do not work on the production tools, so that the data model does not change. Setting the time window will affect the analytics view, showing a more detailed evolution of the concept cloud.

4.4 Group Formation

Group learning is not a characteristic of the Go-Lab learning approach per se. However, collaboration is in the nature of inquiry-based learning. In Go-Lab scenarios, collaborative learning can take place, but is not mediated or supported by the system. Then, group learning has to be orchestrated in the classroom, with the inquiry-learning space as a shared resource. However, we defined approaches that augment the Go-Lab layer with computer-supported mechanisms to have automatic group formation based on learner-generated content that has been created using the Go-Lab system.

As Go-Lab does not support groups by default, there is no notion of groupings in the system. Particularly the user management on the part of the underlying Graasp platform does not handle groups. However, group work can be emulated on the conceptual level. Using a group ID as a shared login similar to an individual login helps to circumvent the shortcomings of the Go-Lab learning environment on the technical level. The collaboration is then coordinated in the classroom setting, while the ICT is used to support the orchestration and initialization the collaboration, for example, by displaying the groupings, giving instructions or even displaying learner or group models. In the first approaches for skill-based group formation in Go-Lab (cf. chapter 5), this mediation is done externally based on the data in the learning analytics server. For the semantic group formation (see section 4.4.2), this is realized using an app for group support and group formation that is embedded into the learning space. On the architectural level, it is integrated into the learning analytics infrastructure following the concept cloud approach.

Learning scenario The envisioned learning scenario for group formation in Go-Lab is a two-phase design. Initially, learners use Go-Lab individually in order to create artifacts to express their knowledge or demonstrate their skills. These artifacts are processed and interpreted by the system. We employ two different pedagogical goals and respective learner models as a baseline for group formation. The first approach, a skill-based group formation, infers skills from the analysis of the learner-generated artifacts and builds heterogeneous and homogeneous groups. We used this approach to test how groups, depending on the heterogeneity, perform regarding learning gain.

The second approach, semantic group formation, takes the learner-generated objects and creates a shared group knowledge model from the approach of this work rather than inferring skills. Using this knowledge model, groups are formed satisfying a different criterion, namely knowledge complementarity. This follows the more inclusive view on diversity, which emphasizes that we do not group highly-skilled students with low-achievers, but assumes that we can define groupings where all learners benefit from each other, as they have different levels of expertise.

4.4.1 Skill-based Group Formation

This section covers the framework for the skill-based group formation. The model consists of a composition of several features that correspond to performance related characteristics such as learners' skills like text writing or concept mapping. Complementary to the skills, affective components, particularly motivational scores have been included into the feature set of the model. In this part of the work, an overview of all measurements used and their backgrounds is provided. The conceptual model has been applied in an experiment using the Go-Lab environment. This experiment is described in section 5. This exploratory work is tailored to the scenario of the evaluation and thus does not provide an exhaustive set of features that suits all possible scenarios. This is rather a cutout of several features that have been identified as relevant for this particular Go-Lab learning space in the frame of this group formation scenario presented.

Group Formation Processing Chain

We define a heterogeneous learning group as a learning group, where each member has different performance characteristics. The learners produce artifacts during an inquiry-based learning scenario as described in the experimental setting. The artifacts, particularly learning objects, and the assessment of motivational scores form the data set for the group assignment. These characteristics span a feature space. The vector, which contains the scores for a single student, is called feature vector and is an element of the feature space. To use simple Euclidean distance measurements in such a vector space, the feature vectors are normalized.

In total, we capture the performance characteristics through six artifact-related and three motivational scores, leading to a nine-dimensional feature space. In terms of classroom size, the dimension is too high to produce meaningful clusters. To tackle this curse of dimensionality, we perform a feature selection to minimize the dimension to a plausible number derived from the number of students and groups.

Features for the Group Formation

To decide whether a group is heterogeneous or homogeneous, different performance characteristics that serve as features for the group formation have been captured. These incorporate not only artifact-based assessments on concept maps and small texts, but also motivational assessments based on the SMTSL questionnaire (Tuan et al. 2005). The following section lines out different measurements of these performance characteristics. However, for the experimental study of the skill-based group formation, a subset of these features has been consulted, which is presented in chapter 5. Although the composition covers all of the skill categories presented in the following, the number of features is restricted in the experiment due to a low sample size compared to the relatively high dimensionality.

Concept maps Concept mapping can be characterized as a technique for the externalization of knowledge structures (cf. section 2.1.2). A learner creates a concept map by connecting concepts that are considered important for a given domain by labeled relations. Since concept maps reflect individual learners' structures of domain knowledge, these artifacts are particularly suitable to characterize students in addition to knowledge tests (Stoddart et al. 2000). In order to use the concept maps of students as parameters for group formation, a quantification or rating of the learner-generated artifacts is needed. One approach is to compare a concept map to a reference map created by a teacher, tutor or (domain) expert (Conlon 2004; McClure et al. 1999). This requires a matching of concepts between both maps by comparing labels. This can be done automatically by using computer linguistic methods (Conlon 2004; Hoppe et al. 2012). However, this is not trivial and can lead to a wrong matching. For the study presented in section 5 it was important to measure the impact of the group formation on student performance as accurately as possible. Therefore a manual processing step has been introduced in order to avoid biases introduced by an automatic matching. However, the processing of concept maps among other artifacts for the knowledge-based approaches (semantic group formation, concept cloud) has been performed by using semantic technologies.

Formally, concept maps were described as multi-graphs and thus can be characterized using measures from (social) network analysis (cf. section 3.2.3). A concept map cm consists of a set of concepts N_S and a set of relations E_S. Each concept map is compared to an expert map (e.g., by a teacher) with the concepts N_E and relations E_E. Five different measures were calculated, whereas the last one is a combined measure based on a regression model:

Node precision Node precision measures the ratio between correctly defined concepts and student concepts. Correctly defined concepts are student concepts that can be matched to the expert concept map.

$np(cm) = \frac{|N_S| \cap |N_E|}{|N_S|}.$

Node recall This measure indicates to which extent the concepts in the expert map are covered by the student map.

$nr(cm) = \frac{|N_S| \cap |N_E|}{|N_E|}.$

Edge precision Edge precision characterizes the fraction of connections (edges) in the student concept map that can be also found in the expert concept map.

$ep(cm) = \frac{|E_S| \cap |E_E|}{|E_S|}.$

Edge recall Edge recall is defined as the fraction of edges in the expert concept map that can be found in the student concept map.

$er(cm) = \frac{|E_S| \cap |E_E|}{|E_E|}.$

HEW measure Hoppe et al. (2012) introduced a quality indicator for concept maps based on the comparison of a concept map to a given ontology. Such an ontology extends a static expert map by inferred items such as synonyms. The measure was obtained based on empirical observations of structural properties that correlate with expert quality judgments.

$hew(cm) = \frac{|N_S|}{1+3|N_S \cap N_E|} + \frac{7|E_S|}{1+6|E_S \cap E_E|} + \frac{|N_S| \cdot |E_S \cap E_E|}{1+6|E_S| \cdot |N_S \cap N_E|}.$

As such, this set of measures is limited regarding a reliable assessment of the students' actual skills, as it rarely covers domain knowledge. However, these measures contribute to the creation of heterogeneous and homogeneous student groups by providing discriminating factors of learners that match performance criteria. In this sense, the features do not necessarily answer the question of which students produce better concept maps but they give insights into which students produce different concept maps, and thus have different characteristics. The approach of semantic group formation incorporates knowledge models into the grouping (cf. section 4.4.2).

Text writing Since text analytics and approaches of text mining still had huge deficits regarding the automatic evaluation of artifact quality or writing skills, a non-automatic measurement for the text quality characteristics has been used. Those shortcomings specifically occur in short learner-generated texts in STEM fields (Leeman-Munk et al. 2014), while promising technologies to extract structural semantic information such as DBpedia Spotlight put their focus on representing declarative knowledge (Manske and Hoppe 2016). As a caveat, there is a difference between automatic quality assessment of learner-generated texts and an automatic extraction of keywords. Particularly (automated) semantic technologies have the potential to produce much better results, which is presented in section 7.9. The assessment of text quality used in this work has been performed by measuring the concept coverage based on teacher solutions and manual coding of the concepts from the student and teacher solutions. The coverage simply counts the ration between correct learner concepts and teacher concepts. Another feature extracted is the number of concepts in the student text and the number of words.

Motivation Curiosity or personal preferences might have an impact on the learning results, particularly in collaborative science-learning. It is obvious that motivation is one of the key ingredients for a successful group work. However, a variety of motivational components have been identified in the literature. To measure motivation towards science, Tuan et al. (2005) developed a questionnaire using six scales: self-efficacy, active learning strategies, science learning value, performance goal, achievement goal, and learning environment stimulation.

To assess science motivation suited to the Go-Lab scenario that has been evaluated in section 5, the SMTSL questionnaire has been shortened. In the version used it measures three different categories of motivation: (a) self efficacy, (b) science learning value, and (c) learning environment stimulation. Although motivation is one of the features used for the clustering, it can be seen more as a filter or a side condition to ensure that each group consists of enough motivated students. The questionnaire also consists of items that are of interest regarding group activities. For example, the category self efficacy contains the item: "During science activities, I prefer to ask other people for the answer rather than think for myself." For the featurization, a score is assigned to each category corresponding to the underlying scale of the questionnaire.

4.4.2 Semantic Group Formation

The learning spaces in Go-Lab help to orchestrate inquiry-based learning activities which are structured in subsequent inquiry phases. Spaces can be enriched with resources and supportive apps to scaffold the activity. Specific apps enable the students to generate knowledge artifacts such as texts, concept maps or hypotheses. Due to the heterogeneity between different artifacts of these types, the interpretation of learner-generated content is challenging for students and teachers, which underlines the need for an aggregated representation of the content to support learning and knowledge building on an epistemic level (Zhang et al. 2009; Wilson 1996). The use of computational methods of semantic extraction has the potential to support students and teachers in understanding and reflecting on their activities in the Go-Lab environment (Manske and Hoppe 2016). Particularly when it comes to the orchestration of heterogeneous learning groups, it is challenging for a teacher to compose groups of complementary knowledge.

To meet this objective, a framework system to enable semantic extraction from learner-generated artifacts has been proposed (Manske and Hoppe 2016). Facilitating this approach and the given Go-Lab environment, an app and an algorithm for group formation in this sense have been developed. The semantic group formation uses semantic extraction from learner-generated content to create heterogeneous learning groups in terms of knowledge diversity. In contrast to former score-based quantitative approaches (Manske et al. 2015c), this follows the idea that a group benefits from its students' complementarity and diversity (Hoppe and Ploetzner 1999). Figure 4.10 shows the idea of grouping learners regarding knowledge complementarity. Traditional approaches for group learning, such as the Jigsaw approach by Aronson et al. (1978), rely on the idea that learners complement each other by having different areas of expertise. The original Jigsaw model consists of three different group phases. For the first phase, Jigsaw groups are created, where each learner has an individual topic or task referring to his or her field of expertise. After a certain time, each expert will be assigned to an expert group, which consists of all experts for a certain topic. During the second phase, the experts exchange their knowledge and results, or elaborate further on the tasks. Finally, for the third phase, the experts go back into their Jigsaw group and report on the results. To retain the findings, this phase might be enriched by additional tasks.

However, the approach of the semantic group formation goes beyond the Jigsaw model, as it provides a more general approach that does not need an adaptation of the pedagogical design (different tasks for different experts in Jigsaw). Further, the semantic group formation quantifies knowledge and knowledge diversity by analyzing learner-generated content. As a consequence of not changing the pedagogical model, all the students have the same prerequisites when finishing the individual phase (expert

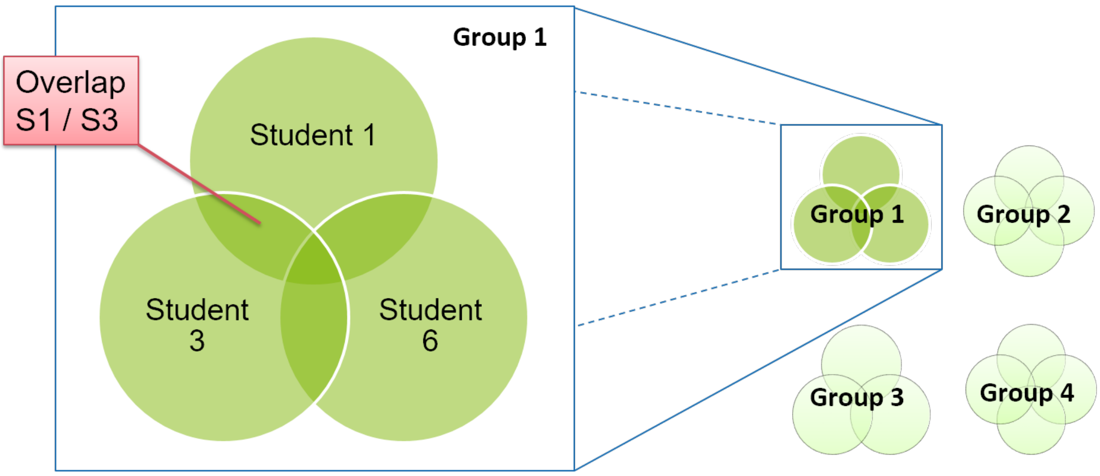


Figure 4.10: The goal of semantic group formation is to maximize knowledge complementarity.

phase in Jigsaw). Since the semantic group formation performs an optimal grouping regarding the knowledge diversity, this leads to a fair and explainable group formation. Thus, the semantic group formation relies on the principles of analytic models (Hoppe and Ploetzner 1999) and Jigsaw (Aronson et al. 1978). The expert knowledge that is usually created during the expert phase in Jigsaw is assigned by the semantic extraction. Then, the semantic group formation algorithm calculates groupings that benefit of situations of complementary knowledge, but with the flexibility of a variable group size.

Data Model

In this section, an algorithm for semantic group formation and the developed app that facilitates this algorithm in Go-Lab is described. It follows and advances the approach of score-based grouping (Manske et al. 2015c). In the former approach, features have been extracted automatically from student artifacts using quality indicators. In this sense, a skill is a composition of scores that represents artifact-specific features, e.g., the complexity of a student's concept map, or the number of teacher-defined key concepts covered by a student's text. A deficit in this approach is that similar scores do not imply similar fields of knowledge. Therefore, heterogeneity rather refers to diversity in skills than to diversity in knowledge. To effectively form groups, the knowledge diversity over all groupings is maximized.

Depending on the approach of the shared group knowledge model (cf. section 4.2), each artifact type needs its own mechanism to extract the relevant concepts. The general approach for a concept extractor is to map an artifact to a list of concepts that represent its knowledge items. All concepts of each student are aggregated into a unified data model. The data model mainly represents the relation R_{CC} over the sets of learners (L), artifacts (A), phases (P), and extracted concepts (C):

$$R_{CC} \subset L \times A \times P \times C$$

On the technical level this data model is encoded in a structured JSON-based format with contextual information corresponding to the virtual learning environment and redundancy, e.g., for a client-side/hybrid caching module and retrieval.

Semantic Group Formation App

The semantic group formation app (see figure 4.11) has been implemented using web technologies and has been integrated into the Go-Lab environment. The app provides additional information to the teacher such as the participation of learners (e.g., artifacts that are not produced or required artifacts). If the app is present in a Go-Lab learning space, it can be used to calculate a group formation. Based on the learner-generated content in the space it creates a concept cloud and the corresponding data model. Following an optimistic learner model approach (cf. section 4.1.4), the data model approximates the total knowledge of the cohort providing an upper bound. To perform the group formation, the data model is then retrieved from the server and used to calculate the optimal groupings.

The design of the group formation app has been influenced by several pragmatic decisions. In such a heterogeneous environment as Go-Lab, teachers have the flexibility to orchestrate learning spaces according to their own and their students' needs. However, learners do not necessarily perform all tasks they have been prompted to. Some of them produce minimal or non-existent solutions or leave out several tasks. However, a group formation app that will be used in large-scale implementation projects such as Go-Lab has to produce valuable output in most of the cases. Therefore, the app has several degrees of freedom to deal with missing artifacts or even with the exclusion of specific learners from the grouping algorithm. Teachers might want to specify that certain artifacts shall be replaced with another solution (from another student or an artifact from the teacher). Therefore, the app provides information about the state of missing artifacts and the percentage of participation in certain tasks. Apart from the grouping, this app provides valuable information to monitor the progress of learners in terms of artifact creation and participation in the different inquiry phases.

Two of the big challenges for algorithms that are used to support orchestration of learning, are trust and transparency of the used representations and data structures.

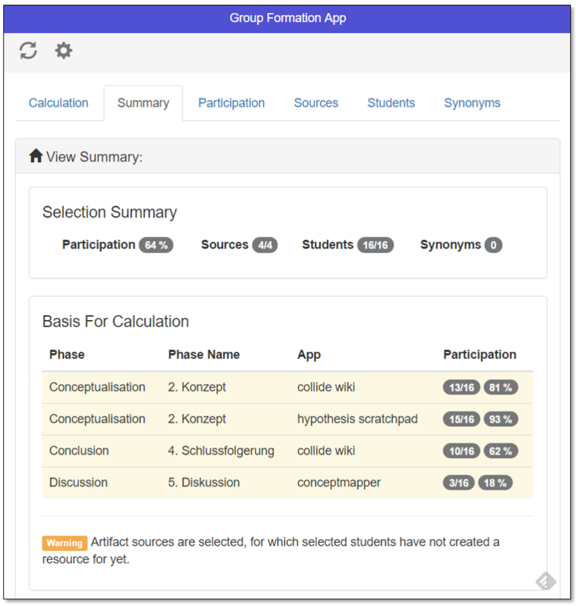


Figure 4.11: The user interface of the semantic group formation app.

This problem of a missing *interpretability* is persistent in different fields, where automated decision making takes place, for example in machine learning (Doshi-Velez and Kim 2017). Modern approaches of machine learning and artificial intelligence reach out into our everyday lives and have a big societal impact (Abdul et al. 2018). Apart from the urgent need in many fields to provide cutting edge and high performance algorithms to achieve near-realtime behavior, *explainable AI* (XAI) is another recent trend in the research (Adadi and Berrada 2018). In other fields like educational technology, there seems to be a consensus that a weaker algorithm (in terms of precision or accuracy) is preferable when the results have a better interpretability. Beyond the developers and researchers, teachers and students as the first-class stakeholders need to have trust in the technological artifacts in order to adopt technology in their working practices. This has two practical implications:

1. We do not use a *greedy algorithm* to form groups. Greedy algorithms might produce only a local optimum, which makes it difficult to trust the results in a setting, where the teacher might have some better knowledge about his or her learning group.

2. *Interpretable visualizations* of the results and the approach of the algorithm itself are needed. Therefore we designed a visual representation that helps to explain the results by explicitly showing the knowledge complementarity and the knowledge overlap. As a consequence of the interpretability, unwanted, missing or incorrect results might raise the need on the part of the teacher to place corrections, which can be done in the web interface. This can be achieved, for

example, by adding synonyms for concepts, as well as by editing or excluding particular student artifacts.

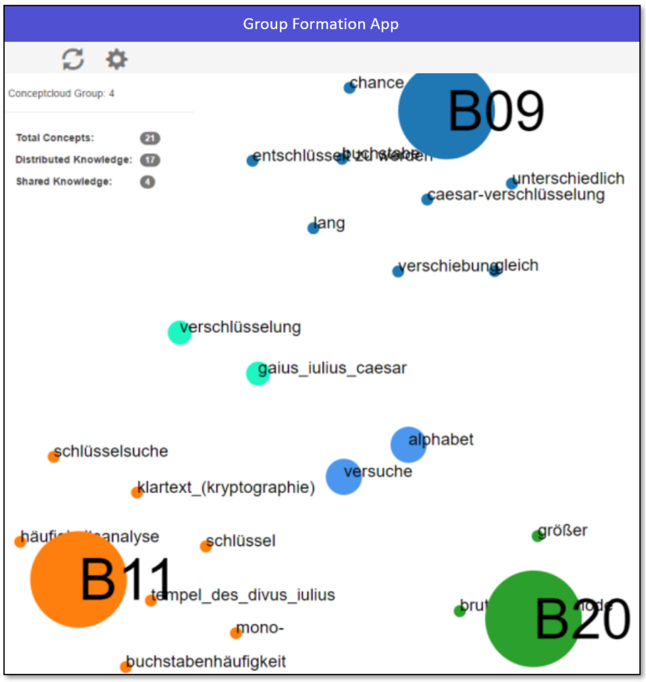


Figure 4.12: An explicit representation of the shared and distinct concepts in the semantic group formation app.

The visual representation of the entities in the group formation can be seen in figure 4.12. It shows the distribution of knowledge items in a learning group of three students (B09, B11, and B20). Some concepts are used by several students and thus overlap, which is indicated by the concept location and color. The relatively small overlap and the magnitude of unique concepts in this group stand for a high knowledge diversity according to the presented definition. This visualization helps to identify distinct and unique concepts and also explains the results of the algorithm. Such a visualization of cognitive information can be used in combination with instructions as part of the educational or pedagogical design as a cognitive group awareness tool. This approach has been evaluated using the semantic group formation and an additional visualization that makes the distinction of learners' knowledge easily quantifiable by displaying knowledge differences in a bar chart (cf. section 7).

Diversity Score

The measure for the diversity of a cohort is the product of the groups' diversity scores. The groups' diversity scores follow a model where the complementarity of knowledge items between two students in terms of the concepts they used influences the scoring positively. While similarity scores such as the Jaccard index quantify the overlap of two sets, their complement can be used to characterize the diversity, namely the Jaccard distance. This operationalization quantifies the diversity based on the complement of sets. All pairwise calculated diversities between two students' concept sets contribute to the group diversity score D(g). The following formula for the diversity of a cohort L is constructed based on these premises (with $J(a,b)$ representing the Jaccard similarity of pairs of students in the student group S_g)

$$D(L) = \prod_{g \in L} D(g) = \prod_{g \in L} \sum_{a,b \in S_g} \frac{1 - J(a,b)}{C}$$

The weighting C is calculated as the number of possible (pairwise) combinations of students. This normalizes the sum of all diversity scores inside a group respecting that the number of students per group might differ. Otherwise, the scoring leads to major benefits of diversity scores for bigger groups caused by a combinatorial explosion. The number of all pairs of students in a group S_g corresponds to the binomial coefficient.

$$C = \binom{|S_g|}{2} = \frac{(n-1) \cdot n}{2}$$

According to the definition of complement from set theory, the calculated diversity measure using the Jaccard index as proposed, characterizes the relative complement of the sets. Of course, it would be possible to reformulate the calculations based on an absolute complement that considers the whole set of concepts from all learners. However, from an application perspective, it is desirable to create learning scenarios and situations in which the participants benefit from the group knowledge model. This is mediated through social interaction, particularly in explicit knowledge exchange phases and does not necessarily need the whole class as a reference point or as a continuum for the knowledge (cf. section 7). This is different for individual cases of learner support, where the diversity inside a class can be measured to quantify the knowledge distribution without explicitly relying on social interactions.

Algorithmic Complexity

For the estimation of the run-time complexity, we exclude the case of the random sampling for simplicity reasons, as it only introduces a constant factor. To estimate the number of possible groupings for the group formation, we reduce this to a partitioning problem of sets. The pairwise diversity scores are calculated in advance in a diversity matrix using Jaccard distance. As the Jaccard distance is based on calculating an intersection and union of two sets, which can be done in $O(m \cdot log n)$ (Baeza-Yates 2004), the diversity calculation step can be performed in $O(n^3 log m)$. With the precalculated results, the calculation of the diversity for each partition can be boiled down to a lookup of pairwise scores in the matrix, which contributes as constant time $O(1)$ to the complexity. However, in total, the systematic partitioning of the groups contributes the most to the computational complexity.

Let L be the set of learners with $L = \{s_1, s_2, \ldots, s_n\}$, and $n = |L|$ be the number of learners. The bounds of the allowed group sizes are given by a and b, where a is the lower bound and b is the upper bound of the group sizes. In this sense, for the allowed group size k is $a \leq k \leq b$. Furthermore, the flexible group size allows different numbers of partitions n_p with $p \leq n_p \leq q$. The minimal partition size p and the maximal partition size q are calculated as follows:

$$p = \left\lceil \frac{n}{b} \right\rceil, q = \left\lfloor \frac{n}{b} \right\rfloor$$

The Stirling Number of the Second Kind can be used to calculate the count of all partitions within the given bounds. It is defined as follows: Stirling Numbers of the Second Kind, denoted by $S(n,k)$, is the number of ways of partitioning a set of n elements into k non-empty subsets (Abramowitz and Stegun 1965). The closed form of the generating functions to calculate the numbers is the following:

$$S(n,k) = \frac{1}{k!} \sum_{i=0}^{k} (-1) k - i \binom{k}{i} i^n$$

Within the scenario of forming groups, for example, creating dyads from 10 students will lead to a partition into 5 subsets. The Stirling Number of the Second Kind quantifies the different ways to form the groups. Consequently, the flexibility of the group size adds up the different kinds of partitioning related to the given group boundaries p and q. Therefore, the total number of partitions N can be calculated as

$$N = \sum_{k=p}^{q} S(n,k)$$

Corcino et al. (1999) provide an asymptotic approximation for the Stirling Number of the Second Kind with $S(n,k) \sim \frac{k^n}{k!}$, so we can conclude that the run-time complexity of this algorithm is roughly estimated by

$$O\left(\frac{k^n}{k!}\right).$$

Example. For a group of 9 students ($n = 9$), a teacher specified that the group size should be between 2 and 3. With $a = 2$ and $b = 3$ follows $p = \lceil \frac{9}{3} \rceil = 3, q = \lfloor \frac{9}{2} \rfloor = 4$.

With $S(9,3) = 3025$, and $S(9,4) = 7770$, this results in 10795 possible partitionings of the set of students:

$$N = \sum_{k=3}^{4} S(n,k) = S(9,3) + S(9,4) = 10795$$

4.4.3 Comparison

In this chapter, the approach of semantic group formation has been presented. Although Go-Lab is not collaborative per se, it has been shown how group formation can be applied to inquiry-based learning scenarios in Go-Lab. The main benefit of this work is the provision of tools for teachers to manage knowledge diversity in classrooms. On an argumentative level applying such algorithms in a group formation scenario is more plausible than a score-based group formation. Although the semantic complementarity induces a heterogeneity (in knowledge), the differences in skills or scores do not necessarily represent this. However, on a semantic level different concepts might be grouped to similar or even synonymous terms. In future work, ontologies or other external knowledge representations might be used to tackle this issue of incorporating semantic closeness. According to the visualization of the group compositions in the presented "Semantic Group Formation App", similar visualizations that explicitly show semantic closeness or knowledge complementarity can be presented to the learners in a group learning scenario following Bodemer's idea of cognitive group awareness (Bodemer and Dehler 2011).

The main focus of this chapter was to better quantify the notion of heterogeneity in knowledge and to operationalize the concept of knowledge diversity. Therefore, a group formation algorithm that forms groups based on the complementarity of students' knowledge with the goal to maximize the overall coverage and minimize the overlap has been implemented. The grouping has been visualized with respect to the group members' shared knowledge. The user-interface has been evaluated with a focus on this visualization in a small expert group (6 instructors familiar with Go-Lab) in a questionnaire. The experts' questionnaire aimed at evaluating the user experience regarding the visualization of the student models and the shared knowledge inside learning groups. This was generally perceived as relatively useful by the experts (M = 2.20, SD = 0.84, 1 = useful, 4 = not useful). Negative aspects are the color scheme and the comparability across different groups, as the interface does not allow reviewing groups in parallel. However, the approach to feed back the cognitive information into the learning scenario has been evolved to a cognitive group awareness tool (CGAT) in order to benefit from knowledge complementarity in the sense that an explicit knowledge exchange phase has been implemented. Thus, the learner can make use of this constellation during this phase, for example, by asking each other questions and by initializing a fruitful collaboration (cf. section 7).

As a limitation of this work, a focus has been put on skill-based heterogeneity and knowledge-based diversity in the comparison. However, other measures for diversity and heterogeneity exist in the literature. Entropy measures have been applied to measure the diversity in population distributions (White 1986). In the field of ecology, the types of species are of interest and accordingly the operationalization of diversity in this area. There, it is a common task to measure and quantify the variations of species, particularly the *effective number of species*, the so-called "true diversity". A variety of definitions and ways to calculate diversity exist. While there have been efforts to create a unified notion and interpretation of diversity measures based on entropy (Hill 1973), a consensus of different measures is not generally given. They are even declared as being conflicting in their notion and behavior as far as being meaningless (Hurlbert 1971). However, the different diversity indices imply their own interpretation of “diversity” and many have been applied in other fields than ecology. A popular measure is the Gini index for (individual) diversity, which exists in a variety of forms such as the extended Gini and the Gini decomposition that are used for inequality of other distributional phenomena that incorporate a custom weighting (Lerman and Yitzhaki 1984; Xu 2003; Taagepera and Lee Ray 1977). This index can be used to characterize "the degree of diversity of each individual from all other members" (Ceriani and Verme 2015). In contrast to the knowledge management approach presented in this thesis, such measures quantify how one individual contributes to the sum, i.e., it lines out a distribution of inequality. Thus, it rather renders a divergence than a complementarity, which is supposed be the desired operational model for diversity in the foreseen learning scenarios. Particularly, such measures cannot be easily adapted to meet the goal criteria used for the group formation algorithm and though optimize the output. Still, such measures have the potential to serve as a quality indicator for imbalance. However, the diversity scoring presented in this thesis as part of the semantic group formation tends to balance the diversity scores across the groups (compare section 7).

5 Evaluation of the Skill-based Group Formation

There is an increasing interest in student-centered teaching methods with small group learning as an important ingredient. In this section, we present a study in which the performance of heterogeneous and homogeneous learning groups has been compared in a technology-enhanced classroom setting in the area of STEM learning. The group formation was based on learning analytics results that were considered in a semi-automatic formation process. The analytic methods used incorporated different artifact-related characteristics, but also motivational features as input. We observed that the heterogeneous groups outperformed the homogeneous ones in different ways. The results of the study are analyzed using quantitative and qualitative approaches on both the individual and the group level. This section is based on a publication for the CSCL conference in 2015 (Manske et al. 2015c).

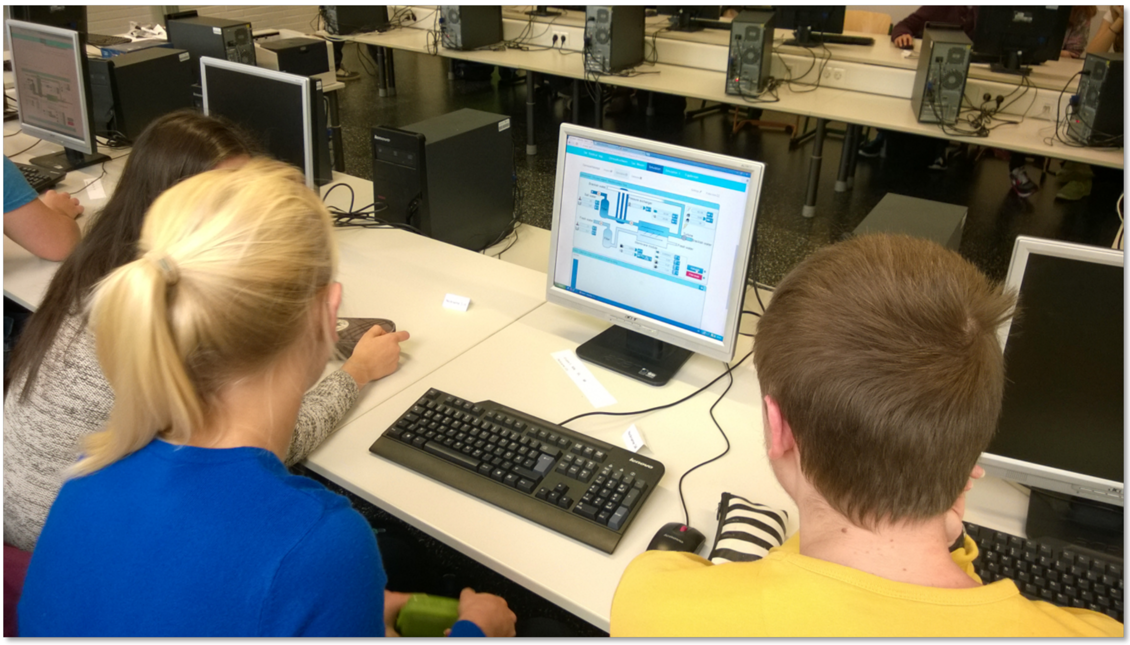


Figure 5.1: The collaboration took place off-the-system when the students experimented in small groups within the Go-Lab learning environment.

5.1 Experimental Setting

This section covers the experimental setting of the study. First, we describe the Go-Lab platform which will be explained in line with the implemented scenario. Apart from the technical system, we explain the didactic goals and the production of learning objects by the students during this scenario. These artifacts were used for the assignment of groups and the assessment of performance characteristics.

The learning activity was split up into two phases: the first phase consisted of individual student work for the initial assessment of performance characteristics. The tasks that the students had to carry out involved writing a short text to describe a simulation, and creating a concept map from different learning resources. A motivational questionnaire captured their interest and motivation in science. These characteristics served as an input for the group formation, which was used in the second phase of the study. Here, the students performed an inquiry-based learning task in groups. The task objective was the online experimentation with a virtual lab of an osmotic power plant.

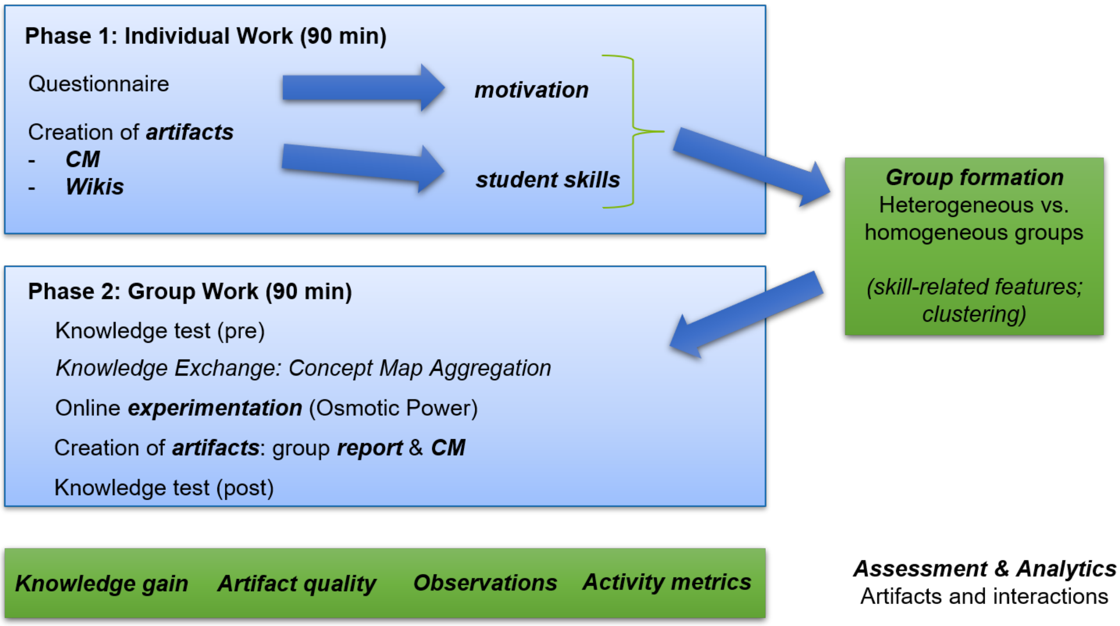


Figure 5.2: The design of the study consists of two parts: the results of the individual phase have been used to form groups for the second phase.

The outcome of the second phase was both a concept map and a written report per group. The concept map should describe the parameter model of the power plant, and for the report the students had to formulate a short summary about their findings. This was supposed to include a critical reflection about the usefulness of osmotic

Table 5.1: The features presented in this table have been used as a basis for the skill-based group formation.

Learner	Cluster	CM: HEW	CM: Node Recall	Wiki: Coverage	Motivation: Self Efficacy	Motivation: SLV	Group
S1	0	0	0.10	0	0.9	0.63	G5
S2	2	0.02	0.10	0.7	0.1	0.75	G4
S3	0	0.12	0.3	0.8	0.7	0.38	G2
S4	0	0.01	0	0.1	0.6	0.25	G6
S5	2	0.42	0.40	0.3	0.6	1	G3
S6	2	0.42	0.50	0.4	0	0.75	G2
S7	0	0.28	0.3	0.2	0.3	0	G1
S8	0	0.001	0.20	0.5	0.7	0.63	G5
S9	2	0.36	0.40	0.5	0.1	0.5	G4
S10	0	0.14	0.20	0.6	0.7	0	G6
S12	1	1	1	0.3	0.5	0.5	G3
S13	2	0.66	0.60	1	0.6	0.75	G1
S14	2	0.58	0.50	0.6	0.3	0.63	G4
S15	1	0.54	0.50	0	0.8	0.63	G1
S16	0	0.41	0.40	0.5	0.4	0.25	G5
S17	1	0.55	0.60	0.1	1	0.13	G2

power under different aspects, e.g. sustainability, effectiveness and dependence of the location. Four explicit assignments guided the students through the scenario and provided a scaffold for the report. However, no formal structure was given in order to promote an open-ended range of possible solutions. The schema in figure 5.2 summarizes the design of the experiment. This schema represents how learner-generated artifacts are used within the two-phase design in order to extract features such as motivational or skill-related scores. Then, the group formation is based on this featurization, whereas the operationalization follows the description in section 4.4.1 of this work.

Only the most distinctive features have been selected using a variance-based dimension reduction to deal with a relatively small set of learners. The following features have been used for the group formation in this study: (1) HEW-measure, (2) node recall, (3) motivation: self efficacy, (4) motivation: science learning value (SLV), and (5) wiki term coverage. The third category of science motivation is left out as it is not distinctive. The learners' score in the category "learning environment stimulation" is inverse to the "science learning value" score and thus does not provide any additional value to the input of the clustering. Table 5.1 shows these features and the respective results of the clustering and the group formation. Groups G1-G3 are heterogeneous groups (selection from different clusters) and groups G4-G6 are homogeneous groups (each group is selected from the same cluster).

The calculation of the actual group formation is based on a k-means clustering using the aforementioned five features. These features were fed into the clustering as an input for the k-means algorithm with three clusters due to a desired group size of three. However, the number of students in the experiment was 16, which cannot be divided equally into triads. Therefore, two dyads (G3 and G6) were created. The three clusters represent different skill-related characteristics derived from the used features. Our underlying definition is that a heterogeneous group consists of different characteristics, whereas a homogeneous group is composed of learners with similar skills. Therefore, these clusters have been used to assign groups the following: Heterogeneous groups are formed using a selection from different clusters, while homogeneous groups are created by sampling from within a cluster of equally skilled learners. In the procedure of the experiment, heterogeneous groups are created first by a random selection from different clusters, whereas the third group (G3) is a dyad to balance the group sizes for the two conditions with a total of 16 learners. Afterwards, homogeneous groups are created by a selection of the remaining students and grouped within same clusters. Cluster 2 only consisted at this point of three learners, which immediately formed a triad G4, while cluster 0 was randomly divided into two homogeneous groups (G5 and G6).

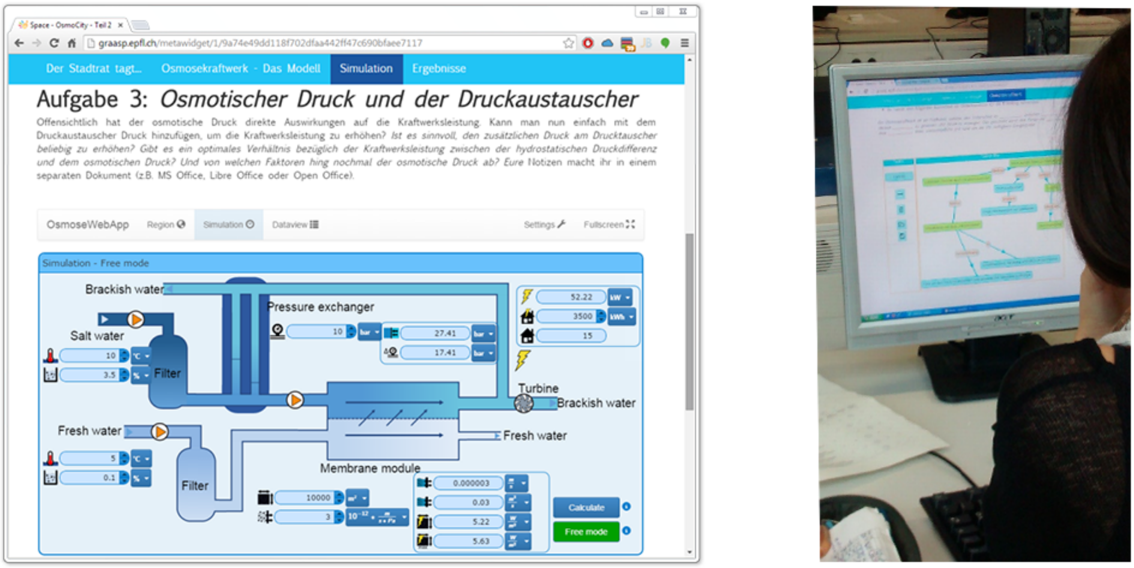


Figure 5.3: The web-based learning environment with the interactive osmotic power plant simulation (a), and a student using the concept mapper in this environment during the study (b).

5.2 Learning Scenario and Goals

Figure 5.3 shows a screenshot of the inquiry learning space of the environment, which has been used for this study, from the students' perspective. The learning activity consisted of different inquiry phases, which are displayed as "tabs" in the navigation bar of the web environment and thus define a guided path through the inquiry process.

The main goal of our learning scenario is to understand the mechanism of osmotic power and how the location of an osmotic power plant influences the power generation. The learning scenario demands multidisciplinarity from the students in a way that knowledge from different subject domains such as biology, chemistry and physics is used. Also competencies from different fields such as text writing, metacognitive skills, concept mapping and inquiry skills are released during this experiment.

Critical thinking skills are demanded in the second phase of the study, where the students perform the group work task. At the beginning of the group phases, they get confronted with the “aggregated concept map” of all students (Manske et al. 2014), which can be seen as a union of all concept maps represented as graphs. Figure 5.4 shows the approach of aggregating concept maps and figure 5.5 presents the aggregated concept map of the individual maps from the experiment. Such a structure contains useful and useless concepts and possibly wrong connections. This enforces a critical group discussion about the correctness of specific parts. In the following,

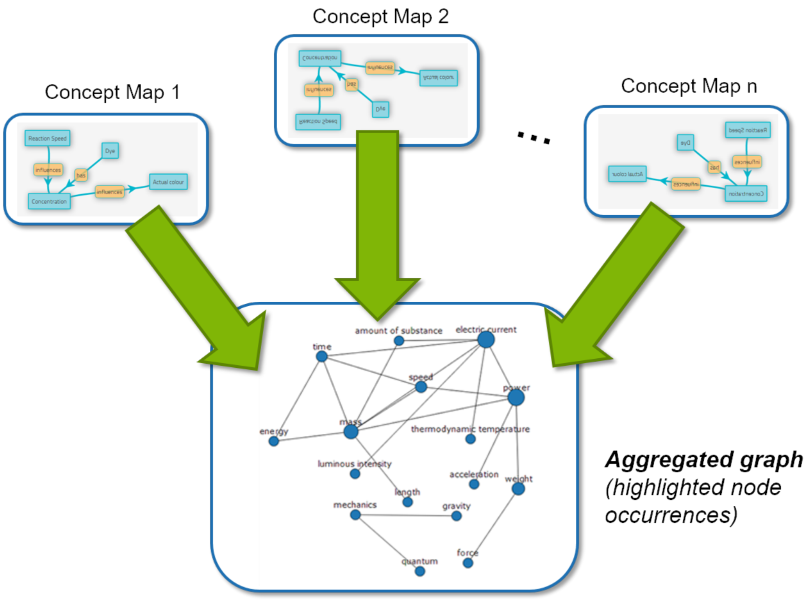


Figure 5.4: The conceptual model of the concept map aggregation.

students take this knowledge to create a new concept map capturing the parameter model of an osmotic power plant, while they are also confronted with some ecological factors of osmotic power and sustainability. Explicit assignments guide them through this scenario although they have to structure a final report by themselves.

Such a complex and multidisciplinary scenario, which incorporates different skills and competencies, possibly leads to a big diversity of the results. The students provide a non-standardized report as a final result, which does not allow a simple and automated assessment. However, the benefits are in the qualitative evaluation of the reports and the group observations, which shows that it is possible to track different competencies and to have a detailed view on the students' performances.

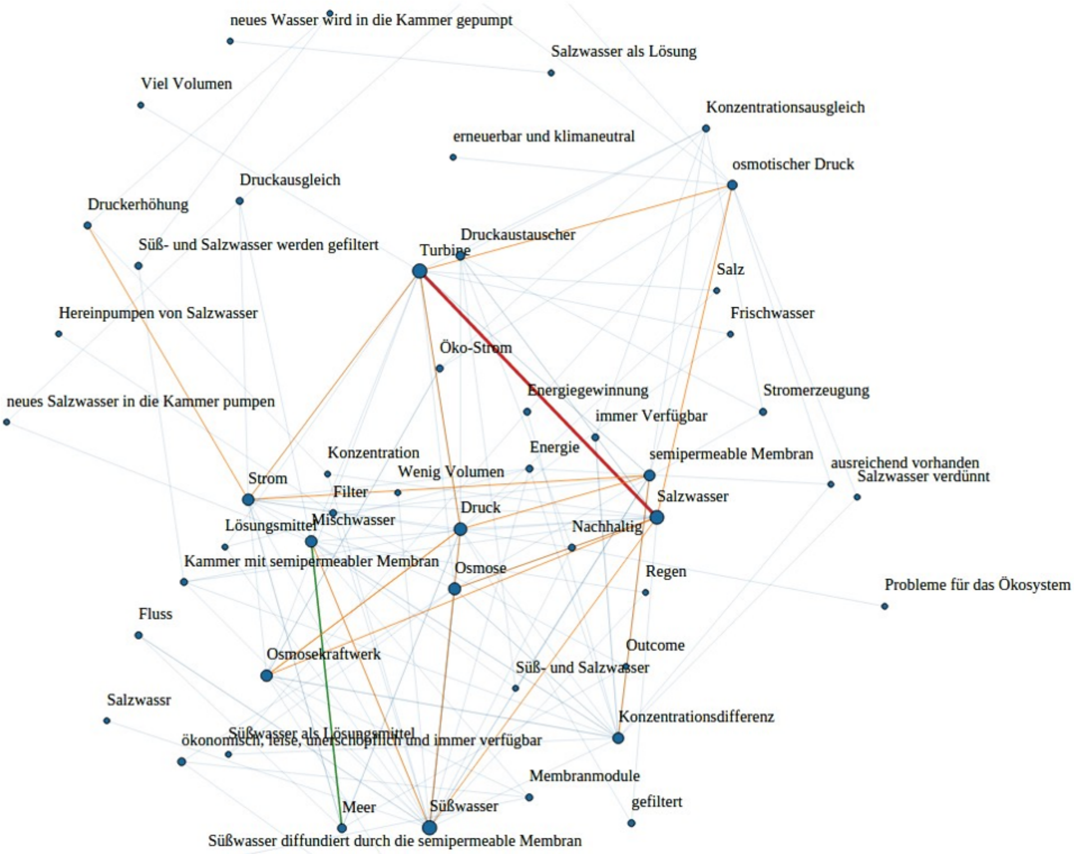


Figure 5.5: The aggregated concept map of the individual maps from the experiment.

5.3 Results

In this study we aimed to explore how group formation affects the practice of students and their performance in collaborative learning activities. To that end, we formed homogeneous and heterogeneous student groups using a multidimensional clustering schema based on artifact-related characteristics and personal traits such as motivational scores, as described above. In order to evaluate the practice of students we used both a qualitative (expert observations) and a quantitative approach (learning analytics). In order to assess the students' performance, we carried out pre knowledge and post knowledge tests. In the following paragraphs, we present the results of the analysis and discuss the findings of the study.

5.3.1 Quantitative Analysis

The interaction of students with the learning platform was recorded in log files. We used the log files to extract metrics of the students' activity and further explore any possible relation with qualitative characteristics and the overall knowledge gain. The scores of the knowledge tests ranged from 0 to 35 points and we used them to assess the learning outcome. Additionally, we defined the activity metrics portrayed in table 5.2 in order to evaluate the interaction of students with the learning platform.

Table 5.2: Activity metrics extracted from action log files.

category	***name***	***description***
learning platform	#actions	number of actions
	duration (min)	overall duration
	avgtimegap (sec)	time gap between consecutive actions (avg.)
concept map	#concepts	number of created concepts
	#relations	number of drawn relations
	#add	number of added objects
	#update	number of updates
	#delete	number of deleted objects

Table 5.3 shows the results of the knowledge tests per group. According to the results, the heterogeneous groups appeared to have a higher knowledge gain than the homogeneous groups. The heterogeneous groups improved their score in the post knowledge test on an average of 33% while the homogeneous groups improved their score about 20%. In the current study, group homogeneity does not ensure that the members of a group share similar knowledge background. For example, the members of group G2 that is considered heterogeneous, scored similarly in the pre knowledge

test (pre-STDEV = 0.5). On the other hand, the pre test scores of the members of group G6 that is considered homogeneous, portray a big deviation (pre-STDEV = 6.50).

Table 5.3: Results of the pre and post knowledge tests for the groups.

	Heterogeneous Groups				**Homogeneous Groups**			
	G1	***G2***	***G3***	***avg (het.)***	***G4***	***G5***	***G6***	***avg (hom.)***
Pre-test score	16.00	15.33	15.50	15.61	12.67	18.67	12.50	14.61
Pre-STDEV	1.41	4.78	0.50	2.23	2.05	1.70	6.50	3.42
Post-test score	23.33	23.00	23.50	23.28	16.17	24.00	15.00	18.39
Post-STDEV	3.40	4.90	0.50	2.93	3.32	1.47	5.00	3.26
Avg gain	7.33	7.67	8.00	7.67	3.50	5.33	2.50	3.78

The results of the knowledge tests were studied in comparison with the metrics of user activity. However, we were not able to draw any plausible conclusion for possible relations. The groups' activity, as portrayed in the log files of the learning platform, was similar for all groups (see table 5.4). A common hypothesis made in similar studies is that collaboration quality and knowledge gain are usually depicted in activity metrics, i.e. intense activity will lead to a solution of better quality (Kahrimanis et al. 2011). This hypothesis however was not confirmed in this study.

Table 5.4: Group activity metrics for heterogeneous and homogeneous groups.

	Heterogeneous			**Homogeneous**		
	G1	G2	G3	G4	G5	G6
#actions	30	57	56	38	56	60
duration (min)	23.82	48.42	21.67	16.77	25.67	29.75
avgtimegap (sec)	49.28	52.94	23.64	27.19	28	30.25
#concepts	13	28	24	19	23	29
#relations	10	25	28	12	23	27
#add	13	24	24	17	24	26
#update	12	24	28	14	23	27
#delete	0	6	1	2	2	4

5.3.2 Qualitative Analysis

During the group phase of the study, the users had to create a concept map based on what they had learned and to write a report. A teacher rated both the concept maps and the reports of the groups. This way, we wanted to ensure the findings from the pre and post knowledge tests. The concept maps were rated in a [0, 8] range and the reports were rated within the range [0, 12]. The ratings of the teachers for the concept maps and the work reports are presented in table 5.5. The results of the ratings with respect to group homogeneity confirm the findings of the knowledge tests. The heterogeneous groups are higher graded than the homogeneous ones for both the concept maps (21.4%) and the final work reports (29.6%).

Table 5.5: Teacher ratings of the concept maps and work reports per group.

	Heterogeneous groups				**Homogeneous Groups**			
	G1	G2	G3	avg(G1_3)	G4	G5	G6	avg(G4_6)
Concept Map Scores	4	7	3	4.67	1	2	2	1.67
Report Scores	10	11	6	9.00	2	6	4	4.00

The practice of the students was recorded in transcripts by two experts who attended the study. In addition, a third expert took general notes of the activity (e.g. notes about the timeline and events that might affect the activity). From the analysis of the transcripts, we identified three main group types: (a) Type A: One student operates the computer, the others comment or guide him verbally, (b) Type B: Group members change roles frequently regarding typing and directing and (c) Type C: One student is actively involved in the task, the others watch silently or do not pay attention.

Two out of three heterogeneous groups (G1 and G3) were identified as type B. The experts stated that even though they started out shyly, they managed to create a common ground and share responsibilities and tasks. They were enthusiastic about the activity until the end and seemed to enjoy it. For the group G3 in particular, the experts noted that they did not communicated openly (talking or arguing, etc.) and sometimes they were hesitant to act. Towards the end of the activity they did not interact between each other, but they carried on working separately even though they shared the use of the computer. This group scored highest in the post knowledge test and the maximum knowledge gain. The third heterogeneous group (G2) was identified as type C. According to the experts' observations, one particular student took over the activity but continuously tried to involve the other members by giving detailed explanations on every step of the process.

For the homogeneous groups, two were identified as type A (G4 and G5). According to the transcripts, the students of both groups were hesitant in the beginning. For group G4, it took them a considerable effort to start communicating and one student took action in order to move forward with the activity. In group G5 one student appeared to be more aggressive and active and dominated the activity from the start. Gradually all students began to participate within their groups. For group G4, however, it was too late to catch up while group G5 lost motivation towards the end. We think that the time the individuals needed to coordinate with the rest of the group members was critical and in the end this is depicted in the learning outcome. The third homogeneous group (G6) was identified by experts as type C. According to the experts' observations, one group member carried out the whole task while the other one was silently watching. Despite the fact that the active student tried to involve the other member in the activity, there was no collaboration or argumentation. Group G6 had the lowest score on average in the post knowledge tests and also had the least knowledge gain.

It is worth mentioning that the groups which were identified as type C had the maximum group deviation in the pre knowledge tests (Pre-STDEV, see table 5.3). This practically means that in both groups there was a "strong" student who eventually dominated the activity. However, in the case of the heterogeneous group the knowledge gain of group average was higher than in the case of the homogeneous group.

5.3.3 Discussion

In this work, the effect of group formation strategies on students' collaboration and the learning outcome have been discussed. Analysis showed that the heterogeneous groups increased their performance and appeared to have a higher knowledge gain than the homogeneous groups. On an individual level, the students who were members of heterogeneous groups had a knowledge gain of 33% on average while the students who formed homogeneous groups improved their individual performance of about 20% with respect to the pre tests. This finding was also confirmed by the teacher ratings of the concept maps and the group reports. Heterogeneous groups were graded higher than homogeneous groups for the quality of the concept maps they provided through the learning platform and for the quality of the written reports.

In order to assess the group activity with respect to collaboration quality, we used activity transcripts where experts recorded their observations. The experts stated that the students of heterogeneous groups adjusted their practice easier than the students of homogeneous groups. They undertook roles and responsibilities faster and without conflicts. Even in the case when they didn't seem to communicate on a satisfactory level, they managed to carry out the task efficiently. On the other hand, homogeneous groups needed more time in order to create a common ground and to collaborate effectively. In some cases it was even seemed to be impossible since some of the students lost interest and others were unable to carry out the task in time.

Additionally, we used the log files of the learning platform to define metrics of user activity. To that end, we followed popular approaches where activity metrics were introduced as indicators of good collaboration quality or efficient group practice (Kahrimanis et al. 2010). However, it was not able to prove any relation between activity metrics and the learning outcome. The group practice was similar in most cases with respect to activity metrics and group homogeneity. It should be kept in mind that, due to the study setup, one could argue that the activity metrics do not reflect group work or collaborative practice and therefore it should not be expected to find a correlation with the overall group picture.

5.4 Conclusion

To tackle the issue of how to engage students in sciences and capture their interest, we propose the usage of rich inquiry-based learning scenarios, as demonstrated in the Go-Lab project. Incorporating online learning with classroom presence leads to blended learning scenarios. This gives the opportunity to take the collaborative parts of the learning into the classroom, with all its benefits and challenges for the teacher.

We propose a way to support the group orchestration through the application of learning analytics, particularly the analysis of learning objects and assessed motivation. Finally, we conducted an experiment and applied methods of sequence and log file analysis to validate our hypotheses through multi-level analysis.

The analysis of the presented results indicates that heterogeneous groups outperform the homogeneous ones and achieve a higher knowledge gain. Thus, there is no benefit of choosing homogeneous groups in terms of performance. Even when having a group with only good performing students, they still do not perform significantly better than heterogeneous groups. Nevertheless, they don't compensate the weaker performance in the other homogeneous groups with weaker characteristics. For the class average, heterogeneous groups are better in sum, while it also covers basic principles of fairness, which is reflected by a lower diversity between the groups' performances. Fairness is both a principle that can influence the motivation of students in a further way but also underpins pedagogical decisions and thus is one of the important steps towards successful internal differentiation of learner groups. The results of this study cannot be generalized due to the small number of participants; however they can serve as indications for group formation.

6 Evaluation of the Concept Cloud App

Explicit visual representations of domain knowledge have the potential to support students who are engaged in scientific inquiry learning activities on an epistemic level. This can be facilitated by using computational methods for the extraction of concepts from student generated knowledge artifacts such as hypotheses, concept maps, or wiki articles. We propose an application of this approach in the context of inquiry learning with online science laboratories. The "concept cloud" is a scaffold to render cognitive information in the sense of an awareness tool. As a cognitive awareness tool, the “concept cloud” presents domain concepts and key phrases to the learners in order to help them reflect on their own learning and knowledge building processes. As part of a learning analytics tool set, the concept cloud also supports teachers in supervising their students’ knowledge building from an epistemic perspective. The approach has been tested in a classroom scenario with 84 secondary high school students. This section is based on a publication for the International Conference on Advanced Learning Technologies (ICALT) in 2016, which has been honored with a best paper award (Manske and Hoppe 2016).

Figure 6.1: Individual work with the Go-Lab environment during this study.

6.1 Experimental Setting

The concept cloud aims at supporting students' reflection and epistemic aspects. It is expected that students revisit and edit their generated artifacts and modify them after spending time on the concept cloud. Therefore, the sequence of actions have been analyzed in order to find these revision patterns. Additionally to the structure, a condition to ensure a temporal closeness to the concept cloud visit has been defined. Regarding these patterns, the time spent on the concept cloud, depending on the different experimental conditions, has been investigated as well as descriptive and activity-related statistics that might have implications on the learning activity. The learners have been assigned to several experimental conditions:

C1: no concept cloud (control group)

C2: concept cloud with DBPedia Spotlight extraction

C3: concept cloud with AlchemyAPI extraction

C4: static concept cloud (pre-rendered) without updates

Condition C1 is defined as the control group. In this condition, the learners did not see the concept cloud and therefore had no additional cognitive information in this learning space. To restrict possible confounding factors, hints to reflect about the generated content have been displayed and placed in the learning space in a similar way as in the other experimental groups. The conditions C2 and C3 use semantic extraction and represent cognitive information as described in the approach (compare section 4.3). In both conditions, different extractors were used: C2 uses the DBPedia spotlight engine, and C3 uses AlchemyAPI. The engine using DBPedia Spotlight extracts domain terminology for which a Wikipedia article exists, which is useful particularly in STEM fields. AlchemyAPI extracts key phrases from the given input. A more detailed description of the extractors can be found in section 4.2.3. C4 can be seen as a second control group, which also uses a tag cloud. Here, a non-interactive visualization that consists of a static, pre-rendered image of the concept cloud from the prior run of C2 has been used. Therefore, this condition shows a comparable visualization (on the content level) that does not correctly match the input from the learner-generated artifacts, although the tasks and the expected results are the same. However, differences are on the level of interactivity. The learners do not get any details such as the information in which phases a particular concept occurs, or a notice about the exact number of term occurrences across the learning group (only approximated as the concept size implies this).

6.2 Learning Scenario

The classroom experiment was conducted at a German secondary school in three computer science classes and three didactic lessons each. 84 students aged fourteen to eighteen years participated in the experiment. The objective of the learning scenario for this study was to learn about basic encryption algorithms and the complexity of decryption. Central to the ILS was an online lab to illustrate and practice encryption algorithms. Figure 6.2 shows the inquiry learning space used in this experiment.

The ILS has been tested in a classroom setting as a pre-study setting (n=15). In a previous version of the ILS, we included the concept cloud as a separate inquiry phase. During the Go-Lab project, there was a tendency to make such a reflection phase explicit in the design of an ILS and to provide on-demand tools for learners to reflect on their learning and their artifacts (Mäeots et al. 2016). However, we found out in the first trial that such on-demand tools have not been used more than superficially, although the learners have been prompted to use it (Schneegass et al. 2016). For a meaningful use of the concept cloud, we would have expected a certain retention time on the app, as well as activity triggered by the concept cloud. This could have been a revision of previously created content by the learners.

During the activity, each student created four short wiki articles, one concept map, and a set of hypotheses. These artifacts were used for the assessment of performance characteristics. The class was split up into experimental groups C2 (n=20), C3 (n=14), and C4 (n=10) and control group C1 (n=40). In the test conditions C2, C3, and C4, the concept cloud appeared in the ILS for the students. In the aforementioned pre-study we found out that it is necessary to force the students to actively engage with the app. Elsewise, if they only have it on demand in an additional phase that is optional for them, they might skip it. The control group C1 used a similar ILS without the concept cloud, but still with the instructions to reflect about previously created artifacts. The teacher was advised to monitor the classroom activity and to stay passive.

6.3 Results

In the ILS about cryptography that has been used to carry out this study, the learners had to create artifacts first. Afterwards, they were instructed to visit the concept cloud app that was embedded into the ILS design. After spending time on it, they could move forward in the ILS. The concept cloud has been calculated dynamically in conditions C2 and C3, whereas C4 has used a static, precalculated version of the concept cloud that is equal to the one from C2. We assume that the learners see a relatively similar concept cloud, due to the fact that the task progression in the lessons has been synchronized explicitly by the teacher. Although the conceptual model for

Figure 6.2: The Go-Lab ILS featured a cryptography lab and content about basic encryption.

the semantic extraction is the same for both conditions C2 and C3, the semantic extractors used are different. In C2, DBpedia Spotlight has been used, while C3 used AlchemyAPI. As a consequence, the information that has been presented to the learners was different. In contrast to the DBpedia engine (C2), the AlchemyAPI version (C3) is capable of extracting phrases that are not exclusively declarative. For example, phrases that encodes a procedural aspects like "needs brute force" (original: "benötigt brute force") or "26 trials" (original: "26 Versuche") have been extracted using this API. Visualizing such information encodes and transports another kind of knowledge (procedural). However, this extraction leads to a lower degree of aggregation, because the phrases cannot be matched due to more differences, only the phrase "alphabet" occurred multiple times. In total, 32 concepts (compound phrases; 71 terms in the visualization) have been displayed to learners in the AlchemyAPI version of the concept cloud for this study. In the DBpedia version, 22 concepts (25 terms) have been displayed where most of the terms occurred several times. A higher degree of aggregation leads to differences in the (visualized) size of concepts due to the number of occurrences. This might put a preference for learners to reflect on larger concepts while disregarding smaller concepts (e.g., "symmetric encryption"). However, these concepts might relate to the cases where most of the learners have problems with.

Contrary, a lower degree of aggregation poses the problem of dealing with a higher number of concepts, therefore, having a higher cognitive load when using the concept cloud.

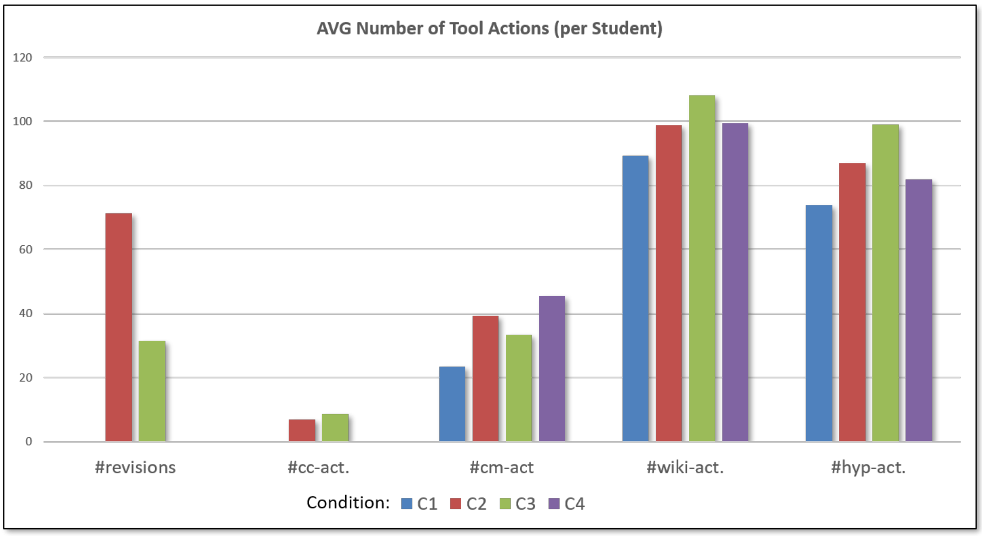


Figure 6.3: App-related user activity: number of revision patterns, concept cloud, concept mapper, wiki and hypothesis tool actions.

An important aspect when dealing with these explicit knowledge representations is to find out if students use them to cheat the system. For example, they could use all the important concepts that have been used by the majority to create artifacts such as concept maps that contain these key concepts without a deeper level of understanding or reflection. Therefore, we analyzed students' action sequences to detect this behavior. Indicators for this are certain revision patterns that involve productions. We investigated revision patterns (a) without non-productive actions, (b) with production-only actions, (c) patterns separated per artifact, (d) time spent on each revision, (e) concept occurrences in the concept cloud and each artifact, and (f) sequences of added and revised concepts.

These metrics and indicators have been applied to the sequence of learners' activities. A revision pattern occurs, when in the action sequence after the visit of the concept cloud (cc) a jump to an prior tool occurs (toolVisit). However, this broad definition of a revision pattern is chosen, to include revisions that show more a revisiting behavior, where learners go back to an artifact after they have seen the concept cloud and then spent some time to align their mental state with the artifact without changing it. Metric (a) detects such action patterns in the semantic of [cc, toolVisit, phaseChange], which indicates such a case of a non-productive revision as mentioned before, when considering the metric for the time spent (d). For the time-based metric, a threshold has been defined according to the observation in the classroom. A student who does not interact with the concept cloud leaves the visualization in less than 30 seconds. A productive action would be marked, depending on the tool, as [cc, toolVisit, tool+], for adding data such as a concept to the tool. Exemplary, the following revision pattern has been extracted for a learner: ["WI+","CC","HY+"]. This encodes the sequence of adding a new revision for the wiki and then switching to the concept cloud app. The CC only occurs in the sequence, if the indicator observed a certain time spent on the concept cloud. Afterwards, the learner goes back to the hypothesis scratchpad and adds a new version. As the hypothesis occurs prior to the concept cloud app in the sequence of phases, this indicates that the learner switched back in the phases and revised the set of hypotheses in the scratchpad. In the data set, the patterns have been isolated for each tool (c) in order to distinguish a more general revision behavior from a targeted refinement of specific artifacts. Indicator (f) represents the revision sequences for each tool. However, this analysis is only possible in the groups, where revision patterns could be extracted (C2 and C3). The control group (C1) and the pre-rendered concept cloud (static; C4) do not provide interactivity, user actions could not be tracked in those conditions. Therefore, a measurement based on the action logging was not possible in these groups. Although hypotheses play a major role in IBL scenarios, there are no relevant findings regarding the artifact analysis of hypotheses for this experiment. We could not observe any significant findings regarding revisions after using the concept cloud, because none of the hypotheses that have been created previously have been discarded.

When investigating the produced artifacts, it turned out that students who used the concept cloud in the proper way created better concept maps regarding the coverage of key concepts. We argue that a meaningful usage pattern would incorporate time and certain consecutive actions. An ideal use of the concept cloud would be the following example: a student creates an artifact during the learning scenario that is later followed by a reflection period when she uses the concept cloud. After a certain time, she revisits the artifact and a) leaves it because she is satisfied with the relative quality compared to others, or b) edits the artifact and proceeds with the learning activity. A problematic pattern can be characterized as an oscillation between artifact modifications and the concept cloud, which indicates the aforementioned “cheating the system”-behavior. However, not a single case could be observed using this pattern on the data from the experiment.

When we compare the ideal use cases of the concept cloud with the non-ideal cases leaving out the “mock”-condition C4, (50 non-ideal, 14 ideal cases), we observe major differences in the artifact creation. Concept maps have been created after using the concept cloud. This implies that the results of the terms visualized in the app might influence the creation of the map. The average number of concepts per concept map is with 5.04 significantly lower compared to the average of 11.29 in the ideal group ($p < 0.01$, unpaired two-tailed t-test). In all of the cases of ideal usages, the concept maps were only with a few exceptions composed of relevant terminology according to the learning design (defined by the teacher), which implies that the concept map quality in terms of coverage increases proportionally to the number of concepts. In conclusion, concept maps created by “ideal” concept cloud users are likely to be better than the maps from students who did not spend any or at least not enough time on the concept cloud.

According to Wilson (1996), the development of epistemic fluency, and the ability to become epistemically fluent, is potentially supported through rich information sources. The concept cloud can be seen as such a rich information source, but only when the learners participate actively. This seems to be in line with the findings about generative knowledge construction from Wittrock et al. (1990). Concept mapping scenarios can be enhanced using the concept cloud application as an additional cognitive scaffold. The finds from the experiment indicate that this combination supports learners’ knowledge construction. Following this, the concept cloud is likely to be a useful scaffold in combination with other production tools or inquiry apps.

6.4 Conclusion

In this work, the use of semantic technologies for learning analytics, particularly to support the guidance of students, has been described. This section presents the evaluation of the concept cloud app, which has been designed based on the conceptual framework presented in section 4. It facilitates the concept of semantic extraction from learner-generated artifacts and aggregates knowledge items into a shared group knowledge model. Therefore, the concept cloud app aims at supporting knowledge construction and reflection on the part of the students. For teachers, it gives insights into knowledge structures of the students in a learning group and therefore it serves as a learning analytics tool for them.

The idea of content-related reflection is facilitated when it comes to the idea of teacher-led inquiry. Teachers perceive and design their own teaching as an experiment, similar to an inquiry-based learning approach and further develop their teaching materials (Emin-Martínez et al. 2014; Clark et al. 2011). The concept cloud app supports this idea with the information presented. In contrast to "regular" mode where learner artifacts are processed, the concept cloud can be used to visualize the extracted concepts from teaching materials. This supports teachers in the supervision and reflection of their own materials, which might positively influence their planning and orchestration of learning materials by reviewing the usage of specific concepts across the different inquiry phases. This has the potential to uncover potential inconsistencies about key concepts that are relevant for a didactic unit or across different lessons. Future evaluations need to prove the use of this application to support teacher-led inquiry and its impact on the teaching quality.

The conducted experiment demonstrated how learning analytics applications in conjunction with inquiry apps support and scaffold learners in their knowledge construction, and teach them epistemic fluency on their way to become 21st century learners. This idea is in line with the concept of relating multiple representations. Further, this concept comprises the active integration of static and dynamic information into external representations. Bodemer et al. (2005) found out that this results in better performance and potentially leads to a more systematic and goal-oriented experimentation in computer-supported inquiry-based learning.

7 Evaluation of the Semantic Group Formation and Group Awareness

One thread of existing research on small group learning has addressed positive effects of grouping learners with complementary knowledge, another one has focused on representing and visualizing knowledge distributions to improve cognitive group awareness. Although a combination of both seems obvious, this has not yet been investigated. Semantic group formation and cognitive group awareness are two concepts that go hand in hand. The grouping is based on learners knowledge complementarity, which aims at supporting the initialization of collaborative processes, of grounding phases, and acts as a transparent and explainable model of the grouping. The information that is responsible for the results of the algorithm, namely the group knowledge models, can be used to form the basis of cognitive group awareness tools (*CGAT*). Thus, the work presented in this section, combined both approaches, namely the semantic group formation as described in section 4.4.2 and a variant of a grouping and representing tool (Erkens et al. 2016a).

An experimental study in a real classroom setting to investigate the effects of support and the level of visualization of co-learners' knowledge has been conducted. This has been evaluated in a 2x2 mixed design approach with the following dimensions: The levels of support have the conditions with group awareness and without group awareness; the level of visualization is high or low.

The work presented in this section is based on the publication in the proceedings of the CollabTech conference (Erkens et al. 2019). A detailed observation of one third of the learning groups (6 out of 18) is presented with an analysis of the communication and coordination behavior of the knowledge exchange, where the learners gain cognitive information throughout the group awareness tool. In addition, this thesis presents an evaluation of the semantic group formation regarding the goal criteria set for the semantic group formation, as described in section 4.4.2. This is driven by the question, if semantic extraction has the potential to approximate and represent knowledge accordingly. As a hypothesis, the distributions of the targeted score (knowledge diversity) are fair and even, if this information is used to form groups. Further, it has been qualitatively examined to what extent learners with group awareness support differ from learners without such support regarding the structuring of their communication and to which extent it can be traced back to the level of own knowledge and to the knowledge distribution in the learning group.

Figure 7.1: The second phase was organized as group work with a shared resource.

7.1 Experimental Setting

The study has been conducted in a real classroom setting with 59 high school students of a German upper secondary school. The 59 students have been distributed across three classes of different sizes. Each class has been supervised by the teacher and three researchers. However, the materials in the Go-Lab environment were intended to serve as a self-regulated learning scenario. The learners used the Go-Lab learning environment during the lessons. The sample had to be reduced to 45 after excluding certain students due to absence in one of the parts (26 men; 19 women; mean age: $M = 16.33$, $SD = 0.67$).

The study was split up into two sessions, each of them had a duration of 90 minutes (two lessons). The first session was organized as an individual phase using the Go-Lab environment. During the work in this environment, the learners created artifacts such as concept maps and wiki texts (cf. section 7.1.1). This was mainly intended for the learners to acquire knowledge and to collect data as input for the semantic group formation, and thus, for the cognitive group awareness tool. The second session was organized as a collaborative phase. The learners were assigned to the groups from the semantic group formation and shared one access to the Go-Lab environment. Therefore, most of the collaboration took place inside the classroom, while the learner-generated artifacts were stored in the Go-Lab portal. In the beginning of the collaboration phase, the learners used the group awareness tool and had an explicit knowledge exchange phase. Figure 7.2 shows the overview of the study design with respect to the sessions and to the artifacts created by learners. The student-generated artifacts have been used as a basis for the second session: (1) to apply the semantic group formation algorithm, which forms groups for the collaborations in the second session, and (2) to visualize the knowledge of all learning partners (see section 7.1.2).The schema of the two phases is inspired by the study presented in section 5. The main differences are the used group formation algorithm (semantic group formation) and a more explicit knowledge exchange phase based on the group awareness information.

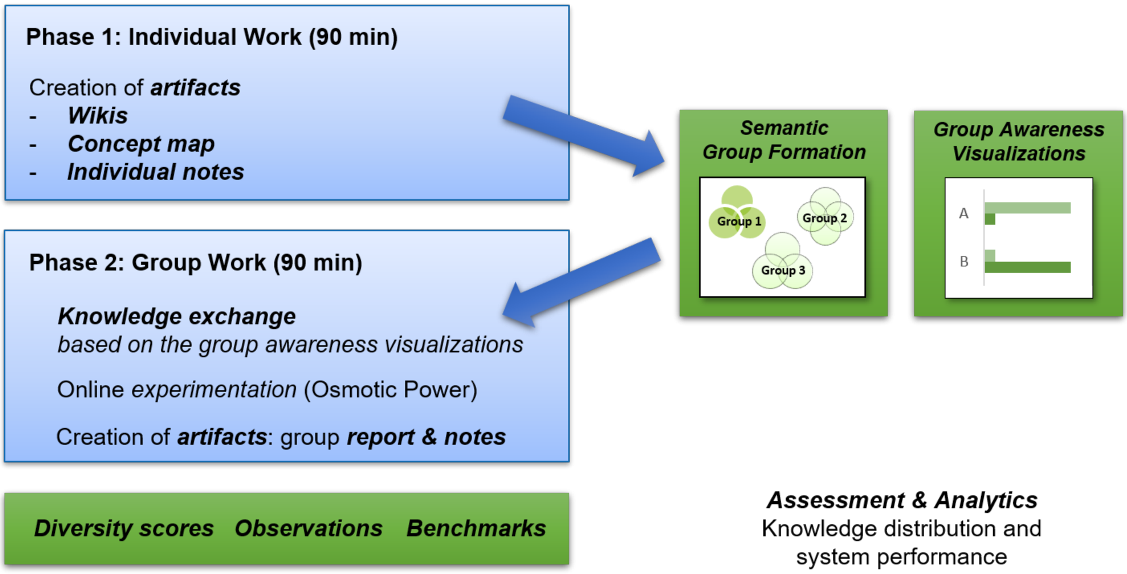


Figure 7.2: The schema of the two phases and the data collection in this experiment.

For conducting the study, the students have been randomly assigned to the control group or experimental group, and have been further divided into dyadic or triadic learning groups based on the semantic group formation algorithm that has been described in the section 4.4.2. Due to the limited ICT resources in the school (technical defects of computers and network equipment) it was preferred to build triadic groups in the larger cohort. To test the hypotheses, a 2×2 mixed factorial design with randomly assigned group membership as a between-subject factor ('CGAT' vs. 'No CGAT') has been used. Learners were asked to self-assess their questioning and explanation behavior (as a dependent variable). Additionally, a detailed observation of three learning groups from each condition regarding their respective sequences and communication patterns has been made in order to verify the hypothesis that the group awareness information presented has an impact on the collaboration.

7.1.1 Learning Scenario

In the first part of the study, each student had access to a prepared inquiry learning space in the Go-Lab environment (de Jong et al. 2014). The specific learning space engaged learners with a project on the issue of energy transition and renewable energies. As a motivation, the learners took the role of a member of the city council of "EnergyCity", a fictional city. The learning scenario was driven by a decision problem, where the learners had to decide whether EnergyCity should build an osmotic power plant. This design is similar to the scenario of the skill-based group formation approach, described in section 5. Changes were made in order to increase the knowledge diversity and to have a better integration of the group awareness information into the learning design. The narrative of the individual phase was that the learner will be interviewed by a local newspaper. The interview was intended to act as a preparation for the decision, but also to introduce argumentative components to foster a discussion and to prepare arguments in advance. To have grounded arguments, the students were provided with learning materials on several related topics, such as osmosis and renewable energies.

Individual phase The tasks of the individual part have been structured as follows: First, learners were asked to write down answers to certain interview questions of a local newspaper. These questions were related to more general topics such as "renewable energy" and "climate change". In parallel, the learners had access to the learning materials. Furthermore, the learners had to create a concept map on "diffusion and osmosis", which is about the physical background of the osmotic power plant. This was necessary to prepare the work with the simulation, as some of the parameters were not part of the standard curriculum, for example "permeability" of a membrane. Afterwards, the learners had to write another short text on general assumptions of the possible functioning of osmotic power plants. The tools to create (wiki) texts are, in contrast to the previous version of Go-Lab, always present as an overlay that can be hidden on demand (see figure 7.3).

Collaboration phase In the second session, which was conducted one week later, students had another 90 minutes to learn collaboratively in their assigned learning groups. In contrast to the individual phase, each group had access to a single computer. Therefore, the collaboration was situated in-class, but the products were provided within the web-based Go-Lab environment. The instructions presented in the ILS posed several collaborative tasks.

At the beginning of the collaboration, the learners have been instructed to had exchange their knowledge about certain topics. The goal of this phase was to create a shared understanding within each group about osmosis, green and blue energy,

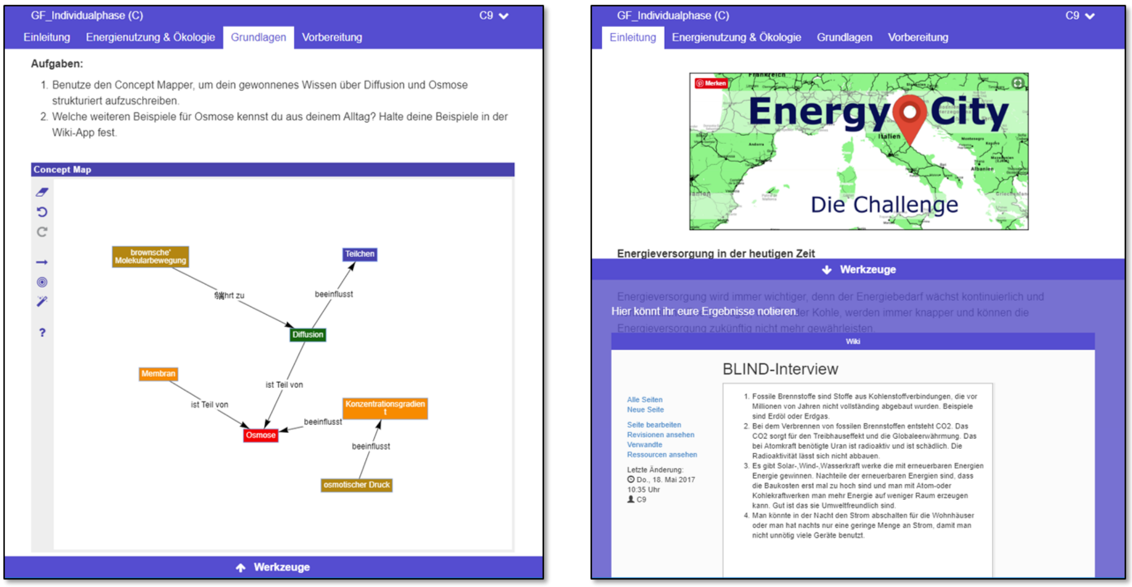


Figure 7.3: The learners used the Go-Lab environment and performed concept mapping and text writing tasks within the ILS.

but also to exchange opinions. Based on the experimental condition (with/without group awareness) they were supported in this process with a group awareness visualization (for an example, see figure 7.5). This visualization presents the knowledge distribution, for instance a quantification of the relative knowledge differences of the co-learners in each topic, visually represented as a bar chart. The control group also received a visualization in form of a topic list, but without the bar chart. Leaving out the visualization completely inhibits the risk of producing possible confounding factors. This phase was restricted to 10 minutes. Afterwards the learners were instructed to collaboratively write a text on the relevance of salt and fresh water for osmotic power plants. For the preparation of the experimentation in the IBL setting, the learners had to create hypothesis regarding the use of an osmotic power plant and the effectiveness of the energy production. They received guiding questions to support the experimentation in the laboratory. Finally, the decision problem framing this scenario could be concluded in a final assignment, where the learners wrote a statement for the city council as a wiki text.

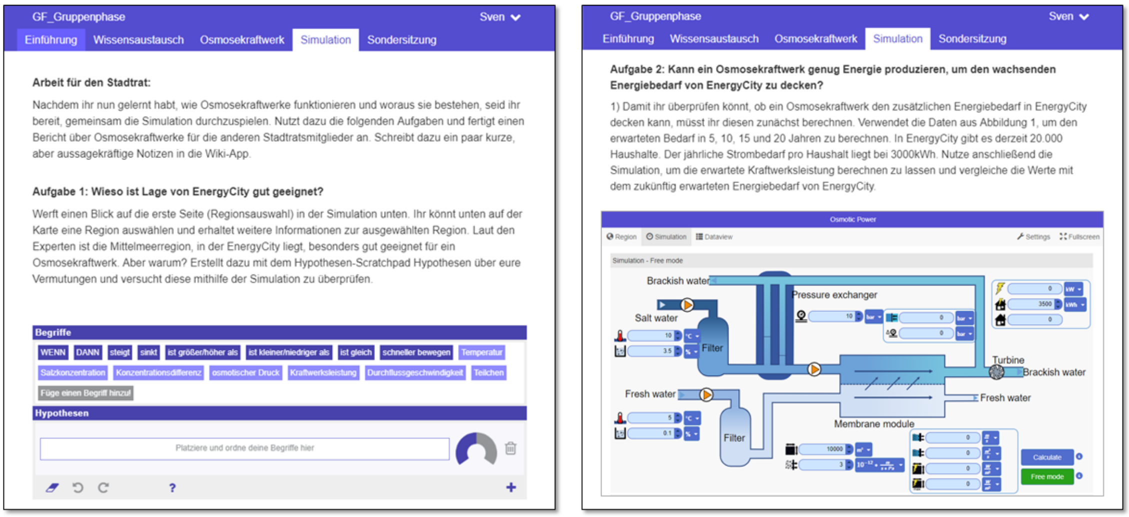


Figure 7.4: During the group phase, the learners had to create hypotheses (left) regarding the osmotic power plant and to solve tasks using the simulation (right).

7.1.2 Group Formation

Grouping of learners with semantic group formation

The grouping is based on the algorithm of semantic group formation described in more detail in section 4.4.2 and published in Manske and Hoppe (2017). The algorithm takes the learner-generated artifacts as an input and creates an optimal grouping as an output. Therefore, it uses the presented architecture to automatically retrieve all artifacts, to process them, and to extract knowledge items using semantic technologies for keyword extraction (cf. section 4.2.3). Although the use case for text extraction is similar to the concept cloud approach, there are some changes in the configuration. The extraction used was an NTA-based extraction using a domain dictionary. It has been enriched by an ontology, which added synonyms to the dictionary. The dictionary-based approach was necessary to render the concept-topic-relation, which is used for the group awareness tool. The initial dictionary was hand-coded in order to create a gold standard for the evaluation of text analytics approaches. This dictionary includes the topics. A topic, can be seen as a labeled category in this context. This induces a relation, where a certain concept is subsumed by the corresponding category. Also, synonyms have been added to the dictionary in a preprocessing step using DBPedia. The dictionary is simply structured as triples $(term, concept, category)$, for example ("CO2","carbon dioxide","Consequences"). This renders "CO2" as a synonym of the concept "carbon dioxide", which is grouped at the category "Consequences".

Based on the extracted knowledge items, the semantic group formation algorithm calculates the optimal grouping according to a diversity measure as described in section 4.4.2. This resulted in 18 learning groups with a complementary knowledge distribution: eight groups in the control condition, of which four were assigned to triads and four to dyads, and ten groups in the experimental condition, of which six were assigned to triads and four to dyads. Further, the results have been used to visualize knowledge distributions.

Knowledge Exchange

During the individual phase of the study, the learners created artifacts, which have been processed by the semantic group formation. The group formation produces an optimal grouping based on knowledge diversity. According to this, the knowledge distribution in each group is diverse. For the second phase of this study, the group awareness visualization is constructed from the underlying data model, which has been created through the semantic group formation. This emphasizes the seamless connection between the model for the grouping algorithm and the cognitive information that is displayed as a group awareness tool. The visualization is represented as a grouped bar chart, which displays the knowledge distribution within each learning group. This provides more detailed information compared to presenting the pure diversity score only. Each group of bars represents a the distribution per category (topic) in the knowledge model. For each category, a certain number of concepts are part of the overlay model. The bar length represents the relative number of concepts matched by the particular learner. The length has been normalized according to the maximal values inside each learning group.

As the experiment uses the group awareness support as a dependent variable, two conditions that influence the information provided have to be distinguished:

1. *CGAT:* Learning groups in the experimental condition had additionally information on their learning partner(s) available. This is presented as bar charts displaying the knowledge distribution for each topic.
2. *No CGAT:* As members of the control group should exchange their knowledge without group awareness support, they were provided with a (static) list of topics corresponding to the individual learning phase. This corresponds to the labeling of the y-axis in the group awareness visualization.

Additionally, concrete absolute or relative (percentage) numbers have been removed from the chart. Otherwise learners might have difficulties with the concrete quantification of the items. By design of this study, it is more relevant to see if there are knowledge differences in a particular topic, on which the learners can negotiate and discuss. Thus, focusing rather on the relative differences to learning partners than

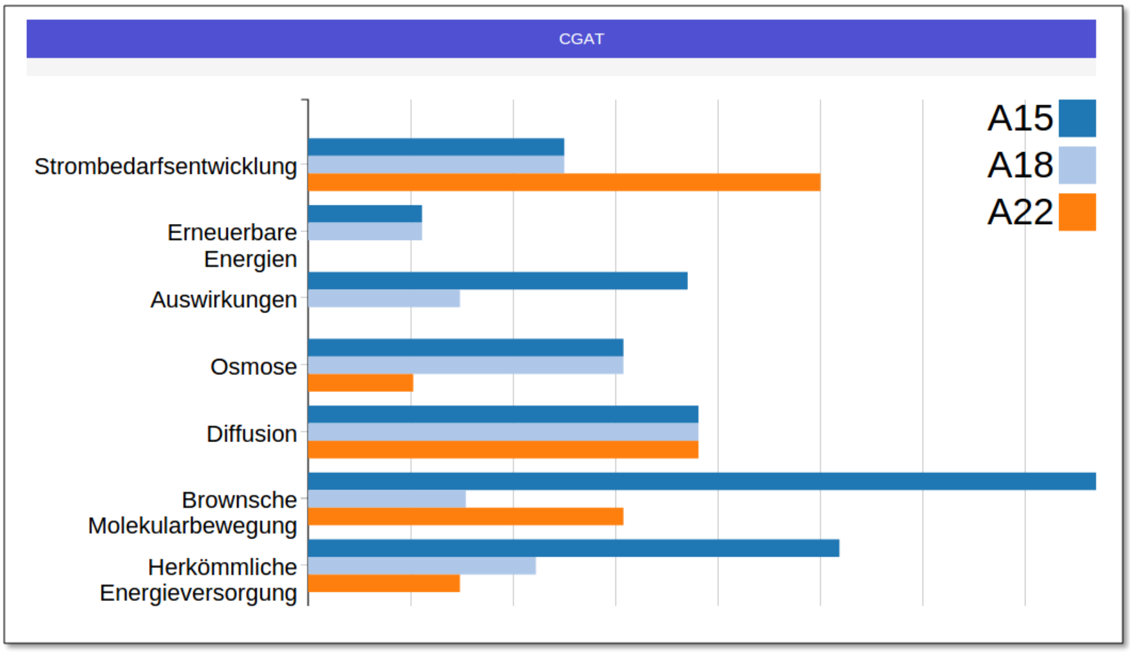


Figure 7.5: Example visualization of a learning group from the experiment with cognitive group awareness support.

on the correct quantification of the algorithm might help to reduce the cognitive load for the learners and prevent a possible bias that results from challenging the system. Figure 7.5 shows an example of the group awareness tool from the experiment, which displays the knowledge distribution of a triadic group.

7.1.3 Observation

During the knowledge exchange, 6 out of 18 groups have been observed in detail to learn more about their collaboration processes, in particular regarding the group awareness information. This was imposed by the question, whether the information influences or fosters the structuring of the knowledge exchange and if this is dependent on a certain knowledge level and a knowledge distribution. In each of the classes, one group with CGAT support and one group without support ('No CGAT') has been selected randomly. For the coding, one researcher has been assigned to a fixed group and evaluated each utterance of the students with regard to their type. For the coding, there were several differentiations:

- *type*: question (Q) or explanation (E),
- *purpose*: content-related (C) and organizational (O),

- *topic*: the particular topic / item in the visualization.

The codings have been transcribed in a temporal order as a sequence of timestamped actions. For example, if student S1 asks a content-related question on osmosis (Os) and student S2 answers this question. Afterwards, S1 explains that they will have to write this down into a wiki tool. In this case, the coding is S1: QCOs, S2: ECOs, S3: EOOs. Using the coding of sequences helps to answer the research questions that have been posed:

1. To which extent influence group awareness tools the learners' structuring of communication?
2. Does the level of knowledge or the knowledge distribution in a group influence the communication and structuring?

Regarding the first research question, an observation of the actual communication sequences can be useful. It can be reasonably assumed that the learners usually structure the communication about the topics in the chronological order of their appearance in the tool. The deviations from this order might give hints if the tool influences learners' self-regulated determination of sequencing. It is questionable, whether these deviations are imposed by other factors or patterns that emerge from a certain level of knowledge or a knowledge distribution.

To investigate the first question, charts representing this sequencing have been created. These line charts represent the order, in which students have talked about topics in relation to the position of the topic in the group awareness tool. For the second question, the number of coded questions and explanations in combination with the level of visualized knowledge has been analyzed. This can be used to identify, whether a learner had a low displayed knowledge and asked many questions or vice versa. The results from the observation are described in section 7.3.

7.2 Results from the Semantic Group Formation

Semantic group formation is quite different compared to more traditional approaches of algorithms, which create heterogeneous and homogeneous groups. To evaluate the algorithm, the goal criteria set by the design need to be revisited. Three goal criteria for the semantic group formation are given:

1. *Optimal diversity* of the grouping.
2. *Fairness* of the distribution of diversity scores.
3. A good approximation of *learner knowledge* through semantic extraction (*text extraction quality*).

4. *Monotonicity*: with a higher number of students given as an input, the average diversity increases.

The first criterion, the optimal diversity, is related to the pedagogical specification of this approach. Groupings are created and favorized, if the learners are complementary regarding their knowledge. Therefore, the term knowledge diversity has been introduced and operationalized. The algorithm fulfills the first criterion, if it creates an optimal grouping regarding the combined diversity of all groups. As described in section 4.4.2, the diversity score of the grouping is calculated as the product of all groups' diversity scores. However, in a real classroom setting, algorithms such as this have to deal with a relatively small number of students, and therefore they will have to create a few small groups. As a consequence, the algorithm has been designed to select a global optimum rather than a local optimum (like a greedy algorithm would select this). Although the asymptotic approximation of the run-time complexity is quite high, there was no need to optimize the run-time behavior in order to find an optimum at all. As the algorithm scans all possible partitionings and calculates scores for all groupings, the proof of the optimality is trivial and therefore omitted in this work.

The second criterion tackles the differentiation between the diversity of the grouping and the groups' diversity and is related to the fairness of the algorithm. Although it creates a globally optimal grouping, there might exist groupings where some groups have a high diversity score while others don't. Therefore, an evaluation of this fairness criterion has been performed in the next section 7.2.1.

As a third criterion for the quality of the group formation, the knowledge extraction quality needs to be investigated. The assumption that the algorithm can be used to approximate the learners' knowledge through semantic extraction is questionable to this extent, where the sufficient and correct concepts can be extracted from the text. Only if that is guaranteed, such a mechanism approximates the learners' knowledge, which is a prerequisite for the cognitive group awareness. If the extraction does not work properly, the group awareness information displayed to the learners is also incorrect, which affects the pedagogical design. The evaluation of the text extraction quality is shown in section 7.2.3.

7.2.1 Distribution of Diversity Scores

The semantic group formation algorithms optimizes the general diversity of the whole cohort. The diversity of the cohort is defined as the product of all group diversities. With the assumption that a higher diversity leads to a better learning situation (in analogy to Jigsaw approaches), it is aspirable to reduce the probability of outliers regarding diversity in a grouping. This is in line with the goal setting to create a more inclusive and fair algorithm. The term unfair refers to an imbalance on the individual level or for a given minority, e.g. a single group. Thus, such an algorithm should not penalize a certain group to satisfy calculatory conditions, for example, to have a better overall score. Therefore, the formula for diversity has been derived in a way that it is stable regarding outliers. However, this aspect is evaluated by analyzing the distribution of diversities. Figure 7.6 shows the distribution of the diversity scores of the all groups from the classroom experiment. The diversity scores for all groups are situated densely around the average, without having outliers. Although the experimental results presented in this figure show that the algorithm does not scatter the diversity scores of the groups a lot, it is questionable if this statement can be generalized. Therefore, the next section presents a benchmark that investigates the question of generalizability of the criterion *fairness*.

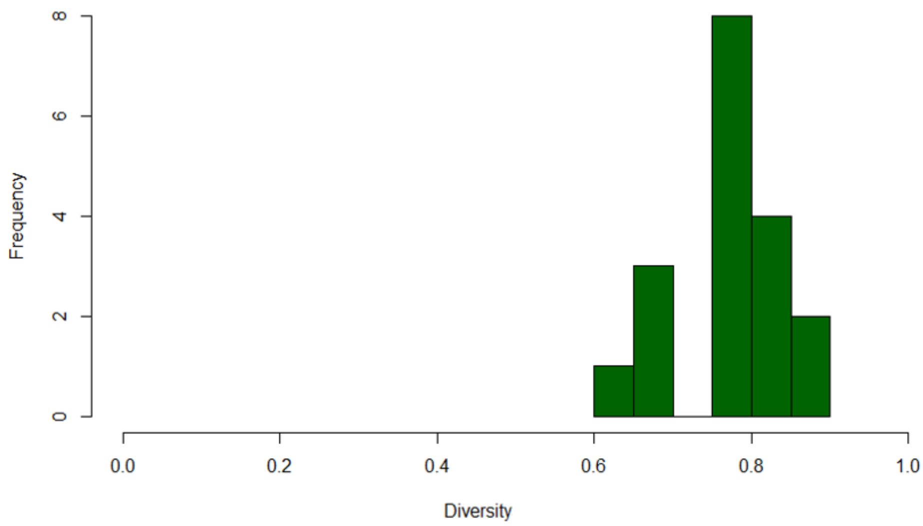


Figure 7.6: The diversity scores per group are not scattered, but distributed closely around the average.

7.2.2 Benchmark of the Diversity Distribution

Although it seems counter-intuitive, there is a focus on simulating smaller cohorts, because the degrees of freedom, namely the number of possible partitionings are much smaller. Therefore it is less probable to find a good solution according to all the criteria with a suitable diversity score. This evaluation has been done using a synthetic benchmark that emulates the classroom setting for a particular cohort size, ranging from $n = 6 \dots 11$, with a variation in the size of the envisioned grouping in dyads or triads. In the cases where it does not add up (e.g. no multiples of the group size), the algorithm balances with the next possible grouping size not equal one.

The benchmark itself is a repeated simulation of the group formation. A test set for the overlay knowledge model in the size similar to former experiments has been created (30 concepts in total). In each run of the simulation, the learners were assigned a random set of knowledge items as a subset from the overlay model. An important output parameter is the diversity of each group in the grouping. According to the aforementioned goal criteria, a good grouping has quite similar groups in terms of performance characteristics. The performance measure for the semantic group formation is diversity. An unwanted grouping will be any grouping with a high diversity of the grouping, but a big variation between the diversity of the inherent groups. For each parameter (group size ranges from 2 to 3, number of students ranges from 6 to 11), the simulation performed 100 group formation runs to create the data set.

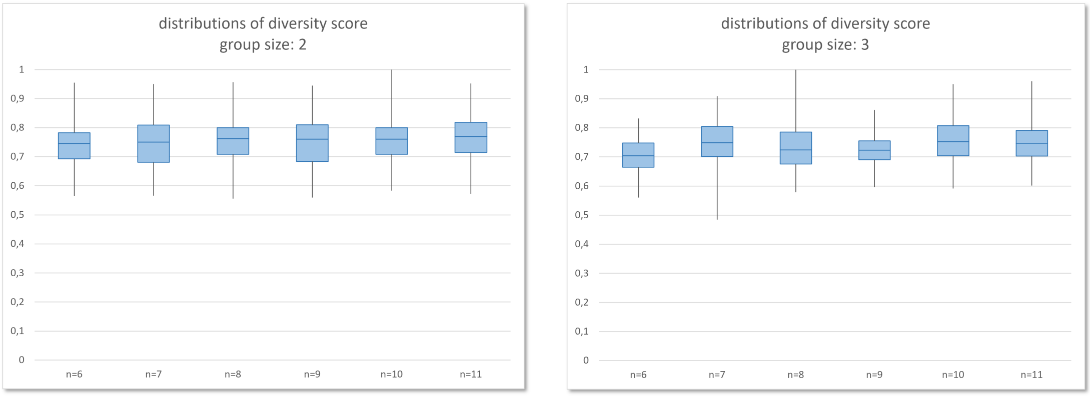


Figure 7.7: The diversity scores per group are not scattered, but distributed closely around the average.

Figure 7.7 shows the distributions of diversity scores across the groups depending on the envisioned group size. Each box plot displays the median and error indicators that characterize the distributions. First of all, it is positive that there are no outliers. Most of the values for diversity are in a good range.

Besides the question of fairness regarding the distribution of diversity scores across groups in a grouping, there is another aspect that should be considered. Particularly in the field of Cognitive Group Awareness research, it is usual to group learners in dyads, because this makes partner modeling more simple and puts the focus for every participant on one partner (Erkens et al. 2016a).

Although the figures presented above indicate, that there is not a big difference between the groupings into dyads and triads, figure 7.8 confirms this hypothesis. Particularly with an increasing n, both averages have a tendency to increase and the standard deviation remains in the interval between $\sigma = 0.051$ and $\sigma = 0.088$. The average is higher in most cases for the dyads, which is plausible because the number of possible groupings is higher for this case. This is in line with the assumption that with a higher number of students the algorithm will produce a more diverse grouping (monotonicity). The biggest differences in the comparison of averages and standard deviations can be observed for the cases n=6 and n=9. This is caused by a relatively big difference in the number of possible groupings due to the number of groups depending on the group size.

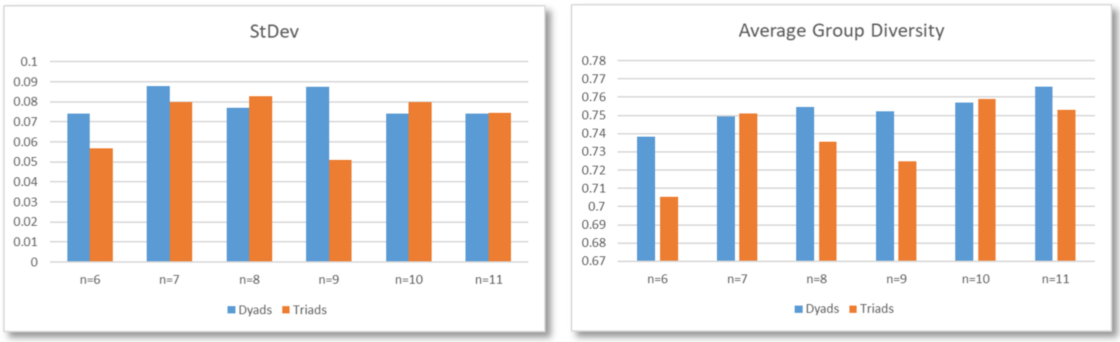


Figure 7.8: Average group diversity and standard deviation of all runs, triads versus dyads.

7.2.3 Text Extraction Quality

Evaluating the quality of the text extraction is crucial for this work. Between the two experimental parts of this study, there is a high interdependence. The Semantic Group Formation forms groups and outputs a shared group knowledge model, which is used to create the group awareness information. Both steps rely on text extraction, because the input of the algorithm is a set of learner-generated artifacts, such as texts or concepts maps. In order to create the knowledge model, these artifacts are processed (cf. section 4.2.2) and relevant concepts are extracted (cf. section 4.2.3).

Texts created by 22 students from the experiment have been used to create a manual coding which serves as a gold standard for this benchmark (cf. section 7.1.2). In relation to this the quality of automatic semantic extraction approaches is checked in comparison to the gold standard. To assess the quality of the text analysis approaches, recall, precision and the F-measure have been used on the sets of extracted concepts of each method (automatic extraction) compared to the set of relevant concepts (manual coding). Precision quantifies the positive predictive value ("true positive accuracy"), which is the fraction of relevant concepts among the extracted concepts. Recall, sometimes called sensitivity or true positive rate, is the fraction of relevant concepts extracted compared to all relevant concepts. The F-measure is a weighted harmonic mean of precision and recall, which combines both aspects. However, in a context of computer linguistics and the evaluation of semantic extraction, recall is

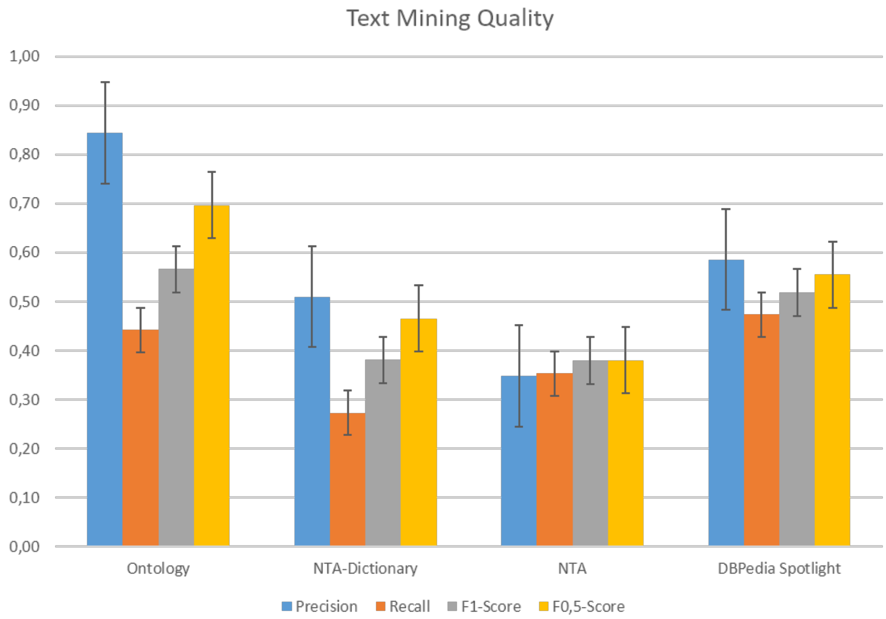


Figure 7.9: The used text mining shows the highest F1 and F0.5 scores compared to related approaches.

often seen as more relevant for benchmarks as it highlights the correctness of extraction methods by quantifying the rate of true positives in all positives. Therefore, the F0.5-Score has been employed, which is a weighted harmonic mean that doubles the impact of recall. The text analysis approaches used for this benchmark are: (1) network text analysis ('NTA'), (2) dictionary-based NTA ('NTA-Dictionary'), (3) ontology-enriched NTA ('ontology'), and (4) DBPedia Spotlight. The ontology-enriched NTA is a regular NTA which uses an ontology created by domain experts in order to increase the accuracy of the automatic semantic extraction. The ontology encodes the domain knowledge structured as synonym-term-category triplets in the domain of the learning context. More details about the DBPedia Spotlight extraction can be found in section 4.2.3. The results (cf. table 7.1) indicate that the ontology-enriched NTA performed best in precision (84.4%), recall (44,2%), and F-measure (F0.5-score 56,6%, F0.5-score 69,6%). Figure 7.9 shows a box plot of the comparison of the different methods.

7.3 Results from the Group Awareness

The main contribution of this part of the thesis is defined in the (technology-based) approach of facilitating knowledge diversity. Particularly for this chapter, this is achieved

Table 7.1: Quality of the evaluated text extraction methods.

Extraction Method	Precision	Recall	F1-Score	F0,5-Score
NTA-Ontology	0.843710211	0.441778457	0.566124775	0.696458921
NTA-Dictionary	0.510026738	0.272666765	0.381111613	0.465575159
NTA	0.348203232	0.353617506	0.380015414	0.380159693
DBPedia Spotlight	0.585997469	0.473666792	0.518386091	0.555032681

by using a shared group knowledge model for the semantic group formation. Although the previous section outlined performance characteristics of the group formation and the corresponding algorithm, the results from the group awareness are of high interest for this work. The work presented is based on the joint publication and summarizes the empirical results from Erkens et al. (2019) which are relevant for the conclusion of this study.

7.3.1 Observation of Communication Sequences

To investigate the first research question regarding the structuring of communication, a line chart visualizing the sequencing of communication based on the aforementioned coding scheme has been created. Figure 7.10 shows the line chart. The x-axis represents the chronological order of the topics mentioned by the students. The y-axis shows the order of the topics as displayed in the group awareness tool. If the learners discuss the topics exactly as presented through the tool, the chart would correspond to the identity function (diagonal). To better illustrate deviations in the sequencing, the identity has been included into every chart as a dark gray line. Another differentiation has been made between an initial mentioning of a topic (light gray) and multiple mentionings (gray).

The results indicate, that the learning groups without CG AT tend to follow the given structure, for instance the given list of topics in the visualization, when organizing their knowledge exchange. This means that, although this is not a group awareness tool or a direct instruction, it implicitly prescribes the behavior of learners. In contrast, groups supported by the CGAT tend to deviate from this order. However, this does not apply to the CGAT group in class C, which shows a similar pattern as the groups without CGAT support. The content of the visualization illustrates that this is a special case, as the visualization indicates that the learners did not have any knowledge on the first topic ("Brownian molecular motion") and had a highly complementary knowledge on the second topic ("development of power demand"). Apparently, the learners have talked at first about the only topic with missing knowledge and then about the topic with the greatest differences before continuing with osmosis, but also in a chronological order.

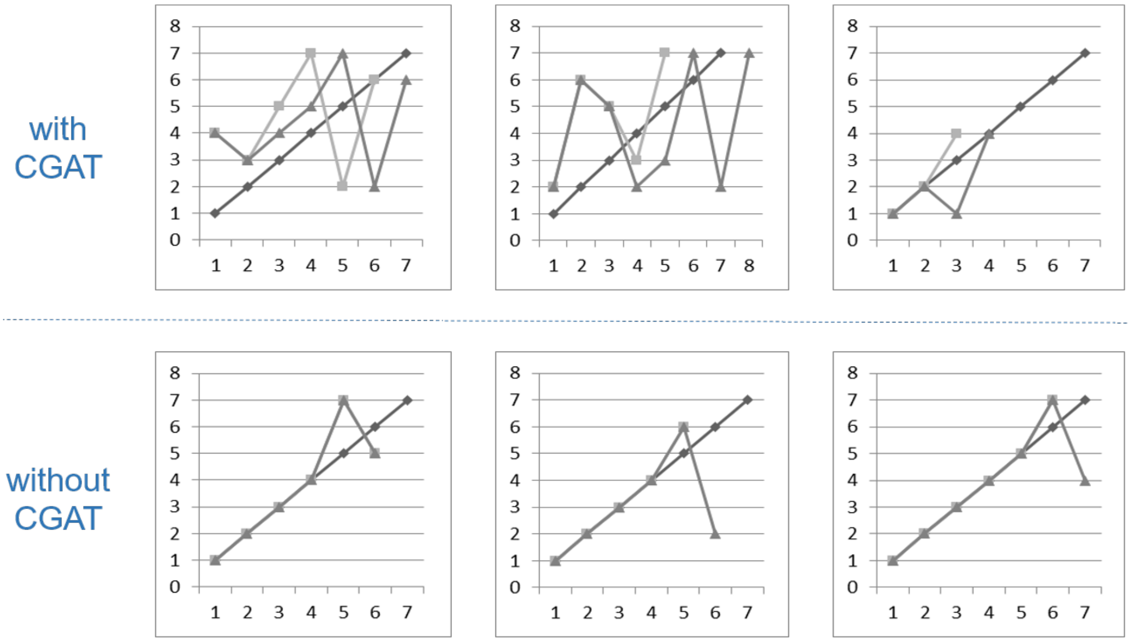


Figure 7.10: Communication sequences of groups with CGAT and without CGAT. X-axes represent the order, in which students talked about topics, y-axes represent the position of topics in the list. Classes are A, B, C from left to right.

To pursue the second question regarding the sequencing, the number of content-related and organizational questions and explanations have been counted and compared to the respective level of knowledge per topic. The level of knowledge can be inferred from the visual distance of the bars in the group awareness visualization. If the length of a bar is less than 50% of a co-learner's bar, then the student's knowledge is classified as *low level*. Vice versa, the co-learner would be classified as having a *high level* of knowledge. Although only the CGAT groups (with GA support) have a visual display of the knowledge, the data from the groups without group awareness have been calculated through the system and have been considered to support the evaluation as well.

Table 7.2 presents the resulting numbers of the questions and explanations. There are almost no differences between students of both experimental groups, since in both groups learners asked nearly the same number of the questions. This might have been biased through the clear instructions about the decision problem and the clear goal, which might have influenced the learners in articulating their own discourse.

By taking own levels of knowledge into account, differences can be found: Students in the CGAT groups asked content-related questions in 71% of the cases (10 out of 14 questions), in which the bar in the visualization was less than 50%. In the groups without CGAT it was just in 46% of the cases (6 out of 13 questions). Regarding the distribution of knowledge in the whole group, similar values in both experimental groups suggest that there are no differences concerning the impact of knowledge distribution on content-related questions. Furthermore, there has been even a counter-intuitive finding: in both conditions, the learners explained more in the cases, in which they had been classified with low knowledge. As the time span between the two experimental phases was quite big, the learners had potentially enough time to close the gaps in their own knowledge, if discovered during the first, individual phase. However, this was not in control of the experiment, which makes it impossible to draw any conclusions about this. In contrast, students in the 'No CGAT' groups only answered in 10% of the cases (1 explanation out of 10) questions, if their bar length was higher than the questioner's bar length and in 50% of the cases, if their bar was the same or higher than the learning partners' bars. This indicates that without an appropriate visualization of cognitive information, the learners tend to underestimate their knowledge. Therefore, such group awareness tools have the potential to support the communication in the group processes. In addition, the tool helps to put organizational aspects into the background: students without support seem to focus slightly more on organizational issues than students with GA support: In the CGAT group, 42% of the questions (10 out of 24 questions) and 12% of the explanations (8 out of 68 explanations) were about the organization of communication. In the 'No CGAT' groups, 55% of the questions (16 out of 29 explanations) and 25% of the explanations (18 out of 71 explanations) were about the organization of communication.

Table 7.2: The number of questions and explanations depending on the own knowledge level and the knowledge distribution.

	CGAT				No CGAT			
	class A (n=3)	class B (n=3)	class C (n=3)	total (n=9)	class A (n=3)	class B (n=2)	class C (n=3)	total (n=3)
# content-related questions	3	7	4	14	7	5	1	13
(a) bar length <50%	1	6	3	10	3	2	1	6
(a) bar length >50%	2	1	1	4	4	3	0	7
(b) smaller bar than the other(s)	1	4	0	5	1	2	1	4
(b) in the middle	1	0	3	4	3	0	0	3
(b) longer bar than the other(s)	1	3	1	5	3	3	0	6
# content-related explanations	22	12	26	60	12	8	33	53
(a) bar length <50%	7	10	19	36	7	5	24	36
(a) bar length >50%	15	2	7	24	5	3	9	17
(b) smaller bar than the other(s)	6	5	7	18	3	5	18	26
(b) in the middle	5	2	11	18	5	0	3	8
(b) longer bar than the other(s)	11	6	7	24	4	3	12	19
(c) smaller bar than the questioner	0	3	1	4	3	2	0	5
(c) same bar as the questioner	2	2	0	4	3	0	1	4
(c) longer bar than the questioner	1	6	2	9	0	1	0	1
# organizational questions	3	2	5	10	7	5	4	16
# organizational explanations	2	1	5	8	6	8	4	18

7.3.2 Impact of the Group Awareness on the Created Artifacts

It could be observed that displaying group awareness information had an impact on the communication and coordination of the groups, particularly the sequencing of topics. In addition to this, it seems plausible that the presentation of cognitive information might have had an impact on the learner-generated artifacts itself. Arguing that the group awareness tool provided a guideline for the communication about particular topics, this might have influenced the group work regarding the decision task. During the collaboration, the groups could take notes in a wiki tool. For the final assessment, they created a short essay-like text that contained an argumentation and a positioning regarding the decision. In total, there was not a big difference between the average word count in the wikis. The following table shows the average wiki word count:

Table 7.3: Average word count in total and without groups with missing members.

Average Wiki Word Count		
	CGAT	NOCGAT
All groups	60.86	58.37
No missing	61.65	33.66

In conjunction with the observation, there were two interesting details that explain the learning processes a bit better. In full awareness of the low sample size, it is valuable to see qualitative indications that help to understand the group processes guided through awareness tools. It could be observed that the learners negotiated about the topics and tried to find a common ground on their knowledge. This applied for the group without group awareness as well. To prevent any confounding effects, a list of topics has been displayed, in particular, the labeling of the domain axis in the group awareness tool (the category labels). This led to a similar mode of group work in the case without support tools. The main difference that could be observed was a different order of the topics. This possibly explains the low differences between the performances of the groups.

However, another bias has been introduced due to the absence of several persons. A few groups (triads) missed a single person that attended the first phase and had been assigned, but did not show up. These groups have been included into the evaluation, because the corrected diversity could be recalculated, but the groups were marked as missing in the table below 7.4. It was symptomatic that exactly these groups with missing persons wrote the longest wiki texts in the control group without the group awareness tool. A possible explanation for this behavior is that they had to compensate the missing person by producing more detailed explanations in the wiki tool. Working in dyads tends to be more effective than working in triads. Additionally, the absence of the group awareness tool might have saved coordination time and efforts, which might have improved their efficiency. To further investigate, whether this phenomenon has its roots in the knowledge exchange phase or much later, would be helpful to clarify this question. Unfortunately, this analysis is not possible with the data set available. The wiki tool only produces a new revision, when the learner explicitly clicks on the save-button. Nearly all texts have only one revision, which makes it impossible to make any detailed statement about the creation time.

Moreover, there were groups with missing students that had group awareness support. That means, the missing learners were still represented as bars in the visualization of the group awareness tool. The missing opportunity to clarify or ground their knowledge based on the (incomplete) information could have led to a worse performance in terms of wiki writing. Table 7.4 shows the full data set of the word count and the corresponding diversity. No correlation between diversity and word count could be observed. However, for future research about the mechanisms of group awareness, it might be interesting to see in more detail, how these communication processes are mediated through group awareness tools. To accomplish this, a fine grained observation and a detailed action logging of user events is necessary.

Table 7.4: Average word count per group.

Group	Participants	Missing	Diversity	AVG Word Count
ANoCGAT1	["A12","A2","A24"]		0.86	29.50
ANoCGAT2	["A14","A21","A9"]		0.83	47.33
ANoCGAT3	["A19","A20","A23"]	A20	0.80	82.33
ANoCGAT4	["A1","A6","A7"]		0.80	17.80
BNoCGAT1	["B15","B3","B9"]	B3	0.77	50.80
BNoCGAT2	["B14","B6","B7"]	B14	0.69	144.50
CNoCGAT1	["C4","C8","C9"]		0.68	31.00
CNoCGAT2	["C2","C6"]		0.81	42.67
ACGAT1	["A10","A13","A5"]		0.89	50.50
ACGAT2	["A17","A4","A8"]		0.79	52.00
ACGAT3	["A11","A16","A3"]	A11	0.76	89.67
ACGAT4	["A15","A18","A22"]		0.75	63.50
BCGAT1	["B11","B12","B8"]		0.80	132.00
BCGAT2	["B13","B16","B5"]	B5	0.76	16.00
BCGAT3	["B10","B4"]		0.80	68.33
BCGAT4	["B1","B2"]		0.76	34.50
CCGAT1	["C1","C10","C3"]		0.64	28.00
CCGAT2	["C5","C7"]		0.65	64.33

7.4 Discussion

The conducted study aimed to investigate the impact of providing groups, which have been formed to maximize knowledge complementarity, with group awareness information during a knowledge exchange. The information displayed to the learners was generated through semantic extraction as a part of the group formation algorithm. The presented results indicate that especially learners who do not have the group awareness support tend to ask more questions when they lack knowledge or when their partners have a higher knowledge. In addition, they also tend to give more explanations when they have a well-grounded expertise or when their co-learner lacks expertise. Findings of qualitative analyses of communication patterns and sequences in the knowledge exchange show on the one hand that the level of own knowledge impacts questioning behavior as learners in the learning groups with support asked more often questions in such cases in which their own missing knowledge was visualized. This indicates an improved cognitive regulation in CGAT groups. On the other hand, the results indicate that the knowledge distribution impacts explaining behavior as learners with group awareness support answered questions more often when they had more knowledge than the questioner. Thus, knowledge distributions were not more important in the CGAT groups than in the control group. However, the visualization of the knowledge appears to foster the decision to answer a question. These results are in accordance with former results that the level of own missing knowledge guides cognitive regulation and that knowledge distribution guides explaining behavior (Dehler et al. 2011).

Learners without available group awareness information mainly followed a chronological order of the topics displayed. Contrary, learners with group awareness support deviate from such a sequencing of topics in the knowledge exchange. However, the question remains which other factors influence the sequencing of co-learners in the groups with group awareness support. One possible explanation might be derived from the visualizations of the sequences: Learners with group awareness support talked more often multiple times about certain topics, especially on the topic "osmosis". since that was of fundamental importance to fulfill the first collaborative task of writing a short text on the potential meaning of salt and fresh water for osmotic power plants. This implies that learners with group awareness support had this task more in mind, related osmosis to other topics and thus better integrated contents of learning material in their exchange of knowledge. Further research should examine to what extent content-related information might be an influential factor, and what value it adds on learning when combined (or not) with group awareness information. Moreover, the differences between groups of two and three learners have not been investigated. This might be relevant regarding co-learners' behavior and should be considered in future studies.

Overall, such an approach of using semantic extraction seems suitable to collect, transform, and visualize cognitive information from educational data to support teachers in their challenging task to form knowledge-complementary groups and to visualize learners' cognitive information for better group awareness. The results have shown that the semantic group formation is a suitable algorithm to fulfill the goals that were set. Compared to the question about heterogeneity and homogeneity in certain skills, an approach based on knowledge diversity promises to be more inclusive. This can also be found in the analysis of the algorithm, particularly the relatively even distribution of diversity scores for different conditions, rendering the result of the semantic group formation as being fair. The concept of group awareness closes the gap of transparency and interpretability of algorithms. Many methods and structures in AI are criticized for the difficulty of understanding, interpreting, and explaining of their results. In contrast, the semantic group formation creates human-understandable knowledge models that even explain the algorithm. Following the concept of group awareness, cognitive information about knowledge distributions can be integrated into the learning scenario. The setting for this study shows an example of how to orchestrate learning tools in order to create a workflow from knowledge extraction over knowledge visualization to group awareness as a part of the learning design. Furthermore, the results of this work show that the use of such a workflow has the potential to improve learning. However, the psychological and social processes during the collaboration phase are very complex and some of the results advise to reenact the scenario with a more detailed level of analysis focusing on the metacognitive aspects.

8 Conclusion

Scaffolding and guiding learners has a long tradition for scientific inquiry and inquiry-based learning. Virtual learning environments and the growing availability of online science laboratories pose new challenges for supporting learners. On the one hand, Go-Lab provides a pedagogical middleware and technological interfaces to effectively promote and orchestrate inquiry-based science education. On the other hand, the magnitude of singleton solutions and decoupling of platforms demand strong technical interfaces that bridge the different technologies in order to support scaffolding and guidance on the architectural level. The rising field of learning analytics investigates in how to support learning and knowledge construction through computational methods.

8.1 Contribution of this Work

The first contribution of this thesis is the creation of an architecture for learning analytics in Go-Lab based on state of the art technologies and standards. The main goal of this architecture is to support a computational level of learning analytics through interfaces and APIs, but also to connect a conceptual level of scaffolding and guidance to the technical framework. The design of the learning analytics server as part of this work has been proven to enrich the Go-Lab infrastructure. It has been used to capture learning processes on the level of apps and ILS during the project time. After two years of using it, we evaluated how teachers and learners used it, based on the learning analytics data. The analysis of the learning processes has been used to give recommendations and to synthesize a recommended inquiry model that incorporates the pedagogical approaches and the end-user perspectives. We found out that there is a need to further support learners through guidance mechanisms and through fostering process awareness and cognitive group awareness.

The second contribution is to provide mechanisms that facilitate the given infrastructure in order to support learners. The concept cloud app is a cognitive scaffold which presents a shared group knowledge model and enables reflective thinking with respect to the knowledge state. We augment the technical layer of Go-Lab in order to incorporate classroom-collaboration support through automatic group formation. For scientific inquiry, knowledge plays an important role, for example when asking questions, posing problems, designing experiments and evaluating the observations. Advancing the technical challenges of the architectural approach to pedagogical challenges of the classroom management, we highlight the role of knowledge diversity as well. In nearly every science class, different people know different things. Without stigmatizing weak learners, nearly everyone can learn something from each other. Therefore, managing knowledge diversity has the potential to improve learning in terms of the learning outcome, but also by creating a more inclusive view about knowledge and performance. The semantic group formation does not create heterogeneous groups with high-achievers and low-performers, it creates groups of complementary knowledge. Such learning scenarios that incorporate this kind of groupings also live from the dynamics of the groupings. For this purpose, we wanted to incorporate the information of knowledge diversity naturally into the learning scenario. To achieve this, we created a cognitive group awareness tool which used the knowledge diversity from the group formation in order to display the complementarity of a group to the learners.

Third, three classroom studies have been carried out using the Go-Lab learning environment in order to evaluate the concepts and effects of skill-based and semantic group formation, cognitive group awareness and cognitive scaffolding.

8.2 Discussion

Although this work contributes to the interdisciplinary fields of CSCL, Learning Analytics and Technology-Enhance Learning in a way that it provides (technical) solutions to better manage and facilitate knowledge diversity in classrooms in order to create a more inclusive setting, this work has some limitations.

Algorithmic complexity We are aware that the algorithms proposed for automatic group formation have a high run-time complexity. The focus of this work is not in delivering a high-performance algorithm for group formation in this case. The nature of the experimental and pedagogical settings in a classroom does not demand a high scalability, as typical classroom sizes do not exceed 30. As a consequence of pragmatic issues regarding the group size, the semantic group formation allows a flexible grouping. The algorithm has a lower and an upper bound for the allowed group size. This increases the complexity of the algorithm to some extent, also regarding the asymptotic approximation. However, to be able to deal with a higher number of students, the algorithm takes several samples by creating random bi-partitions (default: 3) of the set and calculates the diversity for each set as usual. In the end, the sample with the best overall diversity score is taken.

Pragmatic issues for classroom experiments Additionally, pragmatic issues that typically occur in classroom settings introduce new challenges for the algorithm. To scale up and implement technology-enhanced learning at scale, facilitators and researchers need to incorporate the needs of the important stakeholders, foremost teachers and students (Rodríguez-Triana et al. 2015). A typical cliché that teachers draw for researchers in learning sciences is that they are not aware of their working practice, which is true to some extent. This can be circumvented by collecting requirements directly from the stakeholders (Rodríguez-Triana et al. 2015) and by facilitating a participatory design process, as it has been implemented within the Go-Lab project (Heintz et al. 2014). This shifts the focus of us researchers towards more practical and pragmatical aspects, for example, in the design of the semantic group formation app in the context of Go-Lab. Selected aspects are:

1. *Flexible group size.* A cohort cannot always be divided into dyads. It is naive to believe that. In fact, with an even number of students, on the day of the trial students might be absent due to several reasons such as illness. Therefore it is more important to provide a flexible group size that ranges between bounds rather than having a fixed size. Apart from this, ICT resources are limited (e.g., access to computers), so that the group size needs to be set higher. However, the group size influences the collaboration, and the flexibility of the algorithm also its computational complexity.

2. *Missing artifacts.* During our experiments we found out that students sometimes just did not deliver. Although they have been prompted, they forgot, intentionally skipped certain tasks, or even had technical difficulties, for example due to a bad internet connection. Therefore we enabled the teacher to take control of all the mechanisms of the algorithm. The triggering of the group formation calculation can be performed by the teacher, but also the addition of missing artifacts or quickly setting existing artifacts as references. Since the results of the algorithm are visual, the teacher does not lose the control over a black box. By re-running the algorithm through the web-interface, the teacher has the chance to control the results ex-post.

3. *No dependency of reference solutions.* Automatically checking the correctness of learner-generated content has always been of interest in many disciplines such as automated programming assessments. Typical approaches for checking learner-generated content rely on reference solutions. Teachers do not want to or do not have the time to develop and include reference solutions carefully. However, in a large-scale implementation such as Go-Lab, it is impossible to provide reference solutions by members of the consortium. Go-Lab provides through its inventory more than 1000 learning spaces[1] with individual tasks.

[1] Golabz provides 1008 available spaces (2019-06-03). `https://www.golabz.eu/spaces`, retrieved 2019-06-03.

> Most of them are created by teachers and not by members of the Go-Lab consortium. Therefore, we envisioned to create methods and scaffolds that deal without reference solutions in order to provide a more scalable approach. Still, it is possible to deal with reference solutions on demand, which was done in both group formation approaches. The semantic group formation app enabled the teacher to create an overlay model by adding his or her own artifacts as a reference solution.

These are only a few examples of pragmatic aspects that have driven the research particularly for this thesis, and also for the whole Go-Lab project in order to conduct a successful large-scale implementation in more than 1000 schools across Europe during the project time from 2012 to 2016 (Govaerts et al. 2013a).

Learning outcomes and research paradigms In some of the evaluations, we used learning gain to measure the outcome of the setting and to get a quality indicator to quantify learning. The evaluations presented as part of this thesis usually took 3 to 4 lessons in school. On the one hand, such classroom studies are closer to reality and the expectations of teaching and learning than common laboratory studies on learning. On the other hand, such a time span is not enough to quantify learning outcome. We are aware that learning takes more time and demands higher level competencies that cannot be emulated in such a short time (Anderson and Arsenault 2005). The question about how to evaluate learning and how to conduct research in learning sciences was a persistent debate across different disciplines and communities. Researchers focused rather on spotting differences of approaches than on finding a compliant or synergetic discourse (Johnson and Onwuegbuzie 2004). This debate also takes roots in the field of learning sciences, where quantitative research is usually done in a controlled laboratory setting. This suggests the legitimate criticism on how to transfer findings to a pedagogical context in a complex classroom setting with many social parameters and variables. Brown (1992) highlights the need for a paradigm change when creating complex interventions in classroom settings:

> "There was also a dramatic change in what 'subjects' were required to learn, even in laboratory settings, and an awakening to the fact that real-life learning inevitably takes place in a social context, one such setting being the classroom. Psychologists are creatures of their time, and the methods they use to attack such durable problems as learning must be reconsidered in the light of theory change."

Seidman (2006) advocates for more qualitative research, in particular methods based on interviews, which involve multiple stakeholders. These could be in our case both the teachers and the students. He argues that the quantitative paradigm is still predominant in educational research and sometimes referred to as being the scientific standard. Even in the field of CSCL and Learning Sciences, which supports qualitative research methods such as observations per se, Jeong and Hmelo-Silver (2010) found out that most of the research is quantitative. In the meta-study, less than 63% of the studies that investigated face-to-face collaboration have been evaluated following a qualitative approach. For other modalities it was even less.

To make real quantitative statements about learning processes, it needs bigger samples than in the studies presented. However, qualitative methods like detailed classroom observations used in the evaluation of the group awareness (see section 7) or the analysis of complete learning traces (cf. sections 3.4 and 6) have the potential to gain new insights through their exploratory nature. Combining the paradigms of qualitative and quantitative approaches leads to a "third wave" of research methodology using mixed methods (Johnson and Onwuegbuzie 2004).

Teacher-led inquiry The concept cloud has demonstrated its usefulness as a cognitive tool for learners in the context of IBL. It represents a shared group knowledge model of the learners. As such, it can be seen as a "summary" of a lesson. As discussed in the debate of teacher-led inquiry, teaching can also be the subject of an inquiry process. There are certain parameters like learning resources, activities, aspects of orchestration, instructions, and many more that lead to a certain learning outcome. Teachers might use this representation to find common misunderstandings, weaknesses in learning materials or a focus shift. However, in the research of TEL and IBL there has not been much work done on investigating teacher-led inquiry. Moreover, there is a lack of support tools for this particular endeavor.

In analogy, the same applies for group formation. The semantic group formation has, in contrast to other group formation algorithms, a clear output that is interpretable by humans. This shared knowledge model that visualizes knowledge overlap and complementarity can be used as a part of the instructional design, as presented in the evaluation of the semantic group formation and the cognitive group awareness (cf. section 7). However, this work is only the tip of the iceberg. Such cognitive group awareness information about knowledge can be useful in other contexts as well (workplace learning), and also to outline a cognitive or competency-based development over time. A fundamental understanding of the visualizations and results of such approaches requires a basic understanding of algorithms and the associated mechanisms. Particularly the AI-enhanced methods demand for an interpretability to gain trust and to be adopted in a computer-supported learning (and teaching) environment. This is in line with the discussion about explainable AI, which advocates the development of algorithms (in the context of AI and machine learning) that are human-interpretable (Abdul et al. 2018).

8.3 Publications

Figure 8.1 shows the conceptual model of this work, presented as a layered architecture. Each layer of this framework is connected to a specific set of publications used for this thesis. The upper layer (architecture) was dedicated to the definition of a learning analytics architecture and the interfaces to the Go-Lab ecosystem as well as the embedding of learning analytics applications. It contains all publications which are connected to the architectural approach: (Govaerts et al. 2013a; Hecking et al. 2014; Manske et al. 2014; Vozniuk et al. 2014).

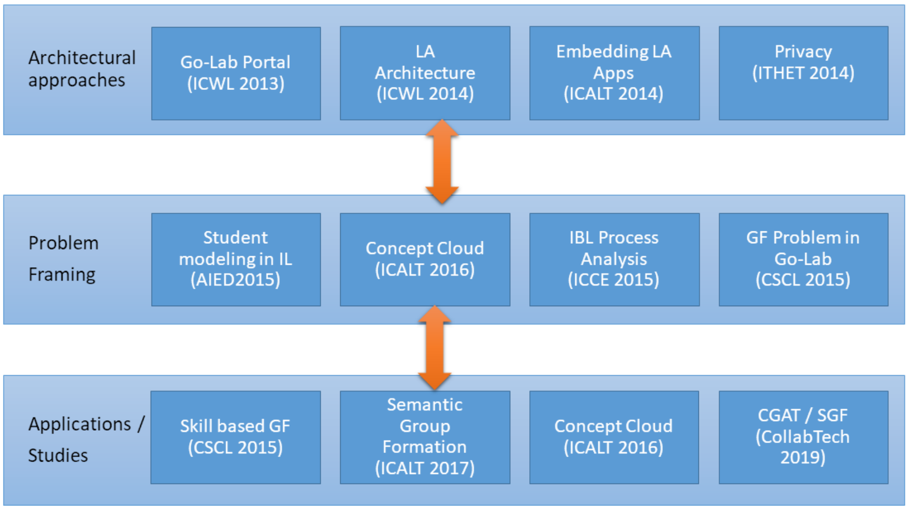


Figure 8.1: The three layer model of the contributions and the connection of the publications for this thesis.

The second layer of the framework includes the publications that encompass the problem framing (Manske et al. 2015b, a, c; Manske and Hoppe 2016). The inquiry model as a basic pedagogical layer has been analyzed in the first learning analytics evaluation of the system. Knowledge is one of the key components in inquiry, and the modeling of knowledge can be seen as one of the distinctive features of this work, including applications of cognitive and metacognitive scaffolding, performance prediction and group formation. The problem framing has been mapped to a conceptual model in figure 8.2.

The application layer shows use cases and their respective evaluation in classroom studies. The evaluation of skill-based and semantic (knowledge-based) group formation shows effective methods for extracting skills and knowledge from learner-generated artifacts in heterogeneous inquiry learning scenarios. The evaluation of the concept cloud as a cognitive tool using shared group knowledge has shown that it successfully scaffolds learning, particularly concept mapping. Such knowledge models can be visualized and represented as cognitive group awareness tools in order to feed back the knowledge diversity actively into the learning scenarios. This framework of the thesis incorporates mechanisms to effectively manage knowledge diversity in inquiry-based science education.

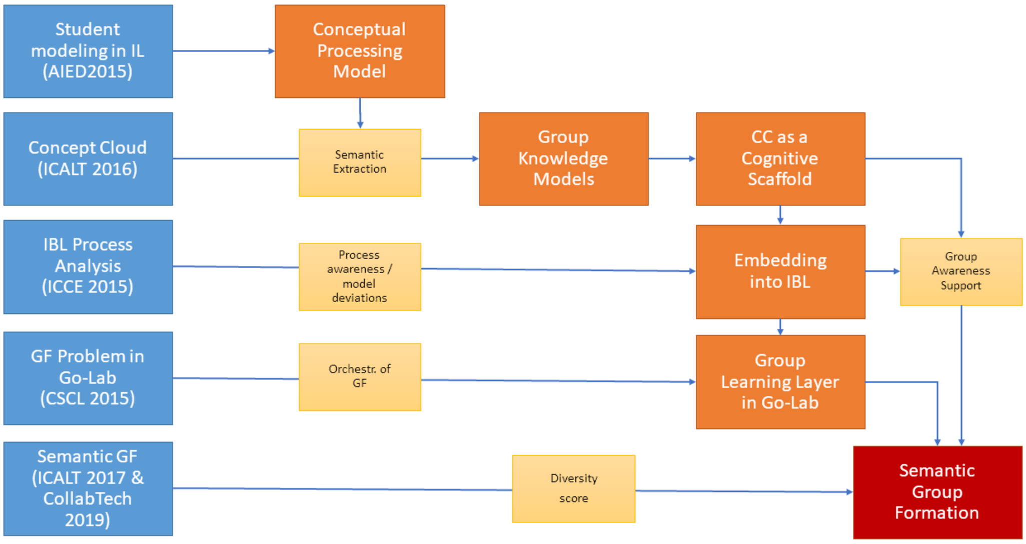


Figure 8.2: The conceptual elements and problems the publications targeted.

8.4 Outlook

The Go-Lab architecture offered a variety of interfaces to connect a learning analytics infrastructure to support the mechanisms to manage knowledge diversity. However, many modern systems used xAPI as an interfacing standard to transfer activity statements to a learning record store. In order to support external applications for further analytics, such mechanisms can be taken into account.

The management of knowledge has shown a lot of benefits, particularly when knowledge models are created and presented as open learner models. The idea of cognitive group awareness has shown an interesting use case of how to interweave computer-supported mechanisms with classroom collaboration and demonstrated how collaboration can be initialized. The notion of knowledge underlay the simple assumption, that concepts, the learners are talking about, are used correctly. Current discussions in the field semantic analyses try to build more intelligent and advanced models. The use of ontologies can improve this, by inferring knowledge from a given ontology. For example, in the context of IBL, Wikipedia represents a lot of knowledge. Rather than just spotting keywords on Wikipedia-based ontologies such as DBpedia, intelligent mechanism can draw conclusions about how certain aspects are presented by learners. Particularly in the field of hypothesis generation, the concept of "qualitative reasoning" could be included to check the plausibility of the input.

Additionally, the dictionary that was used for the CGAT study was created manually. Using knowledge sources such as Wikipedia, which categorize topics in a (pseudo-)hierarchy, can be used to extract topics and their respective categories flexibly. As a consequence, this approach scales for all IBL scenarios without having a prescribed ontology. Extraction methods such as Explicit Semantic Analysis in conjunction with a Wikipedia category graph can be facilitated to support this idea.

As presented in this approach, managing knowledge diversity incorporated learners knowledge, and excluded other variables of diversity. A lot of the research has been investigated in the different dimensions in organizational or economic research. Incorporating other aspects similar to more traditional work in research on diversity might unveil even more potential of diversity management in classrooms. Managing diversity hopefully leads to more inclusive classrooms and a less stigmatizing pedagogy. Apart from schools, there are other educational contexts that start to let digitization thrive. Particularly companies that operate in the STEM fields have an interest that their apprentices acquire scientific skills. Bringing such approaches into their apprenticeship programs will be challenging but promising for companies. This might become a new take on learning and teaching with the support digitization.

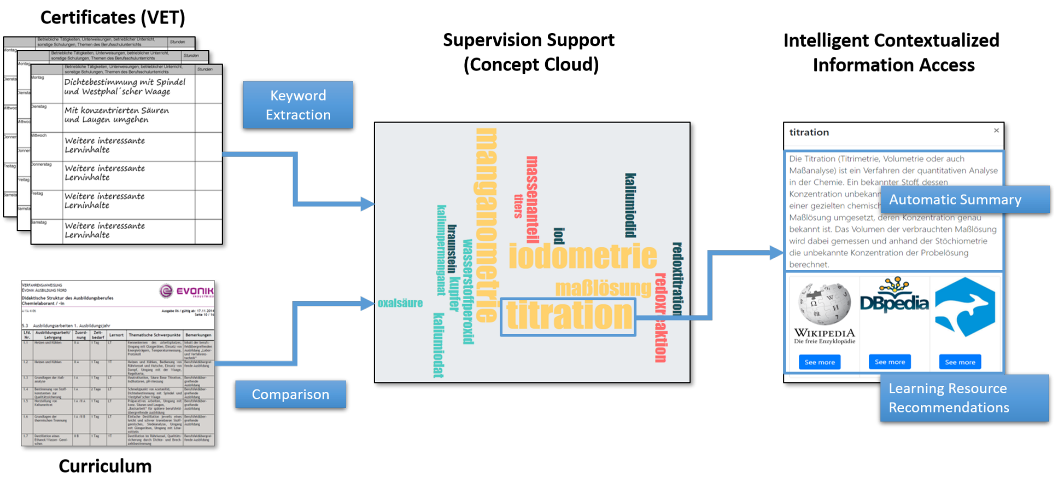


Figure 8.3: The concept cloud app adapted to the VET context, enriched by mechanisms of contextualized intelligent information access.

Additionally, digitization and digital transformation offer many new opportunities for vocational education and training (VET). Although the integration of digital hardware and software tools in educational contexts introduces new challenges for learners and facilitators, the provision of well-adapted technologies has the potential to support learning and teaching using intelligent technologies. In this spirit, the knowledge management approaches that have been conceptualized and presented in this work, can be transferred to other contexts and target domain. As part of an ongoing work, the concept cloud application that has been presented in chapter 4.3 is currently transferred into the context of chemical industry. It has been advanced regarding the color scheme to support certain types of color-blindness, but furthermore, it bridges learners content to semantic technologies providing intelligent contextualized information access. Figure 8.3 shows a prototype of the concept cloud within a VET scenario. The concepts are extracted from formal transcripts of records that are used for the certification of apprentices and filled out by them. The results are displayed in a similar visualization that highlights the differences and comparison to the curriculum. In addition, for each concept the interactive visualization provides information access by displaying a summary that has been automatically extracted from DBpedia and links to external learning resources. Using such information accesses, knowledge management technologies like this can be used to further discover and automatically link open educational resources and thus provide easy and free access to knowledge sources.

Bibliography

Abdul, A., Vermeulen, J., Wang, D., Lim, B. Y., and Kankanhalli, M. (2018). Trends and trajectories for explainable, accountable and intelligible systems: An hci research agenda. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, page 582. ACM.

Abramowitz, M. and Stegun, I. A. (1965). *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*, volume 55. Courier Corporation.

Adadi, A. and Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (xai). *IEEE Access*, 6:52138–52160.

Aleven, V., Koedinger, K. R., and Cross, K. (1999). Tutoring answer explanation fosters learning with understanding understanding. In *Proceedings of the 9th International Conference on Artificial Intelligence in Education*, pages 199–206.

Alfieri, L., Brooks, P. J., Aldrich, N. J., and Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of educational psychology*, 103(1):1.

Anderson, G. and Arsenault, N. (2005). *Fundamentals of educational research.* Routledge.

Anderson, J. R. (1984). Cognitive psychology and intelligent tutoring. In *Proceedings of the Cognitive Science Society Conference*, pages 37–43.

Anderson, J. R. (2013). *Language, memory, and thought.* Psychology Press.

Andresen, M. (2007). Diversity learning, knowledge diversity and inclusion: Theory and practice as exemplified by corporate universities. *Equal Opportunities International*, 26(8):743–760.

Andrus, D. C. (2005). Toward a Complex Adaptive Intelligence Community. page 30.

Angelou, M. (1994). *Wouldn't take nothing for my journey now.* Bantam Books.

Angenendt, K., Bormann, J., Donkers, T., Goebel, T., Kizina, A., Kleemann, T., Michael, L., Raja, H., Sachs, F., Schneegass, C., and others (2015). ConceptCloud - Entwicklung einer Applikation zur Unterstützung von Reflexionsprozessen im Online-Lernportal Go-Lab. In *DeLFI Workshop Proceedings*, pages 132–135.

Aronson, E. (2002). Building empathy, compassion, and achievement in the jigsaw classroom. *Improving academic achievement: Impact of psychological factors on education*, pages 209–225.

Aronson, E., Blaney, N., Stephan, C., Sikes, J., and Snapp, M. (1978). The jigsaw classroom. *Beverly Hills: Sage.*

Auer, S., Bizer, C., Kobilarov, G., Lehmann, J., Cyganiak, R., and Ives, Z. (2007). Dbpedia: A nucleus for a web of open data. In *The semantic web*, pages 722–735. Springer.

Augar, N., Raitman, R., and Zhou, W. (2004). Teaching and learning online with wikis. page 10.

Avison, D. and Shah, H. (1997). *The Information Systems Development Life Cycle: A First Course in Information Systems.* McGraw-Hill Publishing Co., inc. McGraw-Hill.

Azevedo, R. (2005). Computer environments as metacognitive tools for enhancing learning. *Educational Psychologist*, 40(4):193–197.

Baeza-Yates, R. (2004). A fast set intersection algorithm for sorted sequences. In *Annual Symposium on Combinatorial Pattern Matching*, pages 400–408. Springer.

Bakharia, A., Kitto, K., Pardo, A., Gašević, D., and Dawson, S. (2016). Recipe for success: lessons learnt from using xapi within the connected learning analytics toolkit. In *Proceedings of the sixth international conference on learning analytics & knowledge*, pages 378–382.

Balmaceda, J. M., Schiaffino, S. N., and Pace, J. A. D. (2014). Using constraint satisfaction to aid group formation in CSCL. *Inteligencia Artificial, Revista Iberoamericana de Inteligencia Artificial*, 17(53):35–45.

Banchi, H. and Bell, R. (2008). The many levels of inquiry. *Science and children*, 46(2):26.

Barkley, E. F., Cross, K. P., and Major, C. H. (2014). *Collaborative learning techniques: A handbook for college faculty.* John Wiley & Sons.

Barnett, R. M., Johansson, K. E., Kourkoumelis, C., Long, L., Pequenao, J., Reimers, C., and Watkins, P. (2012). Learning with the ATLAS experiment at CERN. *Physics Education*, 47(1):28.

Barrow, L. H. (2006). A Brief History of Inquiry: From Dewey to Standards. *Journal of Science Teacher Education*, 17(3):265–278.

Bell, P. and Linn, M. C. (2000). Scientific arguments as learning artifacts: Designing for learning from the web with kie. *International journal of science education*, 22(8):797–817.

Ben-Ari, M. (1998). Constructivism in computer science education. In *Acm sigcse bulletin*, volume 30, pages 257–261. ACM.

Bergenthum, R., Desel, J., Harrer, A., and Mauser, S. (2012). Modeling and mining of learnflows. In *Transactions on Petri Nets and Other Models of Concurrency V*, pages 22–50. Springer.

Blosser, P. E. (1980). A Critical Review of the Role of the Laboratory in Science Teaching.

Bodemer, D. (2011). Tacit guidance for collaborative multimedia learning. *Computers in Human Behavior*, 27(3):1079–1086.

Bodemer, D. and Dehler, J. (2011). Group awareness in CSCL environments. *Computers in Human Behavior*, 27(3):1043–1045.

Bodemer, D., Ploetzner, R., Bruchmüller, K., and Häcker, S. (2005). Supporting learning with interactive multimedia through active integration of representations. *Instructional Science*, 33(1):73–95.

Bodemer, D. and Scholvien, A. (2014). Providing knowledge-related partner information in collaborative multimedia learning: Isolating the core of cognitive group awareness tools. In *Proceedings of the 22nd International Conference on Computers in Education ICCE 2014*, pages 171–179.

Bogdanov, E., El Helou, S., Gillet, D., Salzmann, C., and Sire, S. (2010). Graaasp: a web 2.0 research platform for contextual recommendation with aggregated data. In *Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems*, pages 3523–3528. ACM.

Bogdanov, E., Limpens, F., Li, N., El Helou, S., Salzmann, C., and Gillet, D. (2012). A social media platform in higher education. In *Global Engineering Education Conference (EDUCON), 2012 IEEE*, pages 1–8.

Bohl, O., Scheuhase, J., Sengler, R., and Winand, U. (2002). The sharable content object reference model (SCORM) - a critical review. In *Computers in Education, 2002. Proceedings. International Conference on*, pages 950–951 vol.2.

Boulos, M. N. K., Maramba, I., and Wheeler, S. (2006). Wikis, blogs and podcasts: a new generation of Web-based tools for virtual collaborative clinical practice and education. *BMC Medical Education*, 6(1):41.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The journal of the learning sciences*, 2(2):141–178.

Bruner, J. (1984). Vygotsky's zone of proximal development: The hidden agenda. *New Directions for Child and Adolescent Development*, 1984(23):93–97.

Bruner, J. (1987). Actual Minds, Possible Worlds — Jerome Bruner | Harvard University Press.

Bruner, J. S. (1961). The act of discovery. *Harvard Educational Review*, 31:21–32.

Bull, S., Jackson, T. J., and Lancaster, M. J. (2010). Students' interest in their misconceptions in first-year electrical circuits and mathematics courses. *International Journal of Electrical Engineering Education*, 47(3):307–318.

Bull, S. and Kay, J. (2010). Open learner models. In *Advances in intelligent tutoring systems*, pages 301–322. Springer.

Bull, S. and Kay, J. (2013). Open learner models as drivers for metacognitive processes. In *International handbook of metacognition and learning technologies*, pages 349–365. Springer.

Caplan, A. J., Ellis, C. J., and Silva, E. C. (1999). Winners and losers in a world with global warming: noncooperation, altruism, and social welfare. *Journal of Environmental Economics and Management*, 37(3):256–271.

Carley, K. M. (1997). Network text analysis: The network position of concepts. *Text analysis for the social sciences: Methods for drawing statistical inferences from texts and transcripts*, 4:79–100.

Cañas, A. J., Hill, G., Carff, R., Suri, N., Lott, J., Gómez, G., Eskridge, T. C., Arroyo, M., and Carvajal, R. (2004). CmapTools: A knowledge modeling and sharing environment. page 9.

Ceriani, L. and Verme, P. (2015). Individual diversity and the gini decomposition. *Social Indicators Research*, 121(3):637–646.

Chan, L. M. and Zeng, M. L. (2006). Metadata Interoperability and Standardization - A Study of Methodology, Part I: Achieving Interoperability at the Schema Level. *D-Lib Magazine*, 12(6).

Chang, K.-E., Chen, Y.-L., Lin, H.-Y., and Sung, Y.-T. (2008). Effects of learning support in simulation-based physics learning. *Computers & Education*, 51(4):1486–1498.

Chatti, M. A., Dyckhoff, A. L., Schroeder, U., and Thüs, H. (2012). A reference model for learning analytics. *International Journal of Technology Enhanced Learning*, 4(5-6):318–331.

Chatti, M. A., Muslim, A., and Schroeder, U. (2017). Toward an open learning analytics ecosystem. In *Big data and learning analytics in higher education*, pages 195–219. Springer.

Chounta, I.-A., Giemza, A., and Hoppe, H. U. (2014). Multilevel Analysis of Collaborative Activities Based on a Mobile Learning Scenario for Real Classrooms. In *Collaboration Technologies and Social Computing*, pages 127–142. Springer.

Christian Maier, M. N. (2010). Lab2go – A Repository to Locate Online Laboratories. *International Journal of Online Engineering (iJOE)*, Vol 6(No. 1).

Christodoulopoulos, C. E. and Papanikolaou, K. A. (2007). A group formation tool in an e-learning context. In *Tools with Artificial Intelligence, 2007. ICTAI 2007. 19th IEEE International Conference on*, volume 2, pages 117–123. IEEE.

Chu, S. K. W., Reynolds, R. B., Tavares, N. J., Notari, M., and Lee, C. W. Y. (2017). Twenty-First Century Skills Education in Switzerland: An Example of Project-Based Learning Using Wiki in Science Education. In *21st Century Skills Development Through Inquiry-Based Learning*, pages 61–78. Springer Singapore, Singapore.

Clariana, R., Engelmann, T., and Yu, W. (2013). Using centrality of concept maps as a measure of problem space states in computer-supported collaborative problem solving. *Educational Technology Research and Development*, 61(3):423–442.

Clark, W., Luckin, R., and Jewitt, C. (2011). Deliverable D5. 1 Methods and Specifications for TISL Components V1. *NEXT-TELL Consortium, EU*.

Clements, P., Kazman, R., and Klein, M. (2002). *Evaluating software architectures: methods and case studies*. SEI series in software engineering. Addison-Wesley.

Clow, D. (2012). The Learning Analytics Cycle: Closing the Loop Effectively. In *Proceedings of the 2Nd International Conference on Learning Analytics and Knowledge*, LAK '12, pages 134–138, Vancouver, British Columbia, Canada. ACM.

Commentz-Walter, B. (1979). A string matching algorithm fast on the average. In *International Colloquium on Automata, Languages, and Programming*, pages 118–132. Springer.

Conlon, T. (2004). 'Please Argue, I Could Be Wrong': a Reasonable Fallible Analyser for Student Concept Maps. In *World Conference on Educational Multimedia, Hypermedia and Telecommunications*, volume 2004, pages 1299–1306.

Corcino, C., Hsu, L., and Tan, E. (1999). Asymptotic approximations of r-stirling numbers. *Approximation Theory and Its Applications*, 15(3):13–25.

Council, N. R. (1996). *National science education standards*. National Academies Press.

Cox, T. (1994). *Cultural Diversity in Organizations: Theory, Research and Practice.* Berrett-Koehler Publishers. Google-Books-ID: Dr7IllPVGJYC.

Crawford, B. A. (2000). Embracing the essence of inquiry: New roles for science teachers. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, 37(9):916–937.

Dahrendorf, D., Dikke, D., and Faltin, N. (2012). Sharing Personal Learning Environments for Widget Based Systems using a Widget Marketplace. In *Proceedings of the PLE Conference 2012 in Aveiro, Portugal, July 11-13, 2012. Aveiro, Portugal.*

David Lowe, Steve Murray, L. W. M. d. l. V. (2009). LabShare: Towards a National Approach to Laboratory Sharing. In *20 th Australasian Association for Engineering Education Conference.*

De Groot, R., Drachman, R., Hever, R., Schwarz, B. B., Hoppe, U., Harrer, A., De Laat, M., Wegerif, R., McLaren, B. M., and Baurens, B. (2007). Computer supported moderation of e-discussions: the argunaut approach. In *Proceedings of the 8th iternational conference on Computer supported collaborative learning*, pages 168–170. International Society of the Learning Sciences.

de Jong, T. (2006). Scaffolds for scientific discovery learning. *Handling complexity in learning environments: Theory and research*, pages 107–128.

de Jong, T. and Lazonder, A. W. (2014a). 15 The Guided Discovery Learning Principle in Multimedia Learning. *The Cambridge handbook of multimedia learning*, page 371.

de Jong, T. and Lazonder, A. W. (2014b). The Guided Discovery Learning Principle in Multimedia Learning. In Mayer, R. E., editor, *The Cambridge Handbook of Multimedia Learning*, pages 371–390. Cambridge University Press, second edition.

de Jong, T., Linn, M. C., and Zacharia, Z. C. (2013). Physical and Virtual Laboratories in Science and Engineering Education. *Science*, 340(6130):305–308.

de Jong, T. and Njoo, M. (1992). Learning and instruction with computer simulations: Learning processes involved. In *Computer-based learning environments and problem solving*, pages 411–427. Springer.

de Jong, T., Sotiriou, S., and Gillet, D. (2014). Innovations in STEM education: the Go-Lab federation of online labs. *Smart Learning Environments*, 1(1):1–16.

de Jong, T. and van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of educational research*, 68(2):179–201.

de Jong, T., van Joolingen, W. R., Giemza, A., Girault, I., Hoppe, H. U., Kindermann, J., Kluge, A., Lazonder, A. W., Vold, V., Weinberger, A., and others (2010). Learning by creating and exchanging objects: The SCY experience. *British Journal of Educational Technology*, 41(6):909–921.

Dehler, J., Bodemer, D., Buder, J., and Hesse, F. W. (2011). Guiding knowledge communication in CSCL via group knowledge awareness. *Computers in Human Behavior*, 27(3):1068–1078.

Dehler Zufferey, J., Bodemer, D., Buder, J., and Hesse, F. W. (2010). Partner knowledge awareness in knowledge communication: Learning by adapting to the partner. *The Journal of Experimental Education*, 79(1):102–125.

Deslauriers, L. and Wieman, C. (2011). Learning and retention of quantum concepts with different teaching methods. *Physical review special topics-physics education research*, 7(1):010101.

Dewey, J. (1910). Science as Subject-Matter and as Method. *Science*, 31(787):121–127.

Dewey, J. (1959). *The child and the curriculum.* Number 5. University of Chicago Press Chicago.

Dillenbourg, P. (1999). *What do you mean by collaborative learning?* Oxford: Elsevier.

Dillenbourg, P. (2002). *Over-scripting CSCL: The risks of blending collaborative learning with instructional design.* Heerlen, Open Universiteit Nederland.

Dillenbourg, P. (2013). Design for Classroom Orchestration. *Comput. Educ.*, 69:485–492.

Dillenbourg, P. (2015). *Orchestration graphs.* Number BOOK. EPFL press.

Dillenbourg, P. and Betrancourt, M. (2006). Collaboration load. *Handling complexity in learning environments: theory and research*, pages 141–165.

Dillenbourg, P., Järvelä, S., and Fischer, F. (2009a). The evolution of research on computer-supported collaborative learning. In *Technology-enhanced learning*, pages 3–19. Springer.

Dillenbourg, P. and Jermann, P. (2007). Designing integrative scripts. In *Scripting computer-supported collaborative learning*, pages 275–301. Springer.

Dillenbourg, P., Järvelä, S., and Fischer, F. (2009b). The evolution of research on computer-supported collaborative learning. In *Technology-enhanced learning*, pages 3–19. Springer.

Dimitriadis, Y., Prieto, L. P., and Asensio-Pérez, J. I. (2013). The role of design and enactment patterns in orchestration: Helping to integrate technology in blended classroom ecosystems. *Computers & Education*, 69:496–499.

Dimitriadis, Y. A. (2012). Supporting teachers in orchestrating cscl classrooms. In *Research on e-Learning and ICT in Education*, pages 71–82. Springer.

Dixon-Krauss, L. (1996). *Vygotsky in the Classroom: Mediated Literacy Instruction and Assessment.* ERIC.

Dobber, M., Zwart, R., Tanis, M., and van Oers, B. (2017). Literature review: The role of the teacher in inquiry-based education. *Educational Research Review*, 22:194–214.

Doshi-Velez, F. and Kim, B. (2017). A roadmap for a rigorous science of interpretability. *stat*, 1050:28.

Drachsler, H. and Greller, W. (2012). The Pulse of Learning Analytics Understandings and Expectations from the Stakeholders. In *Proceedings of the 2Nd International Conference on Learning Analytics and Knowledge*, LAK '12, pages 120–129, Vancouver, British Columbia, Canada. ACM.

Drucker, P. (1992). *The age of Discontinuity: guidelines to our changing society.* Transaction books. Transaction Pubs.

Duffy, P. (2006). The use of blogs, wikis and RSS in education: A conversation of possibilities. page 9.

Dunbar, K. (1993). Concept discovery in a scientific domain. *Cognitive Science*, 17(3):397–434.

Duval, E., Klerkx, J., Verbert, K., and others (2012). Learning dashboards & learnscapes. *EIST*, pages 1–5.

Dwyer, T. (2009). Scalable, versatile and simple constrained graph layout. In *Proceedings of the 11th Eurographics / IEEE - VGTC conference on Visualization*, EuroVis'09, pages 991–1006, Berlin, Germany. Eurographics Association.

Dwyer, T., Henry Riche, N., Marriott, K., and Mears, C. (2013). Edge compression techniques for visualization of dense directed graphs. *Visualization and Computer Graphics, IEEE Transactions on*, 19(12):2596–2605.

Dwyer, T., Koren, Y., and Marriott, K. (2009). Constrained graph layout by stress majorization and gradient projection. *Discrete Mathematics*, 309(7):1895–1908.

Dwyer, T. and Robertson, G. (2010). Layout with circular and other non-linear constraints using procrustes projection. In *Graph Drawing*, pages 393–404. Springer Berlin Heidelberg.

Eaton, D. M., Redmond, A., and Bax, N. (2011). Training healthcare professionals for the future: Internationalism and effective inclusion of global health training. *Medical teacher*, 33(7):562–569.

Eilks, I., Rauch, F., Ralle, B., and Hofstein, A. (2013). How to Allocate the Chemistry Curriculum Between Science and Society. In Eilks, I. and Hofstein, A., editors, *Teaching Chemistry – A Studybook: A Practical Guide and Textbook for Student Teachers, Teacher Trainees and Teachers*, pages 1–36. SensePublishers, Rotterdam.

Emin-Martínez, V., Hansen, C., Triana, R., and others (2014). Towards teacher-led design inquiry of learning. *eLearning Papers*, (36).

Erkens, M. and Bodemer, D. (2017). Which Visualization Guides Learners Best? Impact of Available Partner-and Content-Related Information on Collaborative Learning. Philadelphia, PA: International Society of the Learning Sciences.

Erkens, M. and Bodemer, D. (2018). Improving collaborative learning: Guiding knowledge exchange through the provision of information about learning partners and learning contents. *Computers & Education*.

Erkens, M., Bodemer, D., and Hoppe, H. U. (2016a). Improving collaborative learning in the classroom: Text mining based grouping and representing. *International Journal of Computer-Supported Collaborative Learning*, 11(4):387–415.

Erkens, M., Manske, S., Hoppe, H. U., and Bodemer, D. (2019). A flexible and extendable learning analytics infrastructure. In *International Conference on Collaboration Technologies and Social Computing. To appear.* Springer.

Erkens, M., Schlottbom, P., and Bodemer, D. (2016b). Qualitative and quantitative information in cognitive group awareness tools: Impact on collaborative learning. Singapore: International Society of the Learning Sciences.

Erkens, M., Schneitzer, M. K., and Bodemer, D. (2018). The Difference that Counts: Guiding Knowledge Exchange by Visualizing Levels of Co-Learners' Knowledge. International Society of the Learning Sciences, Inc.[ISLS].

Evans, M. A. and Wilkins, J. J. L. (2011). Social interactions and instructional artifacts: Emergent socio-technical affordances and constraints for children's geometric thinking. *Journal of Educational Computing Research*, 44(2):141–171.

Everitt, B. S. and Hothorn, T. (2009). *A handbook of statistical analyses using R.* Chapman and Hall/CRC.

Falk, D. R. and Johnson, D. W. (1977). The effects of perspective-taking and egocentrism on problem solving in heterogeneous and homogeneous groups. *The Journal of Social Psychology*, 102(1):63–72.

Faltin, N., Dahrendorf, D., and Ternis, J. T. (2013). Responsive Open Learning Environments. *ROLE Report.*

Fardon, R. (2003). *Counterworks: managing the diversity of knowledge.* Routledge.

Ferguson, R. (2012). Learning Analytics: Drivers, Developments and Challenges. *Int. J. Technol. Enhanc. Learn.*, 4(5/6):304–317.

Ferguson, R. and Shum, S. B. (2012). Social learning analytics: five approaches. In *2nd International Conference on Learning Analytics & Knowledge*, pages 23–33.

Fischer, F., Kollar, I., Haake, J. M., and Mandl, H. (2006). Perspectives on collaboration scripts.

Fortenbacher, A., Beuster, L., Elkina, M., Kappe, L., Merceron, A., Pursian, A., Schwarzrock, S., and Wenzlaff, B. (2013). LeMo: A learning analytics application focussing on user path analysis and interactive visualization. In *Intelligent data acquisition and advanced computing systems (idaacs), 2013 ieee 7th international conference on*, volume 2, pages 748–753. IEEE.

Forum, G. S. (2006). Evolution of Student Interest in Science and Technology Studies. Policy Report, OECD.

Frakes, W. B. and Baeza-Yates, R. (1992). *Information retrieval: Data structures & algorithms*, volume 331. prentice Hall Englewood Cliffs, NJ.

Fransen, J., Kirschner, P. A., and Erkens, G. (2011). Mediating team effectiveness in the context of collaborative learning: The importance of team and task awareness. *Computers in Human Behavior*, 27(3):1103–1113.

Fruchterman, T. M. J. and Reingold, E. M. (1991). Graph drawing by force-directed placement. *Softw. Pract. Exper.*, 21(11):1129–1164.

Gabriel, I. (2017). Effective altruism and its critics. *Journal of Applied Philosophy*, 34(4):457–473.

Gago, J., Ziman, J., Caro, P., Constantinou, C., and Davies, G. (2004). Europe needs more scientists. In *Brussels: European Community Conference Increasing Human Resources for Science and Technology.*

Garcia-Zubia, J., Lopez-de Ipina, D., Orduna, P., and Hernandez-Jayo, U. (2006). Experience with WebLab-Deusto. In *Industrial Electronics, 2006 IEEE International Symposium on*, volume 4, pages 3190–3195.

Garcia-Zubia, J., Orduna, P., Lopez-de Ipina, D., and Alves, G. R. (2009). Addressing Software Impact in the Design of Remote Laboratories. *IEEE Transactions on Industrial Electronics*, 56(12):4757–4767.

Garrison, D. R. (2003). Cognitive presence for effective asynchronous online learning: The role of reflective inquiry, self-direction and metacognition. *Elements of quality online education: Practice and direction*, 4:47–58.

Garrison, D. R., Anderson, T., and Archer, W. (1999). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The internet and higher education*, 2(2-3):87–105.

Gelernter, D. (1985). Generative Communication in Linda. *ACM Trans. Program. Lang. Syst.*, 7(1):80–112.

Geraedts, C., Girault, I., Hovardas, T., Julien, R., Kluge, A., Lazonder, A., Lechner, J., Marzin, P., Matteman, Y., Pedaste, M., Piksööt, J., Rinket, M., Sarapuu, T., Zanden, M. v. d., Verkade, A., Vold, V., Wajeman, C., and Zacharia, Z. C. (2010). SCY-mission 1. Public Deliverable DVIII.I, SCY consortium.

Geraedts, C., ST, J. H., UCY, T. H., Julien, R., UT, J. S., van der Zanden, M., UTE, M. Z., UiB, V. V., UiB, C. H., and UT, A. L. (2011). SCY-mission 2. Public Deliverable DVIII. II, consortium.

Giemza, A., Manske, S., and Hoppe, H. U. (2013). Supporting the formation of informal learning groups in a heterogeneous information environment. In *proceedings of the 21st International Conference on Computers in Education*, pages 367–375. Asia-Pacific Society for Computers in Education Bali, Indonesia.

Giemza, A., Weinbrenner, S., Engler, J., and Hoppe, H. U. (2007). Tuple spaces as a flexible integration platform for distributed learning environments. *Frontiers in Artificial Intelligence and Applications*, 162:313.

Gijlers, H. and de Jong, T. (2005). The relation between prior knowledge and students' collaborative discovery learning processes. *Journal of research in science teaching*, 42(3):264–282.

Gillet, D., de Jong, T., Sotirou, S., and Salzmann, C. (2013a). Personalised learning spaces and federated online labs for STEM Education at School. In *Global Engineering Education Conference (EDUCON), 2013 IEEE*, pages 769–773.

Gillet, D., de Jong, T., Sotirou, S., and Salzmann, C. (2013b). Personalised Learning Spaces and Federated Online Labs for STEM Education at School: Supporting Teacher Communities and Inquiry Learning. In *Proceedings of the 4th IEEE Global Engineering Education Conference (EDUCON)*, pages 769–773, Berlin, Germany. IEEE.

Girard, J. and Girard, J. (2015). Defining knowledge management: Toward an applied compendium. 3(1):20.

Goodlad, J. I. (1983). A Study of Schooling: Some Findings and Hypotheses. *Phi Delta Kappan*, 64(7):465–70.

Govaerts, S., Cao, Y., Vozniuk, A., Holzer, A., Zutin, D. G., Ruiz, E. S. C., Bollen, L., Manske, S., Faltin, N., Salzmann, C., and others (2013a). Towards an Online Lab Portal for Inquiry-Based STEM Learning at School. In *Advances in Web-Based Learning–ICWL 2013*, pages 244–253. Springer.

Govaerts, S., Holzer, A., Bogdanov, E., Vozniuk, A., Li, N., Halimi, W., Cholleton, A., Gillet, D., Cao, Y., Bollen, L., and others (2013b). Specifications of the Go-Lab portal and app composer–initial. Technical report, Go-Lab Consortium.

Govaerts, S., Verbert, K., Dahrendorf, D., Ullrich, C., Schmidt, M., Werkle, M., Chatterjee, A., Nussbaumer, A., Renzel, D., Scheffel, M., Friedrich, M., Santos, J., Duval, E., and Law, E.-C. (2011). Towards Responsive Open Learning Environments: The ROLE Interoperability Framework. In Kloos, C., Gillet, D., Crespo García, R., Wild, F., and Wolpers, M., editors, *Towards Ubiquitous Learning*, volume 6964 of *Lecture Notes in Computer Science*, pages 125–138. Springer Berlin Heidelberg.

Govaerts, S., Verbert, K., Duval, E., and Pardo, A. (2012). The student activity meter for awareness and self-reflection. In *CHI'12 Extended Abstracts on Human Factors in Computing Systems*, pages 869–884.

Graf, S. and Bekele, R. (2006). Forming heterogeneous groups for intelligent collaborative learning systems with ant colony optimization. In *International conference on intelligent tutoring systems*, pages 217–226. Springer.

Grimm, S., Hitzler, P., and Abecker, A. (2007). Knowledge Representation and Ontologies. In Studer, R., Grimm, S., and Abecker, A., editors, *Semantic Web Services: Concepts, Technologies, and Applications*, pages 51–105. Springer Berlin Heidelberg, Berlin, Heidelberg.

Gurevych, I., Mühlhäuser, M., Müller, C., Steimle, J., Weimer, M., and Zesch, T. (2007). Darmstadt knowledge processing repository based on uima. In *Proceedings of the First Workshop on Unstructured Information Management Architecture at Biannual Conference of the Society for Computational Linguistics and Language Technology, Tübingen, Germany*, page 89.

Gustavsson, I., Zackrisson, J., H\a akansson, L., Claesson, I., and Lagö, T. L. (2007a). The visir project–an open source software initiative for distributed online laboratories. In *REV 2007*.

Gustavsson, I. and, Z. J., Håkansson, L., Claesson, L., and Lago, T. (2007b). The VISIR project — an Open Source Software Initiative for Distributed Online Laboratories. In *Int. Conf. on Remote Engineering and Virtual Instrumentation*.

Gutierrez-Santos, S., Mavrikis, M., and Magoulas, G. (2010). Layered development and evaluation for intelligent support in exploratory environments: the case of microworlds. In *International Conference on Intelligent Tutoring Systems*, pages 105–114. Springer.

Göhnert, T., Harrer, A., Hecking, T., and Hoppe, H. U. (2013). A Workbench to Construct and Re-use Network Analysis Workflows: Concept, Implementation, and Example Case. In *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, ASONAM '13, pages 1464–1466, Niagara, Ontario, Canada. ACM.

Halabi, Y. (2004). The expansion of global governance into the third world: altruism, realism, or constructivism? *International Studies Review*, 6(1):21–48.

Hannafin, M., Hill, J., Land, S., and Lee, E. (2014). Student-Centered, Open Learning Environments: Research, Theory, and Practice. In Spector, J. M., Merrill, M. D., Elen, J., and Bishop, M. J., editors, *Handbook of Research on Educational Communications and Technology*, pages 641–651. Springer New York.

Harrer, A. (2013). Analytics of collaborative planning in metafora: Architecture, data, and analytic methods. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, pages 255–259. ACM.

Harrer, A. and Göhnert, T. (2015). Integrated representations and small data: towards contextualized and embedded analytics tools for learners. In *Proceedings of LAK*, pages 406–407. ACM.

Harrer, A. and Malzahn, N. (2006). Bridging the gap-towards a graphical modeling language for learning designs and collaboration scripts of various granularities. In *Advanced Learning Technologies, 2006. Sixth International Conference on*, pages 296–300. IEEE.

Harrer, A., Malzahn, N., and Hoppe, H. U. (2007). Graphical modeling and simulation of learning designs. *Frontiers in Artificial Intelligence and Applications*, 162:291.

Harrison, D. A. and Klein, K. J. (2007a). What's the difference? diversity constructs as separation, variety, or disparity in organizations. *Academy of Management Review*, 32(4):1199–1228.

Harrison, D. A. and Klein, K. J. (2007b). What's the Difference? Diversity Constructs as Separation, Variety, or Disparity in Organizations. *The Academy of Management Review*, 32(4):1199–1228.

Harrison, D. A. and Sin, H. (2006). What is diversity and how should it be measured. *Handbook of workplace diversity*, pages 191–216.

Harward, V., del Alamo, J., Lerman, S., Bailey, P., Carpenter, J., DeLong, K., Felknor, C., Hardison, J., Harrison, B., Jabbour, I., Long, P., Mao, T., Naamani, L., Northridge, J., Schulz, M., Talavera, D., Varadharajan, C., Wang, S., Yehia, K., Zbib, R., and

Zych, D. (2008). The iLab Shared Architecture: A Web Services Infrastructure to Build Communities of Internet Accessible Laboratories. *Proceedings of the IEEE*, 96(6):931–950.

Haya, P. A., Daems, O., Malzahn, N., Castellanos, J., and Hoppe, H. U. (2015). Analysing content and patterns of interaction for improving the learning design of networked learning environments. *British Journal of Educational Technology*, 46(2):300–316.

Hecking, T. and Hoppe, H. U. (2015). A Network Based Approach for the Visualization and Analysis of Collaboratively Edited Texts. In *VISLA@ LAK*, pages 19–23.

Hecking, T., Manske, S., Bollen, L., Govaerts, S., Vozniuk, A., and Hoppe, H. U. (2014). A flexible and extendable learning analytics infrastructure. In *International Conference on Web-Based Learning*, pages 123–132. Springer.

Heintz, M., Law, E. L.-C., Govaerts, S., Holzer, A., and Gillet, D. (2014). Pdot: participatory design online tool. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*, pages 2581–2586. ACM.

Helbig, H. (2006). *Knowledge representation and the semantics of natural language*. Springer.

Herlocker, J. L., Konstan, J. A., Terveen, L. G., and Riedl, J. T. (2004). Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems (TOIS)*, 22(1):5–53.

Hernández-Leo, D., Asensio-Pérez, J. I., Dimitriadis, Y., and Villasclaras-Fernández, E. D. (2010). Generating cscl scripts: From a conceptual model of pattern languages to the design of real scripts. In *Technology-Enhanced Learning*, pages 49–64. Brill Sense.

Hicks, D. and Holden, C. (2007). *Teaching the global dimension: Key principles and effective practice*. Routledge.

Hill, M. O. (1973). Diversity and evenness: a unifying notation and its consequences. *Ecology*, 54(2):427–432.

Hinze, U., Bischoff, M., and Blakowski, G. (2002). Jigsaw method in the context of cscl. In *EdMedia+ Innovate Learning*, pages 789–794. Association for the Advancement of Computing in Education (AACE).

Hmelo-Silver, C. E., Duncan, R. G., and Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: a response to kirschner, sweller, and. *Educational psychologist*, 42(2):99–107.

Hodgins, W. and Duval, E. (2002). Draft standard for learning technology-learning object metadata-iso/iec 11404. *IEEE P1484. 12.2/D1*.

Hodson, D. (1993). Re-thinking old ways: Towards a more critical approach to practical work in school science.

Hodson, D. (2001). Research on practical work in school and universities: In pursuit of better questions and better methods. In *Proceedings of the 6th European Conference on Research in Chemical Education, University of Aveiro, Aviero, Portugal.*

Hofstein, A. (2004). The laboratory in chemistry education: Thirty years of experience with developments, implementation, and research. *Chemistry education research and practice*, 5(3):247–264.

Hofstein, A., Navon, O., Kipnis, M., and Mamlok-Naaman, R. (2005). Developing students' ability to ask more and better questions resulting from inquiry-type chemistry laboratories. *Journal of research in science teaching*, 42(7):791–806.

Hollan, J., Hutchins, E., and Kirsh, D. (2000). Distributed cognition: Toward a new foundation for human-computer interaction research. *ACM Trans. Comput.-Hum. Interact.*, 7(2):174–196.

Holton, D. and Clarke, D. (2006). Scaffolding and metacognition. *International journal of mathematical education in science and technology*, 37(2):127–143.

Hoppe, H. U. (1995). The use of multiple student modeling to parameterize group learning.

Hoppe, H. U. (2017). The Handbook of Learning Analytics. pages 23–33. Society for Learning Analytics Research (SoLAR), Alberta, Canada, 1 edition.

Hoppe, H. U., Engler, J., and Weinbrenner, S. (2012). The impact of structural characteristics of concept maps on automatic quality measurement. In *Proceedings of the 10th international conference of the learning sciences (ICLS). Sydney, Australia: ISLS.*

Hoppe, H. U. and Ploetzner, R. (1999). Can analytic models support learning in groups. *Collaborative-learning: Cognitive and computational approaches*, pages 147–168.

Huang, J., Zhou, M., and Yang, D. (2007). Extracting Chatbot Knowledge from Online Discussion Forums. In *IJCAI*, volume 7, pages 423–428.

Huppert, J., Lomask, S. M., and Lazarowitz, R. (2002). Computer simulations in the high school: Students' cognitive stages, science process skills and academic achievement in microbiology. *International Journal of Science Education*, 24(8):803–821.

Hurlbert, S. H. (1971). The nonconcept of species diversity: a critique and alternative parameters. *Ecology*, 52(4):577–586.

Häkkinen, P. (2002). Challenges for design of computer–based learning environments. *British Journal of Educational Technology*, 33(4):461–469.

Ingerman, Å., Linder, C., and Marshall, D. (2009). The learners' experience of variation: following students' threads of learning physics in computer simulation sessions. *Instructional science*, 37(3):273–292.

Inmon, W. H. (2005). *Building the data warehouse*. John wiley & sons.

Janssen, J. and Bodemer, D. (2013). Coordinated computer-supported collaborative learning: Awareness and awareness tools. *Educational Psychologist*, 48(1):40–55.

Jehn, K. A., Northcraft, G. B., and Neale, M. A. (1999). Why differences make a difference: A field study of diversity, conflict and performance in workgroups. *Administrative science quarterly*, 44(4):741–763.

Jeong, H. and Hmelo-Silver, C. E. (2010). An overview of cscl methodologies. In *Proceedings of the 9th International Conference of the Learning Sciences-Volume 1*, pages 921–928. International Society of the Learning Sciences.

Jermann, P., Soller, A., Muehlenbrock, M., and others (2001). From mirroring to guiding: A review of the state of art technology for supporting collaborative learning. In *Proceedings of the European Conference on Computer-Supported Collaborative Learning EuroCSCL-2001. Maastricht, The Netherlands*, pages 324–331.

Jia, J. (2009). CSIEC: A computer assisted English learning chatbot based on textual knowledge and reasoning. *Knowledge-Based Systems*, 22(4):249–255.

Johansson-Stenman, O. (2005). Global environmental problems, efficiency and limited altruism. *Economics Letters*, 86(1):101–106.

Johnson, R. B. and Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational researcher*, 33(7):14–26.

Kahrimanis, G., Avouris, N., and Komis, V. (2011). Interaction analysis as a tool for supporting collaboration: An overview. In *Technology-enhanced systems and tools for collaborative learning scaffolding*, pages 93–114. Springer.

Kahrimanis, G., Chounta, I.-A., and Avouris, N. (2010). Study of correlations between logfile-based metrics of interaction and the quality of synchronous collaboration. *Guest Editors*, 24.

Kamada, T. and Kawai, S. (1989). An algorithm for drawing general undirected graphs. *Inf. Process. Lett.*, 31(1):7–15.

Khambadkone, S. G., Cordner, Z. A., Dickerson, F., Severance, E. G., Prandovszky, E., Pletnikov, M., Xiao, J., Li, Y., Boersma, G. J., Talbot, C. C., Campbell, W. W., Wright, C. S., Siple, C. E., Moran, T. H., Tamashiro, K. L., and Yolken, R. H. (2018). Nitrated meat products are associated with mania in humans and altered behavior and brain gene expression in rats. *Molecular Psychiatry*.

Kienle, A. and Wessner, M. (2006). The cscl community in its first decade: development, continuity, connectivity. *International Journal of Computer-Supported Collaborative Learning*, 1(1):9–33.

Kim, H. J. and Pedersen, S. (2011). Advancing young adolescents' hypothesis-development performance in a computer-supported and problem-based learning environment. *Computers & Education*, 57(2):1780–1789.

Kim, J., Jo, I.-H., and Park, Y. (2016). Effects of learning analytics dashboard: analyzing the relations among dashboard utilization, satisfaction, and learning achievement. *Asia Pacific Education Review*, 17(1):13–24.

Kirschner, P. A., Sweller, J., and Clark, R. E. (2006). Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. *Educational Psychologist*, 41(2):75–86.

Kizilcec, R. F. (2013). Collaborative Learning in Geographically Distributed and In-person Groups. In *AIED 2013 Workshops Proceedings Volume*, page 67.

Klahr, D. and Dunbar, K. (1988). Developmental differences in scientific discovery processes. In *Complex information processing*, pages 129–164. Psychology Press.

Knight, S., Buckingham Shum, S., and Littleton, K. (2013). Epistemology, pedagogy, assessment and learning analytics. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, pages 75–84. ACM.

Kobbe, L., Weinberger, A., Dillenbourg, P., Harrer, A., Hämäläinen, R., Häkkinen, P., and Fischer, F. (2007). Specifying computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning*, 2(2-3):211–224.

Koedinger, K. R., Baker, R. S., Cunningham, K., Skogsholm, A., Leber, B., and Stamper, J. (2010). A data repository for the edm community: The pslc datashop. *Handbook of educational data mining*, 43:43–56.

Konert, J., Bellhäuser, H., Röpke, R., Gallwas, E., and Zucik, A. (2016). Moodlepeers: Factors relevant in learning group formation for improved learning outcomes, satisfaction and commitment in e-learning scenarios using groupal. In *European Conference on Technology Enhanced Learning*, pages 390–396. Springer.

Konert, J., Burlak, D., and Steinmetz, R. (2014). The Group Formation Problem: An Algorithmic Approach to Learning Group Formation. In Rensing, C., de Freitas, S., Ley, T., and Muñoz-Merino, P., editors, *Open Learning and Teaching in Educational Communities*, volume 8719 of *Lecture Notes in Computer Science*, pages 221–234. Springer International Publishing.

Konrad, A. M., Prasad, P., and Pringle, J. (2005). *Handbook of Workplace Diversity*. SAGE. Google-Books-ID: Z1OUnWcWjhYC.

Koren, Y., Bell, R., and Volinsky, C. (2009). Matrix factorization techniques for recommender systems. *Computer*, (8):30–37.

Koschmann, T. (1996). *CSCL: Theory and practice of an emerging paradigm*. Routledge.

Koschmann, T. (2002). Dewey's contribution to the foundations of cscl research. In *Proceedings of the Conference on Computer Support for Collaborative Learning: Foundations for a CSCL Community*, pages 17–22. International Society of the Learning Sciences.

Koschmann, T., Zemel, A., Conlee-Stevens, M., Young, N. P., Robbs, J. E., and Barnhart, A. (2005). How do people learn? In *Barriers and biases in computer-mediated knowledge communication*, pages 265–294. Springer.

Kraan, W., Sherlock, D., and others (2013). Analytics tools and infrastructure. *JISC CETIS Anal Ser*, 1(11):1–24.

Kruckeberg, R. (2006). A Deweyan perspective on science education: Constructivism, experience, and why we learn science. *Science & Education*, 15(1):1–30.

Kuhlthau, C., Maniotes, L., and CASPARI, A. (2007). *Guided inquiry: Learning in the 21st century*. Greenwood Publishing Group.

Lajoie, S. P., Guerrera, C., Munsie, S. D., and Lavigne, N. C. (2001). Constructing knowledge in the context of bioworld. *Instructional Science*, 29(2):155–186.

Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. *Criticism and the growth of knowledge*, 4:91–196.

Lakkala, M. (2010). *How to design educational settings to promote collaborative inquiry: Pedagogical infrastructures for technology-enhanced progressive inquiry*. PhD Thesis, Institute of Behavioural Sciences, University of Helsinki, Finland.

Lazonder, A. W. (2014). Inquiry learning. In *Handbook of research on educational communications and technology*, pages 453–464. Springer.

Lazonder, A. W. and Harmsen, R. (2016). Meta-Analysis of Inquiry-Based Learning: Effects of Guidance. *Review of Educational Research*, 86(3):681–718.

Lee, J. H. et al. (1993). Information retrieval based on conceptual distance in is-a hierarchies. *Journal of Documentation*, 49(2):188–207.

Lee, M. J. and McLoughlin, C. (2007). Teaching and learning in the Web 2.0 era: Empowering students through learner-generated content. *International Journal of Instructional Technology and Distance Learning*, 4(10):21–34.

Leeman-Munk, S. P., Wiebe, E. N., and Lester, J. C. (2014). Assessing Elementary Students' Science Competency with Text Analytics. In *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge*, LAK '14, pages 143–147, Indianapolis, Indiana. ACM.

Lejeune, A., Ney, M., Weinberger, A., Pedaste, M., Bollen, L., Hovardas, T., Hoppe, H. U., and de Jong, T. (2009a). A Graphical Modeling Language for Computer-Based Learning Scenarios. In *2009 Ninth IEEE International Conference on Advanced Learning Technologies*, pages 734–735.

Lejeune, A., Ney, M., Weinberger, A., Pedaste, M., Bollen, L., Hovardas, T., Hoppe, H. U., and de Jong, T. (2009b). Learning Activity Spaces: Towards Flexibility in Learning Design? In *Advanced Learning Technologies, 2009. ICALT 2009. Ninth IEEE International Conference on*, pages 433–437.

Lerman, R. I. and Yitzhaki, S. (1984). A note on the calculation and interpretation of the gini index. *Economics Letters*, 15(3-4):363–368.

Lewoniewski, W., Węcel, K., and Abramowicz, W. (2016). Quality and importance of wikipedia articles in different languages. In *International Conference on Information and Software Technologies*, pages 613–624. Springer.

Lin, D. (1998). An information-theoretic definition of similarity. In *Icml*, volume 98, pages 296–304. Citeseer.

Lipponen, L. (1999). The Challenges for Computer Supported Collaborative Learning in Elementary and Secondary Level: Finnish Perspectives. In *Proceedings of the 1999 Conference on Computer Support for Collaborative Learning*, CSCL '99, Palo Alto, California. International Society of the Learning Sciences.

Lo, R. T.-W., He, B., and Ounis, I. (2005). Automatically building a stopword list for an information retrieval system. In *Journal on Digital Information Management: Special Issue on the 5th Dutch-Belgian Information Retrieval Workshop (DIR)*, volume 5, pages 17–24.

Long, Y. and Aleven, V. (2013). Supporting students' self-regulated learning with an open learner model in a linear equation tutor. In *International Conference on Artificial Intelligence in Education*, pages 219–228. Springer.

Lou, Y., Abrami, P. C., and d'Apollonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of educational research*, 71(3):449–521.

Lowe, D., Murray, S., Weber, L., de la Villefromoy, M., Johnston, A., Lindsay, E., Nageswaran, W., and Nafalski, A. (2009). LabShare: Towards a national approach to laboratory sharing. In *Proceedings of the 20th Annual Conference for the Australasian Association for Engineering Education*, pages 458–463. The School of Mechanical Engineering, University of Adelaide.

Ludvigsen, S., Stahl, G., Law, N., and Cress, U. (2015). From the editors: Collaboration and the formation of new knowledge artifacts. *International Journal of Computer-Supported Collaborative Learning*, 10(1):1–6.

Lunetta, V. N. (1998). The school science laboratory: Historical perspectives and contexts for contemporary teaching. *International handbook of science education*, 1:249–264.

Lunetta, V. N., Hofstein, A., and Clough, M. P. (2005). Learning and Teaching in the School Science Laboratory: An analysis of research, theory, and practice. page 70.

Lunetta, V. N., Hofstein, A., and Clough, M. P. (2007). Learning and teaching in the school science laboratory: An analysis of research, theory, and practice. *Handbook of research on science education*, 2.

Ma, J. and Nickerson, J. V. (2006). Hands-on, simulated, and remote laboratories: A comparative literature review. *ACM Comput. Surv.*, 38(3).

MacAskill, W. (2015). *Doing good better: Effective altruism and a radical new way to make a difference.* Guardian Faber Publishing.

Mäeots, M., Siiman, L., Kori, K., Eelmets, M., Pedaste, M., and Anjewierden, A. (2016). The role of a reflection tool in enhancing students reflection. In *10th annual International Technology, Education and Development Conference (INTED 2016).*

Magurran, A. E. (1988). Why diversity? In Magurran, A. E., editor, *Ecological Diversity and Its Measurement*, pages 1–5. Springer Netherlands, Dordrecht.

Malzahn, N., Ganster, T., Sträfling, N., Krämer, N., and Hoppe, H. U. (2013). Motivating Students or Teachers? In Hernández-Leo, D., Ley, T., Klamma, R., and Harrer, A., editors, *Scaling up Learning for Sustained Impact*, volume 8095 of *Lecture Notes in Computer Science*, pages 191–204. Springer Berlin Heidelberg.

Manning, C., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S., and McClosky, D. (2014). The stanford corenlp natural language processing toolkit. In *Proceedings of 52nd annual meeting of the association for computational linguistics: system demonstrations*, pages 55–60.

Manouselis, N., Drachsler, H., Vuorikari, R., Hummel, H., and Koper, R. (2011). Recommender systems in technology enhanced learning. In *Recommender systems handbook*, pages 387–415. Springer.

Manske, S. and Cao, Y. (2015a). Go-Lab Releases of the Learning Analytics, Scaffolding Services, and Add-on Services-Final. Technical report, Go-Lab Project.

Manske, S. and Cao, Y. (2015b). Go-Lab Specifications of the Learning Analytics, Scaffolding Services, and Add-on Services-Final. Technical report, Go-Lab Project.

Manske, S., Chounta, I.-A., Rodríguez-Triana, M. J., Gillet, D., and Hoppe, H. U. (2015a). Exploring Deviation in Inquiry Learning: Degrees of Freedom or Source of Problems? In *Proceedings of the 23rd International Conference on Computers in Education*. Asia-Pacific Society for Computers in Education.

Manske, S., Hecking, T., Bollen, L., Gohnert, T., Ramos, A., and Hoppe, H. U. (2014). A Flexible Framework for the Authoring of Reusable and Portable Learning Analytics Gadgets. In *Advanced Learning Technologies (ICALT), 2014 IEEE 14th International Conference on*, pages 254–258. IEEE.

Manske, S., Hecking, T., and Hoppe, H. U. (2015b). Predictive Knowledge Modeling in Collaborative Inquiry Learning Scenarios. In *International Conference on Artificial Intelligence in Education*, pages 706–709. Springer, Cham.

Manske, S., Hecking, T., Hoppe, U., Chounta, I.-A., and Werneburg, S. (2015c). Using Differences to Make a Difference: A Study in Heterogeneity of Learning Groups. In *11th International Conference on Computer Supported Collaborative Learning (CSCL 2015)*.

Manske, S. and Hoppe, H. U. (2016). The "Concept Cloud": Supporting Collaborative Knowledge Construction based on Semantic Extraction from Learner-generated Artefacts. In *The 16th IEEE International Conference on Advanced Learning Technologies - ICALT2016*. University of Duisburg-Essen.

Manske, S. and Hoppe, H. U. (2017). Managing Knowledge Diversity: Towards Automatic Semantic Group Formation. In *Advanced Learning Technologies (ICALT), 2017 IEEE 17th International Conference on*, pages 330–332. IEEE.

Marshall, J. A. and Young, E. S. (2006). Preservice teachers' theory development in physical and simulated environments. *Journal of Research in Science Teaching*, 43(9):907–937.

Marušić, M. and Sliško, J. (2012). Influence of three different methods of teaching physics on the gain in students' development of reasoning. *International Journal of Science Education*, 34(2):301–326.

Mason, L., Ariasi, N., and Boldrin, A. (2011). Epistemic beliefs in action: Spontaneous reflections about knowledge and knowing during online information searching and their influence on learning. *Learning and Instruction*, 21(1):137–151.

Mason, L., Boldrin, A., and Ariasi, N. (2010). Epistemic metacognition in context: evaluating and learning online information. *Metacognition and Learning*, 5(1):67–90.

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? *American psychologist*, 59(1):14.

McClure, J. R., Sonak, B., and Suen, H. K. (1999). Concept map assessment of classroom learning: Reliability, validity, and logistical practicality. *Journal of research in science teaching*, 36(4):475–492.

McMahan, J. (2016). Philosophical critiques of effective altruism. *The Philosophers' Magazine*, (73):92–99.

Meira, L. (1998). Making sense of instructional devices: The emergence of transparency in mathematical activity. *Journal for Research in Mathematics Education*, pages 121–142.

Melkonyan, A., Gampe, A., Pontual, M., Huang, G., and Akopian, D. (2014). Facilitating Remote Laboratory Deployments Using a Relay Gateway Server Architecture. *IEEE Transactions on Industrial Electronics*, 61(1):477–485.

Mendes, P. N., Jakob, M., and Bizer, C. (2012). Dbpedia: A multilingual cross-domain knowledge base. In *LREC*, pages 1813–1817. Citeseer.

Mendes, P. N., Jakob, M., García-Silva, A., and others (2011). DBpedia spotlight: shedding light on the web of documents. In *Proceedings of SEMANTiCS*, pages 1–8. ACM.

Miao, Y., Hoeksema, K., Hoppe, H. U., and Harrer, A. (2005). Cscl scripts: Modelling features and potential use. In *Proceedings of th 2005 conference on Computer support for collaborative learning: learning 2005: the next 10 years!*, pages 423–432. International Society of the Learning Sciences.

Miao, Y. and Hoppe, H. U. (2011). Applying learning design to work-based learning. *Learning, Media and Technology*, 36(2):189–205.

Miel, A. (1952). *Cooperative procedures in learning.* Bureau of Publications, Teachers College, Columbia University.

Miller, C. C., Burke, L. M., and Glick, W. H. (1998). Cognitive diversity among upper-echelon executives: implications for strategic decision processes. *Strategic management journal*, 19(1):39–58.

Miner, G., Elder IV, J., and Hill, T. (2012). *Practical text mining and statistical analysis for non-structured text data applications.* Academic Press.

Minner, D. D., Levy, A. J., and Century, J. (2009). Inquiry-based science instruction - what is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of Research in Science Teaching.*

Mitrovic, A. and Martin, B. (2007). Evaluating the effect of open student models on self-assessment. *International Journal of Artificial Intelligence in Education*, 17(2):121–144.

Moody, J., McFarland, D., and Bender-deMoll, S. (2005). Dynamic Network Visualisation. *AJS*, 110(4):1206–1241.

Moon, J. A. (2004). *A handbook of reflective and experiential learning: Theory and practice.* Psychology Press.

Mujkanovic, A., Lowe, D., and Willey, K. (2012). Adaptive group formation to promote desired behaviours. In *AAEE-Annual Conference of Australasian Association for Engineering Education.* Swinburne University of Technology.

Mulder, Y. G., Lazonder, A. W., and de Jong, T. (2014). Using heuristic worked examples to promote inquiry-based learning. *Learning and instruction*, 29:56–64.

Nehring, K. and Puppe, C. (2002). A Theory of Diversity. *Econometrica*, 70(3):1155–1198.

Newman Jr, W. J., Abell, S. K., Hubbard, P. D., McDonald, J., Otaala, J., and Martini, M. (2004). Dilemmas of teaching inquiry in elementary science methods. *Journal of Science teacher education*, 15(4):257–279.

Novak, J. D. (1984). *Learning how to learn.* Cambridge University Press.

Novak, J. D. (1988). Learning science and the science of learning.

Novak, J. D. and Cañas, A. J. (2004). *Building on New Constructivist Ideas and CmapTools to Create a New Model for Education 1.*

O'Donnell, A. M., Dansereau, D. F., and Hall, R. H. (2002). Knowledge Maps as Scaffolds for Cognitive Processing. *Educational Psychology Review*, 14(1):71–86.

OECD, Analysis, D. E., and Division, S. (2003). *OECD Science, Technology and Industry Scoreboard 2003.* Canongate US.

O'loughlin, M. (1992). Rethinking science education: Beyond Piagetian constructivism toward a sociocultural model of teaching and learning. *Journal of research in science teaching*, 29(8):791–820.

Olsen, J. E. and Martins, L. L. (2012). Understanding organizational diversity management programs: A theoretical framework and directions for future research. *Journal of Organizational Behavior*, 33(8):1168–1187.

Olympiou, G., Zacharias, Z. C., et al. (2013). Making the invisible visible: Enhancing students' conceptual understanding by introducing representations of abstract objects in a simulation. *Instructional science*, 41(3):575–596.

Orduña, P., García-Zubia, J., Rodriguez-Gil, L., Irurzun, J., López-de Ipiña, D., and Gazzola, F. (2012). Using LabVIEW remote panel in remote laboratories: Advantages and disadvantages. In *Proceedings of the 2012 IEEE Global Engineering Education Conference (EDUCON)*, pages 1–7.

Ounnas, A., Davis, H. C., and Millard, D. E. (2009). A Framework for Semantic Group Formation in Education. *Educational Technology & Society*, 12(4):43–55.

Overdijk, M., van Diggelen, W., Kirschner, P. A., and Baker, M. (2012). Connecting agents and artifacts in cscl: Towards a rationale of mutual shaping. *International Journal of Computer-Supported Collaborative Learning*, 7(2):193–210.

Page, S. E. (2010). *Diversity and Complexity*. Princeton University Press. Google-Books-ID: Mi6zkXss14IC.

Parker, G. G. (2005). Two-Sided Network Effects: A Theory of Information Product Design. *Management Science*, Vol. 51(No. 10):Pages 1494–1504.

Parker, K. and Chao, J. (2007). Wiki as a Teaching Tool. *Interdisciplinary Journal of E-Learning and Learning Objects*, 3(1):57–72.

Pedaste, M., de Jong, T., Sarapuu, T., Piksoot, J., van Joolingen, W. R., and Giemza, A. (2013). Investigating Ecosystems as a Blended Learning Experience. *Science*, 340(6140):1537–1538.

Pedaste, M., Mäeots, M., Siiman, L. A., de Jong, T., van Riesen, S. A., Kamp, E. T., Manoli, C. C., Zacharia, Z. C., and Tsourlidaki, E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational Research Review*, 14(0):47–61.

Pedaste, M. and Sarapuu, T. (2006). Developing an effective support system for inquiry learning in a Web-based environment. *Journal of Computer Assisted Learning*, 22(1):47–62.

Piaget, J. (1970). Science of education and the psychology of the child. trans. d. coltman.

Piaget, J. and Cook, M. (1952). *The origins of intelligence in children*, volume 8. International Universities Press New York.

Podolefsky, N. S., Moore, E. B., and Perkins, K. K. (2013). Implicit scaffolding in interactive simulations: Design strategies to support multiple educational goals. *arXiv preprint arXiv:1306.6544*.

Press, N. (2004). *Understanding Metadata*. National Information Standards Organization Press, Bethesda.

Puntambekar, S. and Hubscher, R. (2005). Tools for scaffolding students in a complex learning environment: What have we gained and what have we missed? *Educational psychologist*, 40(1):1–12.

Qiu, G., Laurini, R., Leung, C., and others (2007). *Advances in Visual Information Systems: VISUAL, Revised Selected Paper*, volume 4781. Springer.

Quintana, C., Reiser, B. J., Davis, E. A., Krajcik, J., Fretz, E., Duncan, R. G., Kyza, E., Edelson, D., and Soloway, E. (2004). A Scaffolding Design Framework for Software to Support Science Inquiry. *Journal of the Learning Sciences*, 13(3):337–386.

Quinto, I., Buckingham Shum, S., De Liddo, A., and Iandoli, L. (2010). A debate dashboard to enhance on-line knowledge sharing. *VINE The Journal of information and Knowledge Management Systems*, 42:67–93.

Rada, R., Mili, H., Bicknell, E., and Blettner, M. (1989). Development and application of a metric on semantic nets. *IEEE transactions on systems, man, and cybernetics*, 19(1):17–30.

Ramirez, C. and Valdes, B. (2012). A general knowledge representation model of concepts. In *Advances in knowledge representation*. InTech.

Reid, D. J., Zhang, J., and Chen, Q. (2003). Supporting scientific discovery learning in a simulation environment. *Journal of Computer Assisted Learning*, 19(1):9–20.

Renzel, D. and Klamma, R. (2013). From micro to macro: Analyzing activity in the ROLE sandbox. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, pages 250–254. ACM.

Resnik, P. (1999). Semantic Similarity in a Taxonomy: An Information-Based Measure and its Application to Problems of Ambiguity in Natural Language. *Journal of Artificial Intelligence Research*, 11:95–130.

Resta, P. and Laferrière, T. (2007). Technology in support of collaborative learning. *Educational Psychology Review*, 19(1):65–83.

Richter, T., Boehringer, D., and Jeschke, S. (2011). LiLa: A European Project on Networked Experiments. In Jeschke, S., Isenhardt, I., and Henning, K., editors, *Automation, Communication and Cybernetics in Science and Engineering 2009/2010*, pages 307–317. Springer Berlin Heidelberg.

Richter, T., Grube, P., and Zutin, D. (2012). A Standardized Metadata Set for Annotation of Virtual and Remote Laboratories. In *The Seventh IEEE International Workshop on Multimedia Technologies for E-Learning.*

Rieber, L. P. and Parmley, M. W. (1995). To teach or not to teach? comparing the use of computer-based simulations in deductive versus inductive approaches to learning with adults in science. *Journal of Educational Computing Research*, 13(4):359–374.

Rivadeneira, A. W., Gruen, D. M., Muller, M. J., and others (2007). Getting our head in the clouds: toward evaluation studies of tagclouds. In *Proceedings of SIGCHI*, pages 995–998. ACM.

Rocard, M., Csermely, P., Jorde, D., Lenzen, D., Walberg-Henriksson, H., and Hemmo, V. (2007). Science education now: A renewed pedagogy for the future of Europe. Technical report, European Commission: Directorate-General for Research.

Rocard, M., Csermely, P., Jorde, D., Lenzen, D., Walwerg-Heriksson, H., and Hemmo, V. (2006). Rocard Report: Science Education Now: a new Pedagogy for the Future of Europe. Technical report, European Commission.

Rodriguez, A. J. (1997). The dangerous discourse of invisibility: A critique of the National Research Council's National Science Education Standards. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, 34(1):19–37.

Rodríguez-Triana, M. J., Govaerts, S., Halimi, W., Holzer, A. C., Salzmann, C., Vozniuk, A., de Jong, T., Sotirou, S., and Gillet, D. (2014). Rich Open Educational Resources for Personal and Inquiry Learning. Agile Creation, Sharing and Reuse in Educational Social Media Platforms. In *International Conference on Web & Open Access to Learning*, Dubai, United Arab Emirates.

Rodríguez-Triana, M. J., Holzer, A., Vozniuk, A., and Gillet, D. (2015). Orchestrating inquiry-based learning spaces: An analysis of teacher needs. In *International Conference on Web-Based Learning*, pages 131–142. Springer.

Roschelle, J., Dimitriadis, Y., and Hoppe, H. U. (2013). Classroom Orchestration: Synthesis. *Comput. Educ.*, 69:523–526.

Rosé, C., Wang, Y.-C., Cui, Y., Arguello, J., Stegmann, K., Weinberger, A., and Fischer, F. (2008). Analyzing collaborative learning processes automatically: Exploiting the advances of computational linguistics in computer-supported collaborative learning. *IJCSCL*, 3(3):237–271.

Roth, W.-M. and Bowen, G. M. (1995). Knowing and Interacting: A Study of Culture, Practices, and Resources in a Grade 8 Open-Inquiry Science Classroom Guided by a Cognitive Apprenticeship Metaphor. *Cognition and Instruction*, 13(1):73–128.

Rushton, J. P., Chrisjohn, R. D., and Fekken, G. C. (1981). The altruistic personality and the self-report altruism scale. *Personality and individual differences*, 2(4):293–302.

Russell, S. J. and Norvig, P. (2016). *Artificial Intelligence : A Modern Approach*. Malaysia; Pearson Education Limited,.

Rutten, N., van Joolingen, W. R., and van Der Veen, J. T. (2012). The learning effects of computer simulations in science education. *Computers & Education*, 58(1):136–153.

Ryan, J. (2006). *Leading Diverse Schools*. Springer Science & Business Media. Google-Books-ID: cDnhBwAAQBAJ.

Sadeghi, H. and Kardan, A. A. (2014). Toward effective group formation in computer-supported collaborative learning. *Interactive Learning Environments*, ahead-of-print:1–14.

Sancristobal, E., Castro, M., Martin, S., Tawkif, M., Pesquera, A., Gil, R., Díaz, G., and Peire, J. (2011). Remote labs as learning services in the educational arena. In *2011 IEEE Global Engineering Education Conference (EDUCON)*, pages 1189–1194.

Sandoval, W. A. and Reiser, B. J. (2004). Explanation-driven inquiry: Integrating conceptual and epistemic scaffolds for scientific inquiry. *Science Education*, 88(3):345–372.

Santana, I., Ferre, M., Izaguirre, E., Aracil, R., and Hernandez, L. (2013). Remote Laboratories for Education and Research Purposes in Automatic Control Systems. *IEEE Transactions on Industrial Informatics*, 9(1):547–556.

Saunders, W. L. (1992). The constructivist perspective: Implications and teaching strategies for science. *School Science and Mathematics*, 92(3):136–141.

Sawyer, R. K. (2005). *The Cambridge handbook of the learning sciences*. Cambridge University Press.

Scalise, K., Timms, M., Moorjani, A., Clark, L., Holtermann, K., and Irvin, P. S. (2011). Student learning in science simulations: Design features that promote learning gains. *Journal of Research in Science Teaching*, 48(9):1050–1078.

Scardamalia, M. and Bereiter, C. (1991). Higher levels of agency for children in knowledge building: A challenge for the design of new knowledge media. *The Journal of the learning sciences*, 1(1):37–68.

Scardamalia, M. and Bereiter, C. (1994). Computer support for knowledge-building communities. *The journal of the learning sciences*, 3(3):265–283.

Scardamalia, M. and Bereiter, C. (2006). Knowledge building. *The Cambridge*.

Scheiter, K. and Gerjets, P. (2007). Learner Control in Hypermedia Environments. *Educational Psychology Review*, 19(3):285–307.

Schilpzand, M. C. (2010). *Cognitive diversity and team performance: the roles of team mental models and information processing mechanisms*. PhD Thesis, Georgia Institute of Technology.

Schittekatte, M. and van Hiel, A. (1996). Effects of partially shared information and awareness of unshared information on information sampling. *Small Group Research*, 27(3):431–449.

Schneegass, C., Kizina, A., Manske, S., and Hoppe, H. U. (2016). ConceptCloud: Supporting Reflection in the Online Learning Environment Go-Lab. In *14. E-Learning Fachtagung Informatik*, pages 83–83.

Schweingruber, H. A., Hilton, M. L., Singer, S. R., and others (2005). *America's Lab Report:: Investigations in High School Science*. National Academies Press.

Schwendimann, B. (2015). Concept maps as versatile tools to integrate complex ideas: From kindergarten to higher and professional education. *Knowledge Management and E-Learning*, 7(1):73–99.

Seidman, I. (2006). *Interviewing as qualitative research: A guide for researchers in education and the social sciences*. Teachers college press.

Sharan, Y. and Sharan, S. (1992). *Expanding cooperative learning through group investigation*, volume 1234. Teachers College Press New York.

Sharples, M. (2013). Shared orchestration within and beyond the classroom. *Computers and Education*, 69:504–506.

Shawar, B. A. and Atwell, E. (2007). Different measurements metrics to evaluate a chatbot system. In *Proceedings of the workshop on bridging the gap: Academic and industrial research in dialog technologies*, pages 89–96. Association for Computational Linguistics.

Sherin, B., Reiser, B. J., and Edelson, D. (2004). Scaffolding analysis: Extending the scaffolding metaphor to learning artifacts. *The Journal of the Learning Sciences*, 13(3):387–421.

Siemens, G. (2012). Learning Analytics: Envisioning a Research Discipline and a Domain of Practice. In *Proceedings of the 2Nd International Conference on Learning Analytics and Knowledge*, LAK '12, pages 4–8, Vancouver, British Columbia, Canada. ACM.

Siemens, G., Gasevic, D., Haythornthwaite, C., Dawson, S. P., Shum, S., Ferguson, R., Duval, E., Verbert, K., Baker, R., and others (2011). Open Learning Analytics: an integrated & modularized platform.

Sigel, I. E. (1978). Constructivism and teacher education. *The Elementary School Journal*, 78(5):333–338.

Simon, H. (2001). Learning to research about learning. *Cognition and instruction: Twenty-five years of progress*, pages 205–226.

Simpson, E. H. (1949). Measurement of diversity. *nature*.

Simsek, D., Shum, S. B., De Liddo, A., Ferguson, R., and Sándor, A. (2014). Visual Analytics of Academic Writing. In *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge*, LAK '14, pages 265–266, Indianapolis, Indiana, USA. ACM.

Singer, P. (2015). *The most good you can do: How effective altruism is changing ideas about living ethically*. Text Publishing.

Slotta, J. D. and Linn, M. C. (2009). *WISE Science: Web-Based Inquiry in the Classroom*. Teachers College Press, New York, NY, USA.

Smetana, L. K. and Bell, R. L. (2012). Computer simulations to support science instruction and learning: A critical review of the literature. *International Journal of Science Education*, 34(9):1337–1370.

Snell, J. and Prodromou, E. (2017). Activity streams 2.0. *World Wide Web Consortium, Working Draft WD-activitystreams-core-20150722 (July 2015)*.

Sowa, J. F. (2011). Peirce's tutorial on existential graphs. *Semiotica*, 2011(186).

Spörer, N. and Brunstein, J. C. (2006). Erfassung selbstregulierten Lernens mit Selbstberichtsverfahren: Ein Überblick zum Stand der Forschung. *Zeitschrift für pädagogische Psychologie*, 20(3):147–160.

Stahl, G. (2002). *Computer Support for Collaborative Learning: Foundations for a Cscl Community (Cscl 2002 Proceedings)*. Psychology Press.

Stahl, G., Koschmann, T., and Suthers, D. D. (2006). Computer-supported collaborative learning: An historical perspective. *Cambridge handbook of the learning sciences*, 2006:409–426.

Stahl, G., Ludvigsen, S., Law, N., and Cress, U. (2014). Cscl artifacts. *International Journal of Computer-Supported Collaborative Learning*, 9(3):237–245.

Stamper, J., Koedinger, K., Baker, R. S. J. d., Skogsholm, A., Leber, B., Rankin, J., and Demi, S. (2010). PSLC DataShop: A Data Analysis Service for the Learning Science Community. In Aleven, V., Kay, J., and Mostow, J., editors, *Intelligent Tutoring Systems*, pages 455–455, Berlin, Heidelberg. Springer Berlin Heidelberg.

Steele, J. and Iliinsky, N. (2010). *Beautiful visualization: looking at data through the eyes of experts*. " O'Reilly Media, Inc.".

Stoddart, T., Abrams, R., Gasper, E., and Canaday, D. (2000). Concept maps as assessment in science inquiry learning-a report of methodology. *International Journal of Science Education*, 22(12):1221–1246.

Stoyanov, S., Jablokow, K., Rosas, S. R., Wopereis, I. G., and Kirschner, P. A. (2017). Concept mapping—An effective method for identifying diversity and congruity in cognitive style. *Evaluation and program planning*, 60:238–244.

Suthers, D. D. (2006). Technology affordances for intersubjective meaning making: A research agenda for cscl. *International Journal of Computer-Supported Collaborative Learning*, 1(3):315–337.

Suthers, D. D. and Verbert, K. (2013). Learning Analytics As a "Middle Space". In *Proceedings of the Third International Conference on Learning Analytics and Knowledge*, LAK '13, pages 1–4, Leuven, Belgium. ACM.

Sutter, B. (2002). Instructional artifacts. In *Proceedings of the Conference on Computer Support for Collaborative Learning: Foundations for a CSCL Community*, pages 33–42. International Society of the Learning Sciences.

Swaak, J. and de Jong, T. (1996). Measuring intuitive knowledge in science: The development of the what-if test. *Studies in Educational Evaluation*, 22(4):341–362.

Taagepera, R. and Lee Ray, J. (1977). A generalized index of concentration. *Sociological Methods & Research*, 5(3):367–384.

Tchounikine, P. (2008). Operationalizing macro-scripts in cscl technological settings. *International Journal of Computer-Supported Collaborative Learning*, 3(2):193–233.

Terwel, J., van Oers, B., van Dijk, I., and van den Eeden, P. (2009). Are representations to be provided or generated in primary mathematics education? Effects on transfer. *Educational research and Evaluation*, 15(1):25–44.

Thelen, H. A. (1954). Dynamics of groups at work.

Tobin, K. (1990). Research on science laboratory activities: In pursuit of better questions and answers to improve learning. *School science and Mathematics*, 90(5):403–418.

Trimbur, J. (1989). Consensus and difference in collaborative learning. *College English*, 51(6):602–616.

Tsui, A. S., Egan, T. D., and O'Reilly III, C. A. (1992). Being different: Relational demography and organizational attachment. *Administrative science quarterly*, pages 549–579.

Tuan, H.-L., Chin, C.-C., and Shieh, S.-H. (2005). The development of a questionnaire to measure students' motivation towards science learning. *International Journal of Science Education*, 27(6):639–654.

Tudge, J. (1992). Vygotsky, the zone of proximal development, and peer collaboration: Implications for classroom practice.

Turian, J. (2013). Using AlchemyAPI for Enterprise-Grade Text Analysis. Technical report, Technical report, AlchemyAPI (August 2013).

Turnitsa, C. (2005). Extending the Levels of Conceptual Interoperability Model. IEEE CS Press.

Tuzhilin, A. and Adomavicius, G. (2005). Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. *IEEE Transactions on Knowledge & Data Engineering*, 17(6):734–749.

van der Aalst, W. (2011). *Process mining: discovery, conformance and enhancement of business processes.* Springer Science & Business Media.

van der Meij, J. and de Jong, T. (2006). Supporting students' learning with multiple representations in a dynamic simulation-based learning environment. *Learning and instruction*, 16(3):199–212.

van der Vegt, G. S. and Bunderson, J. S. (2005). Learning and performance in multidisciplinary teams: The importance of collective team identification. *Academy of management Journal*, 48(3):532–547.

van Joolingen, W. R. (1998). Cognitive tools for discovery learning. *International Journal Of Artificial Intelligence In Education (IJAIED)*, 10:385–397.

van Joolingen, W. R. and de Jong, T. (1993). Exploring a domain with a computer simulation: Traversing variable and relation space with the help of a hypothesis scratchpad. In *Simulation-based experiential learning*, pages 191–206. Springer.

van Joolingen, W. R. and de Jong, T. (2003). Simquest. In *Authoring tools for advanced technology learning environments*, pages 1–31. Springer.

van Joolingen, W. R., de Jong, T., and Dimitrakopoulou, A. (2007). Issues in computer supported inquiry learning in science. *Journal of Computer Assisted Learning*, 23(2):111–119.

van Joolingen, W. R., de Jong, T., Lazonder, A. W., Savelsbergh, E. R., and Manlove, S. (2005). Co-Lab: research and development of an online learning environment for collaborative scientific discovery learning. *Computers in human behavior*, 21(4):671–688.

van Joolingen, W. R. and Zacharia, Z. C. (2009). Developments in inquiry learning. In *Technology- enhanced learning*, pages 21–37. Springer.

Veermans, K. (2002). Intelligent support for discovery learning.

Veermans, K., van Joolingen, W. R., and de Jong, T. (2006). Use of heuristics to facilitate scientific discovery learning in a simulation learning environment in a physics domain. *International Journal of Science Education*, 28(4):341–361.

Verbert, K., Manouselis, N., Drachsler, H., and Duval, E. (2012a). Dataset-driven research to support learning and knowledge analytics. *Journal of Educational Technology & Society*, 15(3):133–148.

Verbert, K., Manouselis, N., Ochoa, X., Wolpers, M., Drachsler, H., Bosnic, I., and Duval, E. (2012b). Context-aware recommender systems for learning: a survey and future challenges. *IEEE Transactions on Learning Technologies*, 5(4):318–335.

Vogel, O., Arnold, I., Chughtai, A., and Kehrer, T. (2011). *Software Architecture - A Comprehensive Framework and Guide for Practitioners*. Springer.

Vozniuk, A., Govaerts, S., Bollen, L., Manske, S., Hecking, T., and Gillet, D. (2014). AngeLA: putting the teacher in control of student privacy in the online classroom. In *Information Technology Based Higher Education and Training (ITHET), 2014*, pages 1–4. IEEE.

Vozniuk, A., Govaerts, S., and Gillet, D. (2013). Towards portable learning analytics dashboards. In *Proceedings of the 13th IEEE International Conference on Advanced Learning Technologies*, Beijing, China.

Vygotsky, L. (1987). Zone of proximal development. *Mind in society: The development of higher psychological processes*, 5291:157.

Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*. Harvard university press.

Waight, N. and Abd-El-Khalick, F. (2011). From scientific practice to high school science classrooms: Transfer of scientific technologies and realizations of authentic inquiry. *Journal of Research in Science Teaching*, 48(1):37–70.

Wasserman, S. (1994). *Social network analysis: Methods and applications*, volume 8. Cambridge university press.

Webb, N. M., Nemer, K. M., and Zuniga, S. (2002). Short circuits or superconductors? Effects of group composition on high-achieving students' science assessment performance. *American Educational Research Journal*, 39(4):943–989.

Weinberger, A. and Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers & education*, 46(1):71–95.

Weinberger, A., Reiserer, M., Ertl, B., Fischer, F., and Mandl, H. (2005). Facilitating Collaborative Knowledge Construction in Computer-Mediated Learning Environments with Cooperation Scripts. In Bromme, R., Hesse, F. W., and Spada, H., editors, *Barriers and Biases in Computer-Mediated Knowledge Communication: And How They May Be Overcome*, Computer-Supported Collaborative Learning Series, pages 15–37. Springer US, Boston, MA.

Weinbrenner, S. (2012). *SQLSpaces: a platform for flexible language heterogeneous multi-agent systems.* PhD Thesis, Hut, München.

Weinbrenner, S., Engler, J., Bollen, L., and Hoppe, H. U. (2009). Ontologybased support for designing inquiry learning scenarios. In *International Workshop on Ontologies and Semantic Web for E-Learning (SWEL), Brighton, UK.*

Weitzman, M. L. (1992). On Diversity. *The Quarterly Journal of Economics*, 107(2):363–405.

Wertsch, J. V. (1984). The zone of proximal development: Some conceptual issues. *New Directions for Child and Adolescent Development*, 1984(23):7–18.

Wheeler, S., Yeomans, P., and Wheeler, D. (2008). The good, the bad and the wiki: Evaluating student-generated content for collaborative learning. *British journal of educational technology*, 39(6):987–995.

White, B. Y. (1993). Thinkertools: Causal models, conceptual change, and science education. *Cognition and instruction*, 10(1):1–100.

White, B. Y. and Frederiksen, J. R. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and instruction*, 16(1):3–118.

White, M. J. (1986). Segregation and diversity measures in population distribution. *Population index*, pages 198–221.

Whitelock, D., Field, D., Pulman, S., Richardson, J. T., and van Labeke, N. (2013). OpenEssayist: An automated feedback system that supports university students as they write summative essays. In *The 1st International Conference on Open Learning: Role, Challenges and Aspirations.*

Whitelock, D., Twiner, A., Richardson, J. T., Field, D., and Pulman, S. (2015). OpenEssayist: a supply and demand learning analytics tool for drafting academic essays. In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*, pages 208–212. ACM.

Wichmann, A., Engler, J., and Hoppe, H. U. (2010). Sharing educational scenario designs in practitioner communities. In *Proceedings of the 9th International Conference of the Learning Sciences-Volume 1*, pages 750–757. International Society of the Learning Sciences.

Wichmann, A. and Leutner, D. (2009). Inquiry Learning: Multilevel Support with Respect to Inquiry, Explanations and Regulation During an Inquiry Cycle. *Zeitschrift für Pädagogische Psychologie*, 23(2):117–127.

Wilkinson, J. W. and Ward, M. (1997). The Purpose and Perceived Effectiveness of Laboratory Work in Secondary Schools. *Australian Science Teachers Journal*, 43(2):49–55.

Williams, K. Y. and O'Reilly III, C. A. (1998). Demography and. *Research in organizational behavior*, 20:77–140.

Wilson, B. G. (1996). *Constructivist learning environments: Case studies in instructional design*. Educational Technology.

Windschitl, M., Thompson, J., and Braaten, M. (2008). Beyond the scientific method: Model-based inquiry as a new paradigm of preference for school science investigations. *Science education*, 92(5):941–967.

Wise, A. F. (2014). Designing pedagogical interventions to support student use of learning analytics. In *Proceedings of LAK*, pages 203–211. ACM.

Wise, A. F. and Schwarz, B. B. (2017). Visions of cscl: Eight provocations for the future of the field. *International Journal of Computer-Supported Collaborative Learning*, 12(4):423–467.

Wise, K. C. and Okey, J. R. (1983). A meta-analysis of the effects of various science teaching strategies on achievement. *Journal of Research in Science Teaching*, 20(5):419–435.

Wittrock, M., Clark, C., and Peterson, P. (1990). *Vol. 3: Students' thought processes*. New York [etc.]: MacMillan [etc.].

Xu, K. (2003). How has the literature on gini's index evolved in the past 80 years? *Dalhousie University, Economics Working Paper*.

Yoon, H.-G., Joung, Y. J., and Kim, M. (2012). The Challenges of Science Inquiry Teaching for Pre-Service Teachers in Elementary Classrooms: Difficulties on and under the Scene. *Research in Science Education*, 42(3):589–608.

Zacharia, Z. and Anderson, O. R. (2003). The effects of an interactive computer-based simulation prior to performing a laboratory inquiry-based experiment on students' conceptual understanding of physics. *American Journal of Physics*, 71(6):618–629.

Zacharia, Z. C. (2005). The impact of interactive computer simulations on the nature and quality of postgraduate science teachers' explanations in physics. *International Journal of Science Education*, 27(14):1741–1767.

Zacharia, Z. C. and Constantinou, C. P. (2008). Comparing the influence of physical and virtual manipulatives in the context of the physics by inquiry curriculum: The case of undergraduate students' conceptual understanding of heat and temperature. *American Journal of Physics*, 76(4):425–430.

Zacharia, Z. C., Manoli, C., Xenofontos, N., de Jong, T., Pedaste, M., van Riesen, S. A., Kamp, E. T., Mäeots, M., Siiman, L., and Tsourlidaki, E. (2015). Identifying potential types of guidance for supporting student inquiry when using virtual and remote labs in science: a literature review. *Educational Technology Research and Development*, 63(2):257–302.

Zacharia, Z. C. and Olympiou, G. (2011). Physical versus virtual manipulative experimentation in physics learning. *Learning and Instruction*, 21(3):317–331.

Zhang, J., Scardamalia, M., Reeve, R., and Messina, R. (2009). Designs for collective cognitive responsibility in knowledge-building communities. *The Journal of the Learning Sciences*, 18(1):7–44.

Ziegler, S. (1981). The effectiveness of cooperative learning teams for increasing cross-ethnic friendship: Additional evidence. *Human Organization*, 40(3):264–268.

Zutin, D., Auer, M., and Gustavsson, I. (2011). A VISIR lab server for the iLab Shared Architecture. In *Global Engineering Education Conference (EDUCON), 2011 IEEE*, pages 30–33.

Zutin, D., Auer, M., Maier, C., and Niederstatter, M. (2010). Lab2go #x2014; A repository to locate educational online laboratories. In *Education Engineering (EDUCON), 2010 IEEE*, pages 1741–1746.