CrossMark

Composition-Effects of Context-based Learning Opportunities on Students' Understanding of Energy

Sören Podschuweit[1] · **Sascha Bernholt**[1]

Abstract Context-based learning has become a widespread approach in science education. While positive motivational effects of such approaches have been well established empirically, clear results regarding cognitive aspects of students' learning are still missing. In this article, we argue that this circumstance might be mainly rooted in the definition of context itself. Based on this argument, we shift from the issue of *if* contexts are cognitively beneficial to focus on the question of *which composition of contexts* is, at least by tendency, more effective than another. Based on theories of conceptual change, we therefore conducted a small-scale intervention study comparing two groups of students learning in different sets of contexts focusing on the same scientific concept—the cross-cutting concept of energy. Results suggest that learning in a more heterogeneous set of contexts eases transfer to new contexts in comparison to learning in a more homogeneous set of contexts. However, a more abstract understanding of the energy concept does not seem to be fostered by either of these approaches. Theoretical as well as practical implications of these finding are discussed.

Keywords Conceptual change · Context-based learning · Energy · Transfer

Introduction

Although influencing everyday life more and more, natural sciences and mathematics still belong to the least popular subjects in school (Sjøberg and Schreiner 2010). As often discussed, the inherent complexity of scientific topics as well as low relevance felt by students might be two of the main reasons for this issue. Accordingly, several approaches were

✉ Sören Podschuweit
podschuweit@ipn.uni-kiel.de

Sascha Bernholt
bernholt@ipn.uni-kiel.de

[1] Department of Chemistry Education, IPN Leibniz-Institute for Science and Mathematics Education at Kiel University, Olshausenstraße 62, D-24118 Kiel, Germany

⚡ Springer

developed over the past two decades to make scientific topics more understandable, interesting and relevant for students. A very popular and widely implemented attempt in this regard is context-based learning. It is studied in the course of science education for about 35 years. Despite being a quite heterogeneous field, different approaches of context-based learning are unified by the core idea of putting scientific concepts, models or topics in some kind of frame connecting science to everyday life, societal issues, or technological innovations. In a broad definition stemming from linguistics, Gilbert, referring to Duranti and Goodwin (1992), describes a context as "a focal event embedded in its cultural setting" (2006, p. 960). From a teacher's perspective, the focal event might be a scientific process explainable by an important scientific concept or model. The cultural setting framing the event is then often assumed to increase the students' motivation to deal with the content, to raise interest and thus finally to ease the learning process. Criticism on context-based learning, however, comprises that the additional information included in such a cultural framing would hinder conceptual learning due to an information overload curtaining the scientific core ideas (c.f. Gräber 1995; Lye et al. 2001). They favour a more abstract way of learning intended to ease the recognition and understanding of core concepts (Goldstone and Sakamoto 2003; Kaminski et al. 2005).

Nearly as long as context-based learning has been implemented, studies investigate its effects to clarify the above argument. Bennett, Lubben and Hogarth analysed 17 suitable out of 2500 studies concerning context-based learning or science, technology, and society (STS) respectively published from 1980 to 2003 and found that 'context-based/STS approaches result in improvement in attitudes to science and that the understanding of scientific ideas developed is comparable to that of conventional approaches' (2007, p. 1). They also stress that many investigations have particular flaws which make it difficult to compare the reported results across studies. However, recent investigations following very high test standards, for example by Kölbach and Sumfleth (2013), seem to show similar conclusions. Accordingly, one can subsume that there is no clear evidence for a positive effect of context-based learning on students' conceptual understanding so far (c.f. Taasoobshirazi and Carr 2008).

In this contribution, we argue that the inconclusive findings for this effect might, at least to some degree, be grounded in a biased relation of the underlying definition of context and the expected learning results. Following Greeno (2009) and Finkelstein (2005), our key argument is that there is no such thing like a learning opportunity without context. Therefore, we intend to shift the discussion towards the question *which composition of contexts* might be more suitable to acquire knowledge or competence concerning a certain scientific concept in comparison to another context composition. Inspired by theories of conceptual change research, we therefore implemented a small-scale pre-post intervention study comparing two learning environments consisting of different contexts focusing on the same core concept— *energy*. Results indicate a close correspondence between the composition of different contexts and students' learning gains. Differences in test results depending on the learning environment will be discussed with regard to possible consequences for classroom practice.

Theoretical Background

As discussed before, *context* appears to be a quite ambiguous term. Within the following sections, we therefore give a short overview of its most important characteristics and distinctions to then derive our own working definition. On the basis of diSessa's *knowledge in pieces*

(kip) theory, we further discuss what kind of enrichment a conceptual change perspective on context-based learning might offer to finally deduce the setting of this study.

Contexts in Science Education—Definitions and Differentiations

From its linguistic origin—*contextere*—the term context means *to weave something together* or *relationship*. Thus, a context connects several (at least two) things: an implication important to keep in mind when referring to context in educational research where a context is sometimes understood as a simple surrounding of an educational object. However, it is the connection between both—surrounding and educational object—that determines the educational value of the context and, crucially, this connection is individually and dynamically constructed[1] while engaging in a learning environment (van Oers 1998).

Many different definitions for the term context exist in educational discourse emphasising different aspects with no consensus yet on a generally accepted definition (c.f. Bennett et al. 2004; Duranti and Goodwin 1992; van Oers 1998). Nevertheless, Duranti and Goodwin give some points of reference describing context as "a relationship between two orders of phenomena [a focal event and a field of action] that mutually inform each other to comprise a larger whole" (1992, p. 4). Furthermore, they identified four attributes an educational context should possess:

- "a setting, a social, spatial, and temporal framework within which mental encounters with focal events are situated;
- a behavioral environment of the encounters, the way that the task(s), related to the focal event, have been addressed, is used to frame the talk that then takes place;
- the use of specific language, as the talk associated with the focal event that takes place;
- a relationship to extra-situational background knowledge" (Gilbert 2006, p. 960, referring to Duranti and Goodwin 1992, pp. 6ff.)

Shifting this linguistically orientated description more into a conceptual direction, Gilbert (2006, pp. 966ff) distinguishes four models of *context* found in practice with different kinds of connections between surrounding and educational object (i.e. *concept*). The first, "context as the direct application of concepts", describes a post hoc, mono-directional application of a concept on some situation or experience assumed to be of importance for the students. The second model, "context as reciprocity between concepts and applications", more emphasises the mutual influence concept and context have on each other, especially on the level of the cognitive structure. Model 3, "Context as provided by personal mental activity", further includes the personal role of the student in between this connection of context and concept, by, for example, letting the student identify herself with the role of a researcher working on an actual or past scientific problem included in narratives. The last model distinguished by Gilbert is referred to as "context as the social circumstances". Here, a context is "situated as a cultural entity in a society" (Gilbert 2006, p. 970), meaning that primarily actual societal and cultural, not necessarily primarily scientific issues are the focal events of the lesson and teachers as well as students work on them together using scientific concepts and models in some kind of

---

[1] For the sake of terminological clarity, van Oers 1998 suggests distinguishing *context* from *contextualising*, whereby the former refers to an externally discernable amount of contextual elements while the latter describes a person's mental interpretation or perception of the former.

community of practice. Gilbert et al. (2011) only consider the latter model to sufficiently fulfil the four criteria of contextual learning proposed by Duranti and Goodwin (1992).

Coming from *transfer research*, Barnett and Ceci (2002) give a more systematic differentiation of distinct *kinds* of contexts, partially inherent in the descriptions of Gilbert and Duranti and Goodwin. Agreeing on the ability to *transfer* knowledge and skills from one situation in life to another as one of the most desirable *technical* goals of schooling, it is not only important to know what a context might be or what makes a learning opportunity more or less contextualised. On top of that, a differentiation *in between contexts* is needed to evaluate the quality of transfer achieved by a student. Barnett and Ceci (2002) propose such a scheme which is also broadly divided into parts of *surrounding* and *educational object*, denoted by *context* and *content*. As possible categories to differentiate between *contexts* (with the goal to provide answers to the question when and where knowledge is transferred from and to), they list "knowledge domain" (e.g. science vs. history), "physical context" (e.g. school vs. home), "temporal context" (e.g. recalling a month later), "functional context" (e.g. academic vs. informal questionnaire), "social context" (e.g. individual vs. large group) and "modality" (e.g. book learning vs. oral exam). The content part in turn consists of the "learned skill" (e.g. procedure or principle), "performance change" (e.g. speed or accuracy) and "memory demands" (e.g. execution only or recall, recognition and execution) (Barnett and Ceci 2002, p. 621). Within each category, the content categories are not necessarily regarded as ordinal whereas the context categories are. For example, referring to the knowledge domain category, the transfer of conceptual knowledge from physics to chemistry is not as far as the transfer from physics to the arts or to history. The same comparison seems hardly appropriate for a learning skill (Is the transfer of a skill harder than the transfer of a principle?). This issue of transfer distance in turn seems to be at the core of the current discussion on transfer, conceptual change or competence studies (Dori and Sasson 2013). Transfer and widespread applicability of scientific concepts is the ultimate goal of schooling but *what* and *where* to apply these concepts often seems subtle.

In contrast to the classifications mentioned above, Finkelstein (2005) describes *hierarchical layers* or *shells* educational contexts are always embedded in. At the core, he puts the *task formation*, consisting of *task (problem)*, *student* and *concept*, whereas all three of these parts are interconnected. This task formation is embedded in a *situation*, for example a *class activity* like *homework*. This situation in turn is framed by the *idioculture* of learning—for example a *pre-med physics course* probably constitutes a different learning context than a usual *university physics course*. Besides describing these different shells, Finkelstein especially emphasises the interconnectedness of all of them: "Student learning in physics is **always** intertwined with and constituted in context, and inherent in a given context are certain features that promote or inhibit construction of content understanding" (2005, p. 1195). However, no predication is made about the transferability of this content understanding or about the generalisability of promoting or hindering features across contexts (Fig. 1).

Further characterising these features that '*promote or inhibit understanding*', van Vorst et al. (2014) put emphasis on the affective components of context-based learning which of course are connected to the other characteristics mentioned above. It is often expected that context-based learning motivates students more than 'non-context-based learning' (whatever this might be, c.f. Greeno 2009) so that the contextualisation increases students' conceptual understanding, mediated by the increased motivation. According to Keller (1987), students' motivation in a learning situation is determined by interest, satisfaction, expectations and relevance. In learning contexts, van Vorst et al. (2014) expect these criteria to be determined by
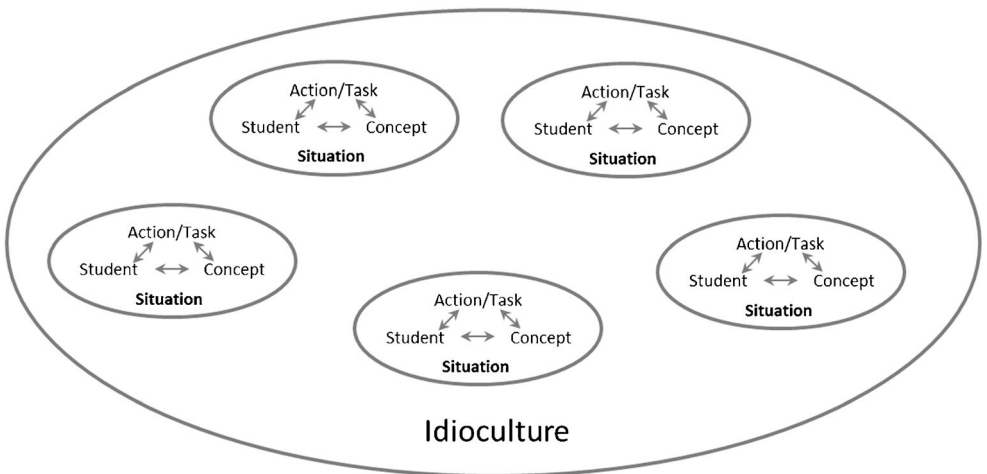
Fig. 1 Schematic depiction of task formations embedded in an idioculture (adapted from Finkelstein 2005)

the context's *authenticity* and its *prominence (ordinariness or peculiarity). Authenticity*, for example, results from the *complexity* and the *form of representation* of a situation and the credibility for the student. In consensus with Finkelstein's (2005) description of task formation, the authors heavily emphasise the affective constitution of contexts to be a joint result of task and student characteristics.

Based on this short review on theoretical aspects of the context term, it becomes obvious that—as Duranti and Goodwin (1992) already concluded 24 years ago—there is no clear definition. Albeit most of the presented approaches can be seen as trying to specify the relation between content (or more general: the focal event) and context (the field of action, c.f. Duranti and Goodwin 1992), they differ in the focus of this specification, ranging from the educational function (Gilbert 2006), the classification of different kinds of and the transfer between contexts (Barnett and Ceci 2002) or the inspection of promoting and hindering features of the context regarding students' learning of the content (Finkelstein 2005; van Vorst et al. (2014)). Consequently, the definition of context often depends on the research question at issue, making the reception and comparison of different research approaches difficult. As context-based learning of science meanwhile has a substantial history of about 35 years (c.f. Bennett and Lubben 2006), there is a plethora of studies investigating its effectiveness. However, despite the large number of studies, the results are not very coherent. This ambiguity might not seem surprising taking into account the ambiguity in context definitions depicted above. Within the following section, we intend to discuss some crucial issues between the poles of theoretical expectations and empirical outcomes in research on context-based learning.

Contexts in Science Education—Empirical Findings

Within the frame of probably the most comprehensive review of studies on context-based learning, Bennett et al. (2007) identified 61 out of about 2500 studies worth interpreting concerning the effectiveness of context or STS (science, technology, and society, c.f. Aikenhead 1994) approaches. Seventeen out of these 61 were judged to be based on an appropriate experimental design. Two of these reviewed studies further allowed calculating

effect sizes honouring context-based or STS approaches as superior concerning science understanding. Thus, there is an urgent need of methodologically more rigorous research in this field in the first place. Yager and Weld (1999), evaluating the so called *Iowa SS&C project* (scope, sequence and coordination) on a large scale, found the STS-orientated approach to be superior in comparison to textbook-orientated lessons in terms of teachers confidence and constructivist attitudes as well as student achievement ($d = 1.52$; as reported in Bennett et al. 2007). The intervention programme lasted about 4 years. Restrictively, it has to be mentioned that in part different test instruments were used for the two comparison groups and also different teachers taught the classes. The second study was conducted by Winther and Volk (1994) in the field of chemistry education. Within the frame of this nearly yearlong intervention, also a collaborative learning and STS-orientated course has been compared to a 'traditional' textbook-orientated teaching approach. They found the experimental group to be superior concerning chemistry understanding ($d = 0.63$; as reported in Bennett et al. 2007). Comparable to the first study, some restrictions have to be mentioned: different teachers taught the groups, important control variables like cognitive abilities have not been assessed and also different teaching styles confounded the design (the experimental group experienced a more group-orientated and collaborative teaching style). Additionally, both studies mention having predominantly investigated STS courses and not specifically the effectiveness of context-based learning, an important difference.

Extending this review of studies on the effectiveness of context-based learning, Taasoobshirazi and Carr (2008) by contrast focused more on the idea of context as depicted above. Departing from a definition of context-based science education as "using concepts and process skills in real-world contexts that are relevant to students from diverse backgrounds" (ibid. p. 157 citing Glynn and Koballa 2005, p. 75), they reviewed the affective and cognitive efficacy of contexts in assessment as well as learning situations. With regard to contexts used in learning situations, as it is the focus of the present study, they identified only two studies being relevant to some extent. Wierstra and Wubbels (1994) compared two groups of students taught either in a contextualised or a traditional manner over a period of 4 weeks. Results on a traditional physics post-test did not differ significantly, whereby the absence of a pre-test restricts the interpretation of this finding. Using pre-post measurements, Murphy et al. (2006) found that students learning about radioactivity in a context-based setting outperform students learning traditionally. Unfortunately, the length of the intervention has not been reported in this case. Moreover, pertaining as well some of the studies reviewed by Bennett et al. (2007), context-based learning and collaborative group work were often confounded variables in the different studies. Generally, Taasoobshirazi and Carr (2008) note major limitations of studies investigating the effect of context-based learning and demand more precise pre-post interventions with randomised assignment of participants to exclude confounded independent variables.

Fechner (2009) identifies similar flaws in previous studies and calls for well-controlled experimental designs, since field studies as the ones cited above are faced with the problem of too many variables possibly influencing the outcome measures. In addition to the evidence presented above, she mentions a study by Glemnitz (2007), indicating that students attending *ChiK-classes* (Chemistry in Context, cf. Parchmann et al. 2006) reach higher levels of conceptual understanding and are better able to interlink knowledge than students taught traditionally. However, since this study also did not fit the study criteria required so far, Fechner implemented an experimental pre-post-follow-up study comparing small groups of students working either in *real-life* or *laboratory* contexts for one school lesson (45 min) per

day on five subsequent days. Nearly randomised conditions had been assured by a randomised block design. Results indicate that students who learned in real-life contexts outperform the laboratory group in applying the acquired knowledge to formerly unknown real-life situations. Vice versa, both groups performed equally well in applying knowledge to formerly unknown laboratory situations. Generally, Fechner (2009) assigned small to medium-sized effects to context-based learning on student achievement measures (knowledge recall: $\eta^2 = 0.02$; knowledge application in real-life contexts: $\eta^2 = 0.07$; knowledge application in laboratory contexts: no significant effect; connected knowledge: $\eta^2 = 0.08$), results worth noticing when also taking into consideration the relatively short time of the intervention. Related studies by van Vorst or Kölbach and Sumfleth tried to clarify and confirm these results. van Vorst (2013), analysing affective features of contextualised tasks, showed that for ninth graders it is not the up-to-datedness of a context that raises the emotional component of the students' interest but its exceptionality. This result emphasises the importance of paying particular attention to the question which contexts are compared to each other in empirical studies, since attributes like *relevant*, *interesting* or *motivating* for students might have been assigned to some context quite non-reflectively in former investigations. Kölbach and Sumfleth (2013) in turn implemented a research design quite similar to Fechner (2009). Balanced by prior knowledge, students were assigned to groups of either *real-life* (bathing lake) or *non-real-life* (student laboratory) contexts. The intervention took about 60–90 min on two subsequent days. Both groups learned with exemplary solutions. Pre-post-follow-up tests revealed significantly higher gains concerning situational interest for the *real-life* context group but no differences regarding knowledge gains.

Summing up the foregoing short review of empirical findings, some points seem to have significance. Firstly, and no one will be surprised so far, there is no clear evidence of a general positive effect of context-based learning on students' learning. Positive motivational effects in contrast seem to be robust. Searching for explanations for these findings, it seems like there is a relation between the duration of the intervention and its effect on the learning outcome. Comparing the few studies with acceptable design standards, effect sizes seem to increase with the duration of the intervention. The investigation of Yager and Weld (1999) lasted about 4 years and revealed an effect size of $d = 1.52$ (strong effect; Cohen 1988); Winther and Volk's (1994) nearly yearlong study showed a medium-sized effect of $d = 0.63$; students in Fechner's study experienced context-based learning in an experimental setting on five subsequent days and revealed small to medium-sized effects regarding achievement ($\eta^2 = 0.02$ to 0.08), whereas Kölbach's 2-day intervention showed no significant results. A possible reason for this time-dependence might be that the impact of context-based learning on students' conceptual understanding, mediated by students increasing interest and/or motivation, only comes into effect with larger time scales. This assumption is supported by the fact that most of the cited studies primarily expect this mediation effect, but no further direct cognitive mechanism is explicated.

However, when identical cognitive elements are presented to the students in the different groups of comparison (e.g. real-life contexts vs. laboratory contexts) and when some score of conceptual understanding (i.e. results of an achievement test) is the measure of students' (cognitive) learning gain, we should maybe not expect too much difference between both groups (assuming that motivation-driven benefits fail to appear on short time scales). In fact, keeping cognitive elements constant across both groups of comparison is often intended in studies on context-based learning, e.g. "both [contexts] rely on the same underlying conceptual structure in order to keep the required amount of content knowledge constant" (Fechner 2009,

p. 43). This might explain the limited (or missing) effectivity of context-based interventions in comparison to *non-contextualised* control groups especially found in short-time interventions (c.f. Fechner 2009; Kölbach and Sumfleth 2013). The expectation of differences in learning gains might be especially questionable since, like in case of Kölbach and Sumfleth (2013), it is arguable if a laboratory can serve as a *non-contextualised* control group for comparison, as the laboratory may be the most authentic environment to learn scientific content from the students' point of view.

In summary, cognitive as well as motivational factors should be taken into consideration when planning studies concerning the issue of context-based learning. In addition (and as discussed in the previous section), the definition of context depends on the problem to be studied. Accordingly and with regard to the study to be presented in this paper—a small-scale pre-post intervention study comparing the effectiveness of two learning environments consisting of different contexts—our own definition of context is elaborated within the following paragraphs.

Reflecting Duranti and Goodwin (1992), we agree on the broad characterisation of context as an individually constructed focal event embedded in and interconnected with a field of action. Referencing to Finkelstein (2005), we restrict ourselves to the task formation (consisting of task, student and concept) since these are the main *objects* in school settings. We are aware of the influence social, local, temporal or modal contexts can have on the learning process in general, but these are not investigated further here. Regarding learning in school in the first place, broad applicability of conceptual knowledge, predominantly assessed in classroom talk or written exams is at issue and we therefore focus on this part of the learning environment. This means we aim to analyse contextual features of a learning or assessment situation in which the concept to be learned is embedded. When focusing on one concept, like on the energy concept in this study, there are two context-specific degrees of freedom left to describe within the task formation: the student on the one hand and the task on the other hand (Finkelstein 2005). As the student is the *dependent variable* of education, it is the task we can intentionally manipulate to hopefully ease and foster conceptual understanding. Besides modality, complexity or openness of a task, it is the frame of reference the concept at issue is connected to and which is suspected to hinder or foster students' learning. Mostly, these frames of reference are supposed to connect to the students' everyday life experience or to arouse interest by referencing to technological advancements or societal issues. This in turn is close to Duit's understanding of *context*: "Always, a subject-specific content can only be learned in a context relevant for the student. That's why the context should be chosen in a way that it is 'meaningful' (Muckenfuß 1995) for the learner. Such contexts can be topics from the students' everyday life, like every day and natural phenomena, or technical applications, but also aspects of the meaningfulness of physics for technology and society." (2010, p. 1).

In other words, a context can be seen as a topic (or frame of reference) a content (or concept) is embedded in. Accordingly, for our study, we define context as *the topically determined and individually constructed meaning- (and acting-) frame of reference, within which an engagement with a scientific concept occurs*. For instance, wind turbines are seen as a scientific context which can be addressed with different conceptual foci (energetic aspects, mechanics, material properties, as well as their interplay), providing the opportunity to engage in theoretical and experimental investigations, and allowing students to engage in personal, societal and/or technical perspectives with regard to a topic that has a high (in this case societal) relevance (cf. Gilbert 2006; Gilbert et al. 2011). Hence, contexts are defined in close correspondence to specific topics, in contrast to rather specific situations (e.g. Thaden-Koch

et al. 2006, consider a single- and a two-lane marble run as two distinct contexts) or larger thematic areas (e.g. fuels or cleaning detergents; Parchmann et al. 2006). When investigating the effect of context-based learning on students' conceptual learning, a definition of *context* is one side of the problem, specifying ones' understanding of *concept* the other.

Concepts and Conceptual Change Research

Within science education research, the issue of students' (un)successful development of domain-specific concepts is widely addressed under the label of conceptual change. Within this debate, Brown (2014) identifies three major, apparently controversial, commitments concerning how student conceptualizations are characterised: as *misconceptions* (e.g. Kruger 1990), as *coherent systems of intuitive ideas* (Vosniadou and Brewer 1992), and as *intuitive fragments* (diSessa 1993). In the first view, misconceptions are regarded as unitary ideas at odds with the accepted scientific norm that therefore needed to be replaced. The coherent systems view takes a more detailed look and identifies students' conceptions (be they either correct or false) to consist of several elements like presuppositions, beliefs and mental models which coherently interact in a system (Vosniadou et al. 2008). According to the intuitive fragments perspective, conceptions are also regarded as consisting of various elements, whereas the elements are supposed to be of smaller grain size compared to the ones regarded in the coherent systems view (diSessa 2008). Moreover, proponents of the intuitive fragments perspective suppose these elements to generally act together less coherently. Whereas it seems to become consensus within the science education community that students' conceptions are not solely unitary and simply replaceable, especially the two latter approaches to conceptual change are still discussed controversially. This controversy, in a nutshell, derives from conflicting findings concerning the issue of *context sensitivity of learning* (Özdemir and Clark 2007, p. 570). On the one hand, there is evidence that especially students' alternative conceptions are resistant to change and appear quite coherently across different contexts. This finding is well explainable in terms of the coherent systems view, as coherent conceptual structures should be hard to modify and not be affected by contextual factors too heavily. On the other hand, there is opposing evidence that students' conceptions are affected by contextual factors, which supports the intuitive fragments proponents, as they claim that the relatively loosely connected elements might respond quite differently to contextual features.

In general, diSessa and Sherin (1998) identify an inexplicitness in the definition of concept and its constituents underlying many studies in science education research and psychology. They point out that "researchers have uncritically applied the term concept to what may be very different entities - dog, force, or number" (ibid, p. 1188). Even though they do not give a final definition of what a concept is, they introduce a *kind of* concept, supposed to be of explanatory value concerning students reasoning in science, more specifically in reasoning about force. It is called *coordination class*, "a particular kind of concept whose principal function is to allow people to read a particular class of information out of situations in the world" (ibid, p. 121). In contrast to other more or less explicit conceptions of concepts and also with regard to the other conceptual change approaches mentioned above, the coordination class approach towards conceptual learning pays particular attention to the influence of the learning context on the learners' conception.

Generally, diSessa assumes conceptual knowledge to consist of atomistic elements of a grain size smaller than concepts or theories (diSessa 1993). Accordingly, his approach is often referred to as *knowledge in pieces (kip)*. Within this approach, much emphasis is put on the

role of naïve conceptions students are expected to possess without having a comprehensive overarching framework connecting them in the first place (diSessa and Wagner 2005). These naïve conceptions appear in form of plenty of simple explanatory ideas, mainly stemming from everyday experience—so called *phenomenological primitives (p-prims)*. They are assumed to be evident in themselves and therefore would need no further explanation. As they are not connected in a fixed conceptual system, p-prims cause *wrong* scientific explanations by being applied in the wrong situation. For example, *Ohm's p-prim*[2] is valuable in explaining a lot of everyday processes (e.g. pushing objects) but fails to explain the increased pitch of a vacuum cleaner with a covered nozzle (c.f. diSessa 1993). If a particular p-prim is activated or not depends on individual as well as contextual factors. Some p-prims are intrinsically more likely to be articulated then others, for example because of a high accountability in everyday life (cf. Ohm's p-prim). Others in turn might be cued or suppressed by contextual factors (for example, students tend to see *springiness*[3] in the context of a book placed on a spring but not in the case of a book placed on a table; cf. Kapon and diSessa 2012).

In later stages of learning, p-prims might get organised in the structure of a coordination class. This concept's primary function, as stated in the definition above, is to gather information in certain situations. The range of contexts in which certain information is determined is called *span*, a performance specification of a coordination class. Furthermore, the specification of *alignment* describes if the *same* information is reliably gathered within a given span of contexts. Often inferences based on observations have to be made as certain information is not always directly accessible in a given context. diSessa and Sherin (1998) introduce the terms *readout strategies* and *causal net* to describe this process of inferring certain information. Readout strategies in this terminology are the ways observable information is gathered and the causal net is the set of inferences necessary to determine the (not directly observable) information at issue. Especially in early stages of learning, the set of p-prims a student possesses is supposed to be the main constituent of the causal net (ibid.). diSessa and Wagner (2005) also point out that multiple coordination classes might be involved in processes of information gathering. An elaborated conceptual understanding then can be achieved by gradually restructuring the organisation and the cuing priority of p-prims and other causal net elements or by gathering new readout strategies.

As abbreviated, the approach presented here is just a small slice of the general discussion on what a concept actually is and how it might change in the course of learning. The reason to choose the *kip* approach, or more specifically coordination class theory, for investigating the context specificity in learning about energy is that it, at least in our understanding, more directly addresses the issue of context-specific learning difficulties. It depicts cognition at a grain-size appropriate to understand the complex interplay of individual and contextual influences of the learning process. Moreover, other science educators support the general approach of understanding cognition as dynamic interaction of small cognitive elements (Brown and Hammer 2008; Koponen and Huttunen 2013). Dawson (2014) argues that learners conceptions are best described as neuronal *coexisting* elements, representing both common-sense and normative scientific views, both being optionally activated in dependence of contextual factors. In accordance with diSessa and Wagner (2005), he concludes that "instead

---

[2] "The major function of this element is to provide for activation (Systematicity A, mutual use) of a set of qualitative relationships among differentials in the effort of the agent (amount of impetus), the resistance, and the result: more effort implies more result; more resistance implies less result; and so on" (diSessa 1993, p. 126)

[3] "Objects give under stressing force. The amount of give is proportional to force" (diSessa 1993, p. 221)

of focusing on conceptual replacement, science educators need to aim more actively at strengthening the learner's executive processes which select contextually appropriate responses and inhibit inappropriate ones" (Dawson 2014, p. 389).

Students' Understanding of the Energy Concept

The following section depicts the main characteristics of the scientific concept energy, as we will apply the discussed theoretical considerations to this concept in the course of the current study. Energy is widely seen as a cross-cutting core concept with enormous political, scientific, societal and practical meaning. It provides a powerful tool to model, analyse or predict phenomena in all science disciplines as it plays a central role in everyday life situations. Thus, energy is not only a core concept within each science discipline (Driver and Millar 1986), but also a concept cutting across the science disciplines (Chen et al. 2014).

One reason why energy is such an important concept is that it is a conserved quantity (Feynman 1970). Being basically a mathematical principle when saying that there is a numerical quantity which does not change when something happens (ibid.), different attempts have been made to describe characteristics of the energy concept in order to make it less abstract. Duit (1986) for instance, proposed to differentiate forms, transfer/transformation, degradation and conservation. This differentiation as well as the general consensus of seeing energy as one of the major scientific concepts have influenced both research as well as policy documents, although to a different extend and with different emphasis (AAAS 2007; Lee and Liu 2010; National Research Council 2011; KMK 2004).

Next to theoretical considerations, numerous studies have tried to shed light on students' understanding of energy. Students' conceptions about energy have been investigated extensively within the focus of conceptual change research. Watts (1983) provided an initial indication of the richness of students' pre-instructional conceptions of energy. Interviewing students based on pictured situations related to energy, Watts identified six different conceptions: (a) energy as a causal agent, (b) energy as an ingredient, (c) energy as (obvious) activity or movement, (d) energy as the output or by-product of some process, (e) energy as a generalised fuel and (f) a flow-transfer model of energy (ibid., p. 216). He also pointed out that students in his study often made use of different frameworks across situations, indicating that students' understanding cannot be matched to a single conceptual framework. Although Watts' results had a major influence on subsequent studies, Harrer et al. (2013) reanalysis of his interview data challenges the unambiguity of interpreting students' verbal reports with regard to students' underlying conceptual framework.

Multiple studies extended this line of research, while also attempting to further systematise students' conceptions (Bliss and Ogborn 1985; Nicholls and Ogborn 1993; Solomon 1983, 1985). Different approaches have also been used to shed more light on the role of language in students' application of the energy concept in diverse situations. Amin (2009) tried to identify *metaphorical construals* in order to conceptualise the shift from a concrete, naïve to an abstract, scientific understanding of energy. He concludes that "developing an understanding of an abstract concept may rely extensively on metaphorical projection from experiential knowledge gestalts and that these projections are invited by verbal metaphors pervasive in both every day and scientific language" (ibid., p. 189), stressing the importance of metaphors to structure conceptual understanding, but also the difficulty that arises when the interpretation of a metaphor differs between everyday life and scientific contexts. Lancor (2014) qualitatively analysed student-generated analogies using the metaphor theory to gain understanding of how

students conceptualise energy in different contexts. She subsumes the results in six different conceptual metaphors (Energy as a substance that can be accounted for; Energy as a substance that can change forms; Energy as a substance that can flow; Energy as substance that can be carried; Energy as a substance that can be lost; Energy as a substance that can be added, produced, or stored; Lancor 2014, p. 16), indicating a large overlap to the results of Watts (1983) and others and again giving "insights into what sorts of metaphors make sense to students and how students connect science with their everyday experiences" (Lancor 2014, p. 16).

In addition to these qualitative approaches, students' understanding of energy has also been the object of quantitative empirical studies. Liu and McKeough (2005) reanalysed the TIMSS 1995 items and used the results to postulate five conceptual levels along which students' understanding of energy seems to develop: activity and work, forms and sources, transition and transformation, dissipation and conservation (see also Lee and Liu 2010). Neumann et al. (2013) extended this line of research. Based on newly developed items, they tried to assess students' understanding of energy in a cross-sectional study in grades 6, 8 and 10. "Findings provided evidence that students from Grade 6 mostly obtain an understanding of energy forms and energy sources. Students of Grade 8 additionally demonstrate an understanding of energy transfer and transformation, whereas only students of Grade 10, and then only some of these students, achieve a deeper understanding of energy conservation" (Neumann et al. 2013, p. 1). The authors further conclude that they "were able to confirm a general progression with respect to the levels described by four conceptions of energy (forms and sources, transfer and transformation, dissipation, conservation), but [they] could not confirm that these conceptions create distinct levels" (ibid., p. 23).

The differentiation of four to five different aspects of the energy concept in these quantitative studies resembles the theoretical considerations depicted above (cf. Duit 1981b, 1987). In addition, each of these aspects of the energy concept had been the subject of empirical investigations (Boyes and Stanisstreet 1990; Driver and Warrington 1985; Solomon 1985).

Generally, students learn about energy much earlier than in middle school physics instruction as they are already confronted with energy in everyday life contexts. Based on everyday life experiences and instruction in elementary school, students develop intuitive conceptions early on, which are reflected in the qualitative results depicted above. However, little is known about individual learning paths in middle school and beyond. Based on the frameworks described by Watts (1983), Trumper (1990, 1993, 1998) conducted a series of studies spanning grades 5 to 11 as well as student teachers at university. In a cross-sectional study from grades 9 to 11, two conceptualisations (*energy as a causal agent* and *energy as the output or by-product*) were particularly frequent (Trumper 1990), a result which was also found for the grades 5 to 9 (Trumper 1993). On tertiary level, Trumper (1998) analysed results of 25 student teachers in Israel over 4 years with a written energy test. He concludes that most alternative conceptions remain stable over time, that the dominant conceptions are *energy as ingredient* and *flow-transfer models* and that most students refuse the concept of energy degradation.

Beyond the analysis of students' understanding of energy, little is known about the influence of contexts and/or domains on students' conceptions. Analysing students' metaphors of energy, Lancor summarises that "there are differences in how energy is conceptualised in different scientific contexts; each system examined here favoured a different set of metaphors, and a different set of characteristics of energy was highlighted or obscured. However, there is also a surprising amount of overlap between systems that may seem very different on the surface (such as electrical circuits and ecosystems)"

(2014, p. 18). In addition, he concludes that each of these different conceptualisations of energy is valid, depending on the particular context, as neither students nor textbooks or experts can define energy without the use of metaphors. With regard to teaching the energy concept, different approaches exist which also make use of a particular terminology. For instance, some authors advocate a transfer-based approach (Brewe 2011; NRC 2011; Swackhamer 2005), involving the consideration of systems and/or fields, while other approaches focus on energy transformation, providing students with specific indicators to identify different energy forms and transformations between these forms (Nordine et al. 2011; Papadouris and Constantinou 2016). Both approaches received considerable criticism. For example, introducing students to the energy concept through the idea of energy forms may be a hindrance for future learning about energy (Kaper and Goedhart 2002), may nurture the idea of energy as a quasi-material substance (Warren 1982) or may teach students labels instead of a deeper understanding (Swackhamer 2005).

Comparing the qualitative and quantitative categorisations depicted above, it becomes obvious that the different approaches focus on differently detailed levels of the energy concept, ranging from metaphors, analogies or explanatory frameworks used by students to describe a specific phenomenon at hand (cf. Lancor 2014; Watts 1983), to broader knowledge areas addressing a specific general principle of energy (cf. Liu and McKeough 2005; Neumann et al. 2013). These different categorisations thus do not necessarily imply a specific explanatory knowledge on the students' side (e.g. recognising which energy form transforms into another energy form in a specific situation vs. knowing why and how this transformation of one energy form into another occurs). Trying to tie in with the conceptual change discourse depicted above, the more coarse-grained categories (e.g. forms, transformation, degradation, conservation; cf. Liu and McKeough 2005) seem to be closer related to a *coherence* perspective of conceptual change, as it is assumed that students who have understood for instance the idea of energy degradation can apply this principle in different contexts and situations. Accordingly, Neumann et al. (2013) "suggest that the description of how students develop an understanding of individual conceptions should build on the idea that mastering a particular level of understanding the energy concept (e.g. 'energy forms and sources') relates to students being able to describe scenarios in a greater variety of contexts (e.g. identify energy forms in more contexts)" (p. 22). This in turns leads to the conclusion that "we should be able to observe a smaller impact of item context on item difficulty for the more able students (i.e. students from higher grades)" (ibid., p. 23). Conversely, the more fine-grained categorisations stemming from the qualitative analyses seem to be more aligned with a *fragmentation* view on conceptual change, as the described conceptions and metaphors used by students are closely related to the context or phenomenon at hand (cf. Lancor 2014).

In summary, the concept of energy reflects the different characteristics of a concept, as stated in the definition above. With regard to specific, energy-related phenomena, the energy concept provides a specific conceptual framework to interpret, understand and predict these phenomena from a scientific perspective. In addition, numerous studies revealed that students use different conceptualisations to make sense of these phenomena. On a more comprehensive level, researchers have proposed different models that intend to structure the energy concept in smaller, specific aspects, with the aim to mediate between students' conceptualisations and the scientifically accepted concept. As it also relates to individual and social questions in current developments, the energy concept thus provides a meaningful knowledge domain for the current study.

Synthesis

In comparison to more traditional content-oriented teaching approaches, multiple empirical studies have shown that explicit context-based learning does not seem to hinder (even though not necessarily to foster) the understanding of scientific concepts embedded in contexts but increases interest and motivation of students. Accordingly, there might not be too much need to further discuss the value of such an approach in general, i.e. it might not be productive to ask *if* context-based or non-context-based learning is more effective. However, the actual influence of contextualised learning environments on students' conceptual understanding is highly complex and dynamic and different, partly controversial theories can be found in the literature trying to explain this influence. Aside from a strong emphasis on the impact of contextualised learning environments on students' interest and motivation, there is no consensus regarding the impact of the contextual features of a learning setting on students' cognition.

Regarding this aspect, the coordination class approach (diSessa 1993) pays particular attention to the influence of the learning context on the learners' conception. A students' conceptualisation of a scientific concept can be regarded as extensive when he or she can apply it correctly in many situations or contexts. From a *coordination class* perspective, it seems to be beneficial to teach specific aspects of a concept explicitly in different contextualised situations since this enables students to comprehend the situational appropriateness or inappropriateness of certain causal net elements or readout strategies (cf. van Oers 1998). In addition, it seems questionable if a learning opportunity without a context actually exists (Finkelstein 2005; Greeno 2009). Consequently, the question might not be *if*, but *which* contexts are more effective than others. Or, taking into consideration the structure learning is usually organised at school, *which* composition of contexts might be more fruitful than another.

Based on the coordination class approach, as sketched above, tackling a certain scientific concept like energy in a variety of contexts is expected to broaden the span of the learners' coordination class of energy. However, educators should pay attention to the conceptual offer a certain context implies. DiSessa and Wagner claim that "a new context that does not require any new causal net elements or new readout strategies is not helpful, [but a context that] provokes use of a new intuitive idea that extends the range of contexts that are seen as relevant […] will likely be productive" (2005, p. 149). At this point, it becomes obvious that it is necessary to tackle the issue of conceptual learning in context-based learning environments on the level of the concepts' elements. Consequently, we consider it promising to engage students in learning environments that are heterogeneously contextualised (covering a broad span of contexts), as more heterogeneous contexts should imply a variety of new causal net elements or new readout strategies that need to be incorporated in the students' prior causal net, broadening the span of the coordination class, and thus making transfer of conceptual knowledge easier than it usually is (Barnett and Ceci 2002; diSessa and Wagner 2005).

In addition to this cognitively motivated claim for a heterogeneous composition of contexts, we also see affective benefits in such an approach. Albeit the design of the current study is driven by diSessa's knowledge in pieces approach (implying a more cognitively driven reasoning), it is of importance to consider motivational aspects as well as they are inherently interwoven (cf. Taasoobshirazi and Sinatra 2011; Zembylas 2005). When educators or test designers select contexts for learning environments, they should consider the effect that some groups of students might be advantaged or disadvantaged by certain types of contexts, be it in situations of learning or assessment. McCullough (2004) for example found the Force Concept

Inventory (FCI—Hestenes et al. 1992) to include predominantly stereotypically male contexts. She developed a modified version including stereotypically female contexts. This revision resulted in a significantly lowered gender gap. With regard to context-based learning environments, this finding prompts covering a broad field of interests assumed to appeal to different interest groups as far as possible (be it gender, religion or ethnicity) included in a learning community when choosing proper learning contexts. With reference to Krapp's (1999) model of interest, motivation and learning such an approach is more likely to evoke the emotional component of an individual student's interest and hereby increase motivation and learning success. Due to a higher potential of personal affect, heterogeneously assembled contexts might also increase the probability that students award relevance[4] to a certain concept embedded in this contextualisation—be it for example personal, ecological or societal (cf. Stuckey et al. 2013).

Regarding energy as an abstract cross-cutting concept which possesses relevance not only in the scientific but also in societal discourses, educators should pay even more attention to the specific contexts students are supposed to get engaged in. In contrast to other scientific concepts, energy is not directly observable. It only can be inferred with the help of its connections to other concepts and by mathematization. Furthermore, these inferences differ within scientific and societal areas and are constrained by a specific use of language. Accordingly, energy appears to cut both ways: it offers the chance of simplification by unification but conversely there is a risk of confusion and misinterpretation. Thus, the constant struggle of finding the proper balance of simplification and specification when teaching energy will remain.

In conclusion, cognitive as well as affective aspects of learning are beneficially interconnected in an approach of heterogeneously structured context-based learning about energy. To test this assumption, we put our main research questions as follows:

- Does learning about the energy concept in a heterogeneous learning environment increase students' achievement in a test on the energy concept more than in a homogeneous learning environment?
- Does learning about the energy concept in a heterogeneous learning environment increase the *span* of contexts students are able to successfully work on?

Hypothesis

Considering the theoretical issues concerning the nature of concepts and contexts, we assume learning in a more heterogeneous field of contexts to be more effective in comparison to a more homogeneous field of contexts. In a *mechanistic* regard, we assume this effect to be evoked by the cognitive potential of the contextualised learning settings used, since motivational aspects of context-based learning are supposed to be of little relevance given the short term of the intervention (see next section). The students' ability to transfer conceptual knowledge, i.e. the *span* students are able to successfully perform in (diSessa and Sherin 1998), is expected to be larger due to the increased number of conceptual elements aligned

---

[4] We are aware that what is *relevant* for whom and for how long is always difficult to determine and we do not want to contribute to this discussion here but argue that at least for time periods manageable to react on as a school, teacher or science educator, relevant contexts for the main scientific topics can be found consensually.

more or less consciously when engaging in such a heterogeneous field of contexts. Since when successful transfer, as for example Barnett and Ceci (2002), diSessa and Wagner (2005), Gentner (1989) or Salomon and Perkins (1989) claim, predominantly depends on contextual distance and conceptual as well as contextual familiarity, the former should be reduced by a knowledge base build from a broadly structured learning environment.

In opposition, one could also argue that a span of contexts which is too broad might hinder students to see the conceptual bridges in between them. However, in our opinion, if the conceptual basis is made sufficiently explicit, this will not be a risk and with the help of a well reflected, heterogeneous structure of contexts, conceptual and contextual familiarity can be increased and thus potential distance of further transfer tasks decreased. This argument is also supported by Marton, arguing that "rather than focusing on relations between two isolated situations, we should focus on relations between sets of situations that have certain relevant aspects in common" (2006, p. 503). These common relevant aspects are represented by features of the energy concept in this study and the sets of situations are reflected by the differing kinds of contextualisation (as depicted in the next chapter).

Research Design and Methodology

Departing from the working definition of context as contentual surrounding elaborated above, we planned a pre-post-intervention study to test the idea of heterogeneous contexts as learning opportunity for conceptual learning. The conceptual focus of the study is the energy concept.

To elicit active learning and a considerable portion of communication, the setup of the study included pairs of students working on experiments in a laboratory environment. Due to organisational reasons, students only came to the laboratory once. We therefore had limited time available (about 3 h in total with 2 h of active engagement with the learning environment), but still enough for learning to occur (cf. Dawson 2014). Before and after working on the experiments, students had to solve an energy test (see below), as well as questionnaires on cognitive abilities and interest. Students were videotaped during their work in the laboratory for further qualitative analysis of the learning process.

Selection of Participants

Teachers in ten schools in the larger rural area of Kiel (Germany) were addressed to participate in this study. Four out of these 10 schools officially replied and students of these schools were invited to the laboratory (after providing parental consent). The distribution of the 32 students across schools and gender is depicted in Table 1:

Table 1 Distribution of participants across schools and gender

Gender	School (Nr.)				
	1	2	3	4	Total
Male	8	2	2	1	13
Female	8	2	4	5	19
Total	16	4	6	6	32

Sixteen participants worked on each learning setting, whereby 10 female students worked on the heterogeneous setting (63%) and 9 female students worked on the homogeneous one (56%). The acquisition of participants stretched over a period of about 4.5 months (Feb 28 to Jul 10 2014). Correlation analyses indicated that there was no systematic relation between pre-test scores and testing date ($r = 0.27$, $p = 0.14$), as might have been expected due to the ongoing teaching at school.

Learning Settings

A major challenge in answering the research question at issue was to develop learning environments that are different in their contextualisation but comparable concerning other influencing factors like structure, modus or difficulty. Sub-contexts relating to a common frame topic were chosen to ensure a basic level of contentual comparability. With reference to recent media discourses concerning energy, we selected the frame topic *sustainable usage of energy*. Within this topic, appropriate sub-contexts were differentiated. With the intention of fostering active collaborative engagement with these contexts, a general premise was that students could work in these sub-contexts experimentally. Furthermore, the experiments needed to include different basic elements of the energy concept, as elaborated above (e.g. forms, transformations, degradation and conservation).

As there is no explicit framework characterising the homogeneity or heterogeneity of contexts, at least to our knowledge, a heuristic related to ontological categories was used. Homogeneity in this sense means that the sub-contexts can be described in one ontological category whereas heterogeneity exceeds this category. We developed sub-contexts fitting into the category *power plants* to constitute the homogeneous learning setting. The heterogeneous setting is compiled of sub-contexts not being compatible under a single category. Table 2 depicts the structure of the learning settings and some of their contentual features:

All the experiments in both settings focus on the energy concept. The experiments in the homogeneous setting, however, predominantly allude to physics content, whereas the experiments in the heterogeneous setting also touch chemistry and biology content (c.f. Barnett and Ceci 2002). Moreover, the energy-related societal contexts are more narrow in the homogeneous case, as different types of power plants mainly cope with aspects of energy supply. In addition, the parameters students had to measure during the experiments were more diverse in the heterogeneous setting. Based on these measurements, calculations of the energy amounts converted during the experiments had to be made. To be clear, heterogeneity or homogeneity, respectively, in this contribution is a result of compiling several differently contextualised experiments to a set and not a specific feature of each context. Each of the experiments on its own is not intended to be more or less heterogenic compared to one another; they only differ in their context.

Besides ensuring the learning settings' contextual heterogeneity and homogeneity, respectively, warranting comparable cognitive demands and affective potential is also essential with regard to the research question. Therefore, we made the settings as comparable as possible, even though being aware that exact experimental conditions cannot be achieved as contextuality is an individual construction of the interdependent relation of a scientific content and its frame of reference. Structural comparability was achieved by implementing scripts guiding students through the process of experimenting and analysis. Each sub-script (one for each sub-context) consisted of an introductory text accompanied by an illustrative picture, a picture of the experimental setup, a description of experimental conduction and five

Table 2 Structure and contentual features of the learning settings

Setting	Experiment	Concept	School subject	Content focus	STS context	Measured parameter
Homogeneous setting	Wind turbine	Energy	Physics	Mechanics	Energy supply	Current and voltage
	Coal power plant	Energy	Physics	Mechanics/thermo dynamics	Energy supply	Current and voltage
	Water power plant	Energy	Physics	Mechanics	Energy supply	Current and voltage
	Solar power plant	Energy	Physics/chemistry	Quantum physics	Energy supply	Current and voltage/temperature
Heterogeneous setting	Wind turbine	Energy	Physics	Mechanics	Energy supply	Current and voltage
	Photosynthesis	Energy	Biology	Metabolism	Energy transformation	–
	Eco-fuel	Energy	Chemistry	Chemical reaction	Regenerative vs. fossil fuels	Temperature
	Power-to-gas	Energy	Chemistry	Electrolysis	Energy transport and storage	Volume

subsequent analysis questions (cf. appendix A). Text length and complexity (Flesch 1948) were kept comparable across settings in the introduction. They also included a comparatively high amount of energy-related words (e.g. forms of energy—for an overview comparing the introduction text parameters see appendix B).

Students were confronted with five guiding questions for the analysis after conducting the experiment. These questions were designed according to criteria of complexity (Bernholt and Parchmann 2011) and energy aspects (c.f. Neumann et al. 2013). Each of the five questions corresponded to one of the five levels of complexity introduced by Bernholt and Parchmann (2011): (1) unreflected everyday experience, (2) facts, (3) description of processes, (4) univariate causality and (5) multivariate causality—for examples, see appendix A. The questions were presented in increasing complexity to first give students easier questions and to then increase the affordance level within each context systematically. Furthermore, each set of analysis questions demanded the description of an energy transformation and the calculation of the experiments' energy conversion efficiency including the identification of possible spots of energy degradation. The questions are not considered part of the assessment but were mainly used to stimulate the students' conversation about the phenomena, experiments and conceptual aspects at hand.

Regarding the selection of contexts in this study, the design of the rather structured learning environments, as well as the formulation of the assessment items, we restricted ourselves to features that we expected students are familiar with. In consequence, these elements are rather traditional and less in line with innovative conceptions of context-based learning approaches (e.g. Prins et al. 2016). The rationale behind this decision was to enable students to focus on the experiments and tasks presented in the learning environments in a way they are used to from their experience in school, and to minimise any conflicting aspects induced by unfamiliar methodological or content features of the learning environment.

Summing up, each pair of participants experienced the following schedule: after arriving, a pre-test given 25 min to answer was written (the test was identical in both the homogeneous and the heterogeneous setting; see below). Afterwards, the students worked on the four experiments on energy (depending on condition, being contextualised either heterongenically or homogenically). Each experiment was introduced by a short text on a work sheet to set out the context. The experimental tasks were also given on this sheet. Having finished the experimental part, participants answered the five guiding questions to the experimenter (first author). Each pair of students worked on four differently contextualised experiments, one of which is identical between the two settings (wind turbines). In consequence, students in both settings (homogeneous and heterogeneous) perceive the same sequence and are confronted with almost the same cognitive demands (in terms of the affordances of the introductory texts, the experiments and the analysis questions). The only difference was that students either worked on four power plant contexts (homogeneous setting) or with four more diverse contexts (heterogeneous setting: wind turbine, photosynthesis, eco-fuel, power-to-gas), thus being confronted with different contextual and cognitive elements due to the different contexts. Having finished the discussion of the experiments, the post-test was assigned (given 20 min to work on, again identical in both settings).

Assessment

In addition to characterising students' learning process, conditional variables have to be assessed. Deemed as important variables determining successful learning, cognitive abilities, interests in science and scientific self-concept were surveyed.

The figural scale of the 10th grade cognitive ability test was used to partially assess differences in the individuals' general intelligence (Heller and Perleth 2000). Data concerning scientific interests and self-concept was gathered with the help of an adapted version of a test developed by Klostermann (2012). The test on the energy concept is a compilation of items adapted from Neumann et al. (2013), Bodzin (2012), and the Energy Concept Inventory (ECI 2011) as well as newly developed items relying on quality criteria proposed by Haladyna (2004).

As the structure of the test on the energy concept is of major importance for the interpretation of the results presented within the next section, it will be described in more detail here. The test consists of 25 items referring to the sub-contexts covered in the learning settings as well as new contexts. Three items cover the sub-context of wind turbines, which has been worked on by both groups. Additionally, two items always refer to each of the other six sub-contexts (*homogeneous setting*: photovoltaic power plant, coal power plant, water power plant; *heterogeneous setting*: photosynthesis, eco-fuels, power-to-gas technology). Further, six items cover contexts that neither of the two groups has encountered during the study (referred to as transfer-contextualised) and lastly, four items from the ECI have been implemented to include some kind of abstract, non-contextualised tasks[5] (further referred to as transfer-abstract). Likewise to the questions embedded in the actual learning setting, criteria of complexity (Bernholt and Parchmann 2011) and energy aspects (Neumann et al. 2013) were equally distributed across the test items.

The test was used both as pre- and post-test. Cronbach's Alpha as an indicator of test reliability showed acceptable results (pre-test: $\alpha_{raw} = 0.80$; $\alpha_{std} = 0.81$/post-test: $\alpha_{raw} = 0.77$; $\alpha_{std} = 0.75$), especially taking into consideration the small sample size ($N = 32$) (Cronbach 1951). Table 3 gives an overview of the test structure. Exemplary items reflecting the underlying structure can be found in appendix C.

Methodology

To tackle the research question in a quantitative manner, we conducted analysis of covariance (ANCOVA). This approach enables to check for the influence of the setting on the post-test score under control of the pre-test score. As students had been assigned to the learning environments randomly, the use of ANCOVA should have the highest test power, compared to other possible statistical procedures (e.g. ANOVA with repeated measures; Maxwell and Delaney 2004). All items and single groups of items referring to the sub-contexts, respectively, were calculated.

As this is, at least to our knowledge, the first attempt to study the influence of composition-effects of context-based learning settings, it might be reasonable to raise the common alpha-level of significance (i.e. $p < 0.05$) to not miss the chance of detecting effects in this exploratory study, especially regarding the limited sample size. Otherwise, as we intend to analyse if there is both an overall effect concerning students' learning gain and a differentiated effect concerning the different groups of items (i.e. in terms of transfer), the alpha-level might also need to be decreased in order to control for multiple testing. As it is difficult to balance out both effects, we will stick to the common alpha-level of $p = 0.05$ for deciding about significant

---

[5] As elaborated on above, there might not be items or learning settings that are not contextualised at all but we feel these specific ones to be more abstracted or less contextualised (c.f. van Oers 1998)

Table 3 Item structure of the test on the energy concept

Item group	Number of items	Sub-contexts/topics
Wind turbine	3	Wind turbine
Homogeneous contexts	6	Photovoltaic power plant, coal power plant, water power plant
Heterogeneous contexts	6	Photosynthesis, eco-fuels, power-to-gas technology
Contextualised transfer	6	Car, campfire, falling apple, battery, sports, energy drink
Abstract transfer	4	Rolling ball (×2), temperature of metal and wooden blocks, falling soft ball

effects, but we will descriptively report other effects that are close to this threshold to provide a more detailed summary of our findings.

Results

Due to the small sample size, we checked for equal sample distribution and possible differences in pre knowledge of the two groups in a first step. In the pre-test, students reached on average a score of 12.03 out of 25 (SD = 5.02). *T* test revealed no significant differences between both groups in the pre-score ($p = 0.757$). After the intervention, students obtained a mean score of 15.84 (SD = 4.27). Accordingly, students' learning gain during the intervention can be characterised by an increase in scores from pre- to post-test of 3.81 (SD = 4.27). Questionnaires of student's interest and self-concept in science, captured on a four-point Likert-scale, revealed a mean of 2.90 (SD = 0.40), and 2.60 (SD = 0.72), respectively, whereby the groups also did not differ significantly ($p = 0.52$, and $p = 0.34$, respectively). Also, no differences in cognitive abilities had been found ($M = 16.47$ of 25 points, SD = 3.31, $p = 0.80$).

Regarding the overall test results, the group membership does not seem to have an influence on the post-test result. There was no significant effect of the learning setting on the post-test score after controlling for pre-test score ($F(1,29) = 0.62$, $p = 0.44$, see also Fig. 2).

Additional ANCOVA analysis also taking covariates into consideration revealed that cognitive abilities ($p = 0.87$) and interest ($p = 0.87$) in science do not influence the learning gain significantly (p values stem from a three-predictor model also including setting and pre-score).

To further investigate if learning in one of the two settings increases students' performance on particular subsets of the test items, we conducted comparable analysis steps for these subsets of items. Firstly, we calculated an ANCOVA on the post-test results controlling for pre-score, but only used the items corresponding to the respective learning setting. In this case, results reveal a clear advantage of the respective group for solving the corresponding items (c.f. Figs. 3 and 4). Learning in a homogeneous setting increases the probability to solve the test items directly associated with this setting ($F(1,29) = 4.71$, $p = 0.038$, $\eta_{Setting} = 0.14$, see also Fig. 3) and learning in a heterogeneous setting acts likewise, although revealing a greater effect size ($F(1,29) = 10.56$, $p = 0.003$, $\eta_{Setting} = 0.27$, see also Fig. 4).

As these results appear to be less surprising, it is of greater interest how the two groups perform on items referring to contexts that have not been covered explicitly in either of the

Fig. 2 Pre-post-test results of the two intervention groups including all 25 test items
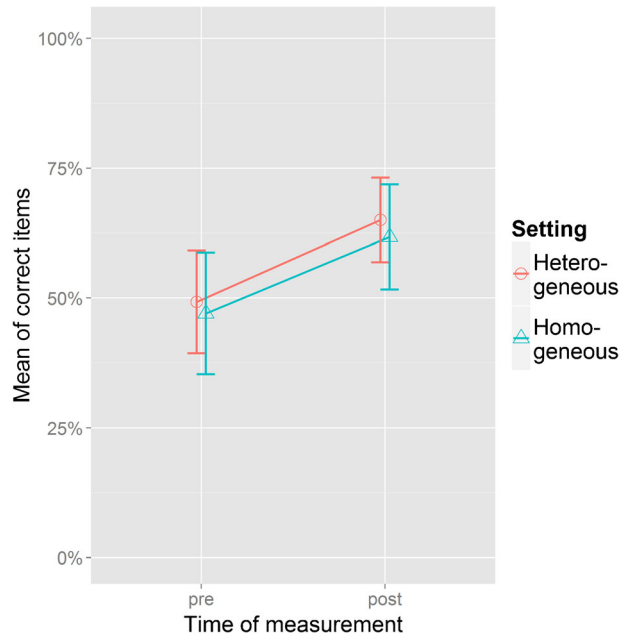
learning settings. Using only the contextualised transfer items, ANCOVA show a significant advantage in favour of the heterogeneous group under control of the pre-score ($F(1,29) = 7.24$, $p = 0.012$, $\eta_{\text{Setting}} = 0.20$, see also Fig. 5). Interestingly, even though only in tendency, the heterogeneous group tends to perform worse on average on the more abstract transfer items after instruction; the homogeneous group remains on the same level ($F(1,29) = 3.46$, $p = 0.073$, see also Fig. 6).

Fig. 3 Pre-post-test results of the two intervention groups including only the six 'homogeneous items'
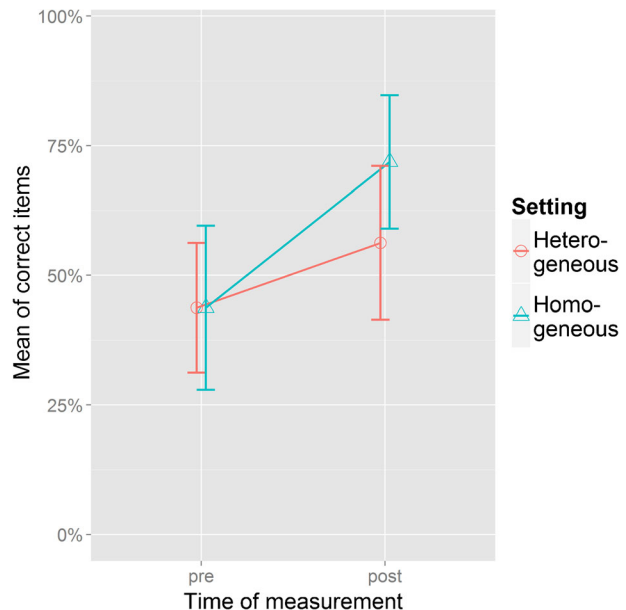
Fig 4 Pre-post-test results of the two intervention groups including only the six 'heterogeneous items'
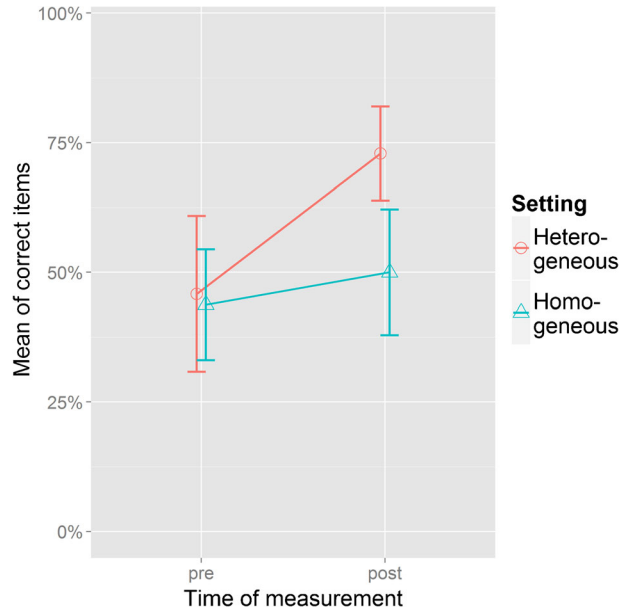
Discussion

Within the current contribution, we investigated possibilities to foster context-based learning based on theories of conceptual change. More explicitly and mainly relying on diSessa's Knowledge in Pieces approach and the connected idea of coordination classes, we implemented an intervention study to analyse if sets of contextualised learning opportunities should be closely related to each other or should cover a broader thematic *span* when having transfer of conceptual knowledge to new contexts in mind as a main learning goal.

The assessment items were carefully selected to cover both the different contexts addressed in the two learning settings and other contexts that were not part of the two learning environments. The statistics obtained from these assessment items indicate that the tests' internal consistency is sufficiently high to interpret the results. In addition, the comparison of pre- and post-test results indicates that students perform better in the assessment after the intervention. This can be interpreted as learning gain, supporting the assumption that our preparation of the different learning environments also was successful at least with regard to provide learning opportunities for the students. While it is also possible that students perform better on the post-test due to repetitively solving the same items (as pre- and post-test are identical), this effect is assumed to be low because of the short time span between both tests. In addition, the effect of increasing scores from pre to post fails to appear in case of the transfer-abstract items which supports the assumption that the increase in test scores can mainly be attributed to learning gains on the students' part and less to repetition effects.

Overall, taking the entire set of test items into consideration, the results of the students' test performance indicate that there is no benefit for either of the groups (homogeneous vs. heterogeneous learning environment). Different explanations might be plausible for this result. First, the absence of differences in the learning gains between both learning settings might be caused by an absence of relevant differences between both settings. For instance, the compilation of the different contexts might not have been successful to drive the assumed cognitive

Fig. 5 Pre-post-test results of the two intervention groups including only the six 'contextualised transfer items'
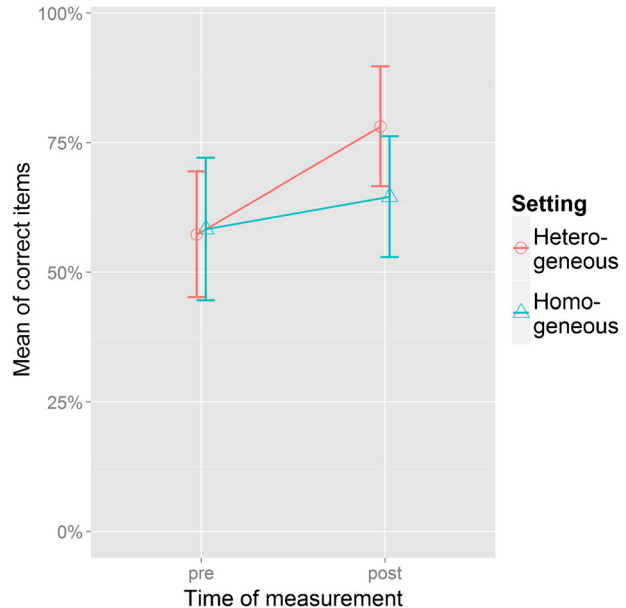

mechanism and to increase the *span* of contexts students in the heterogeneous learning setting are able to process. Alternatively, the assumed cognitive mechanism might just not exist. In addition, the duration of the intervention might be too short so that the benefit of the heterogeneous compilation of contexts does not come into effect in the respective learning setting. Second, there might have been differences in the learning gains between both learning settings, but the assessment was not sensitive enough to detect these differences. We will discuss these possible explanations in the following.

As pointed out in the theoretical section, a major issue in conceptual change research is the question whether students' conceptualisations of normative scientific concepts are coherent or fragmented. Whereas Vosniadou and Ioannides (2001) argue in favour of a coherent view, diSessa (2008) proposes the fragmentation stance. Regardless of the detailed mechanisms supposedly underlying the respective positions, we will discuss the findings of this study in light of this issue on a less detailed level.[6] Taking a coherence perspective, no difference should be expected between the two groups since both had the same learning time and number of contexts to engage with for developing a more sophisticated coherent conceptualisation of energy, which again should be applicable equally well in any context. Results revealing no difference in the total test score between the two learning settings support this hypothesis. Following diSessa and Wagner (2005), the heterogeneous group should be ahead because the more distinct the included sub-contexts are, the higher the chance that new ideas, causal net elements or readout strategies are involved and aligned. Additionally, they argue that this might increase the relevance students see in these certain contexts as well. Even if the test

---

[6] In contrast to the in-depth level of individual learning analysis these theories originated from, the quantitative results of the current study only allow profound discussion on a macro-scale (i.e. inter-individual level).

Fig. 6 Pre-post-test results of the two intervention groups including only the four 'abstract items'
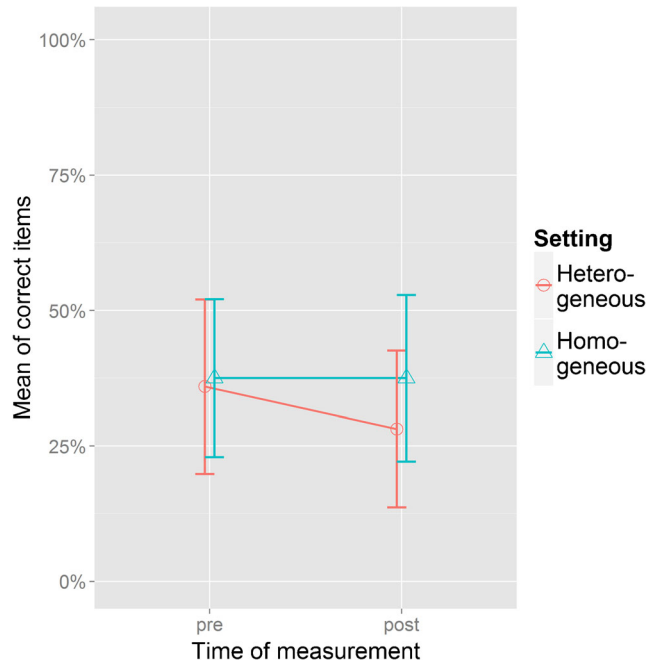
results as a whole do not explicitly confirm these assumptions, results on the level of subsets of items from the assessment provide evidence for this perspective.

Taking a closer look at the differently contextualised subsets of items, different tendencies occur. First, and probably less spectacular, learning about energy in a certain context increases the probability to correctly solve tasks related to the same context. Accordingly, students in both settings perform better on those items that are related to the contexts they just worked in. A similar effect is mentioned by Bennett et al., pointing out possible biases in the connection between intervention and assessment: "*[…] performance on assessment items is linked to the nature of the items used*, i.e. *students following context-based/STS courses perform better on context-based questions than on more conventional questions*" (2007, p. 362).

Second, on items that are unrelated to the set of contexts they learned in, students who learned in a heterogeneous set of contexts outperform their peers who learned in the homogeneous set of contexts. This effect is also nearly stable considering all off-context items (i.e. the transfer-contextualised items and the homogeneous items for the heterogeneous group of students, and vice versa; $F(1,29) = 3.92$, $p = 0.057$, $\eta_{Setting} = 0.12$). As no differences between both groups were found in the pre-test scores in the entire test as well as on all subsets of items, differences in post-test score can be attributed to effects of the intervention. Hence, group-related differences on the different subsets of the achievement test support both the compilation of the different contexts as well as the assumed cognitive mechanism underlying this compilation. In summary, learning in a more heterogeneous set of (sub-)contexts seem to increase the *span* of contexts students are able to process, making it easier to transfer conceptual knowledge to formerly unfamiliar contexts, as indicated by higher scores on the contextualised items in the achievement test.

However, albeit students in the heterogeneous group seem to have acquired a larger *span*—in terms of the number of contexts they are able to apply their knowledge to—it seems that this

span cannot be equated to an abstract, de-contextualised or context-independent understanding of energy: no differences in scores and learning gains were found when comparing between students in both experimental groups on the abstract items. This result is in line with other research, indicating that knowledge acquired from contextualised learning does not easily transcend the learning domain (Kaminski et al. 2005), making it difficult for students in both groups to generalise and *de-contextualise* the general principles and concepts standing behind the context-based learning scenarios they have encountered. Nevertheless, the fact that abstract transfer failed to appear in both experimental groups might also have a background in the limited time of the intervention as well as in the limited number of contexts students were engaged in. Increasing learning time and learning opportunities might end in positive effects as well.

Limitations

On a more general level, the decision to predominantly use multiple choice items in the test limits the generalisability of the obtained results, as a closed format might be insufficient to fully reflect students' conceptual knowledge. In addition, we designed the two learning settings to be internally more (heterogeneous) or less (homogeneous) distinct concerning their contextualisation but at the same time, jointly, comparable to some degree. Nevertheless, we are aware that the criteria mentioned cannot be regarded as exact in the sense of a psychological experiment.

On the students' side, we are faced with the limitations of a comparatively small sample and the aspect that students always worked in pairs. The latter implies that students heavily influence each other and thus also influence the amount of learning taking place during the intervention. The decision to ask students to work in pairs was mainly based on prior experiences, indicating that students working together with peers are much more engaged in talking and elaborating on their ideas than students working alone or when being interviewed by a researcher. Students' verbal engagement in the learning environment, however, was a key requirement for the qualitative analysis of this study, which will be conducted as the next step of this project. Furthermore, experimental work in pairs can be regarded common practice in (especially context-based) science education and therefore a criterion of external validity.

Lastly, it is possible that affective components might be more important during the learning process than expected, albeit we tried to point out that such effects should rather be expected in the longer term. Assessment of situational interest in the experimental sub-contexts (based on a five-point Likert-scale) revealed no significant difference between the settings, but the heterogeneous sub-contexts tended to be slightly more appealing ($M_{\text{ho}} = 3.56$, $M_{\text{he}} = 4.00$, $p = 0.12$). Possibly, further qualitative analysis may give some hints concerning the complex interrelatedness of affection and cognition in this case,

Conclusions and Implications

The current study suggests that the composition of contexts in which students learn about a scientific concept has an influence on the degree of understanding of the concept under discussion. A broader learning span seems to ease transfer of conceptual knowledge to new contexts. We elaborated this stance in advance from a mainly cognitivist point of diSessa's *KiP*

perspective (diSessa 1993), or more precisely, the *coordination class* idea (diSessa and Sherin 1998), emphasising the value of more cognitive considerations of context-based learning. However, the quantitative results provide evidence with regard to different aspects of debates and problems discussed in the theoretical background, related to both the field of conceptual change research as well as research on students' understanding of energy.

Conceptual Change Discourse

The debate whether students' conceptualisations of normative scientific concepts are rather coherent (Vosniadou and Ioannides 2001) or fragmented (diSessa 2008) remains tricky. As shown, results of this study allow an interpretation in favour of either perspective. The absence of differences in the total test score between the two learning settings supports the coherence perspective in broad terms, explaining this result as being due to the same amount of learning time and the same amount of learning opportunities. Conversely, the intervention-related effects on subsets of items might also support the fragmentation stance, arguing that the heterogeneous compilation of sub-contexts has led to triggering and aligning new ideas, causal net elements or readout strategies, resulting in a larger *span* of contexts on the students' side. This study as well as other studies describing concepts as dynamic systems, suggest that both are true—independent of context, repetition of central conceptual features helps internalising relevant knowledge (cf. Brown and Hammer 2008), but transfer as application of these central features in new contexts remains difficult (diSessa and Wagner 2005).

As our study suggests, transfer might be eased when decreasing the distance between the target context and the more familiar contexts, whereby there is certainly need for clarification concerning what actually defines such distance (cf. Dori and Sasson 2013). To us, the taxonomy for transfer proposed by Barnett and Ceci (2002) gives important basic criteria but further empirical validation and specification in the level of detail is needed. Such progress might enable one to describe what particular contextual elements determine a certain distance and how these context elements interfere with the students' knowledge elements.

Learning and Testing the Energy Concept

Describing students' understanding of scientific concepts and especially the development of this understanding requires prior description of the actual concept at issue. In the present case of the energy concept, we relied on a compartmentalised understanding of the energy concept consisting of forms of energy, energy transformations and energy conservation. Other researchers applied similar descriptions (c.f. Duit 1981a; Neumann et al. 2013). Relying the normative energy concept on these compartments implies mastery of tasks and problems centred on these or a combination of these in a certain variety of contexts to be associated with an elaborated conceptualisation of energy. Testing the degree of elaboration for various age groups revealed two major finding: *the average understanding of energy is low*, and *the degree of understanding depends on the test-context*.

Firstly, it is not surprising that, for example according to Euler et al. (2011) or Neumann et al. (2013), the *common understanding of energy is low* since the normative scientific concept of energy is complex, hard to perceive and widespread in its possible applications. Of course it is *involved* in everyday life, but that does not necessarily imply that anybody is reflecting on it. Likewise, grammatical rules or mathematical algorithms are used or even actively applied without explicit conscious thought. Also, researchers' decisions on what

explicit composition of (factual, conceptual, procedural) knowledge to include in a test heavily determine the outcome of such a test (Neumann et al. 2013). Is it for example more important to know *which energy source is likely to run out first?* (Bodzin 2012) or *which form of energy a ball possesses when pushed up a hill*? Or is it, as we primarily required in this study, sufficient to simply describe a certain energy transformation in terms of energy forms without explaining underlying domain-specific concepts like heat, force or chemical reaction and without explaining *why* this happens and *why* we can assign a specific form of energy. Such questions are hard to answer and depend, amongst others, on the researcher's understanding of energy, concepts, relevance or common knowledge. As a result, tests vary dramatically in (conceptual) content and difficulty, thus making fair comparisons difficult. This situation might not be easy to overcome but requires more transparency and clarification.

Secondly, not only the learning process but also the testing of energy concept knowledge is dependent on context (Neumann et al. 2013). Is it fair to ask students for de-contextualised or abstract applications of the energy concept when having learned in a predominantly contextualised manner in advance and vice versa? And even if we agree on testing in contexts, which contexts should we include—technical, everyday life, societal, biological, chemical or physical? These decisions are important since students' interests are connected to students learning and testing (McCullough 2004), and these interests are significantly different (Sjøberg and Schreiner 2010). Applicability of conceptual knowledge in a variety of contexts is a crucial characteristic of competency (Hartig et al. 2008); but as it becomes more evident that students conceptualisations are at least not absolutely coherent and therefore full applicability of all possible contexts is impossible, it should be talked about which direction to head in, in learning as well as in testing: "With the *target* of conceptual change in the case of energy not well understood, little progress is being made in understanding the *process* of change and how to induce it via instruction" (Amin 2009, p. 174). Moreover, we should be more precise about the fact that these targets might change in accordance with technological and societal development. Therefore, science education research as well as schools should make more effort to prepare for plurality and change instead of sticking to apparent coherence and stability in the long term. In conclusion, we agree with Finkelstein (2005), demanding more detailed analysis of the mechanisms involved in context-based learning, but also how effective contexts might be "incorporated in a large-scale, practical manner" (p. 1206). The design of this study was intended to provide a first step in this direction. Relying on quantitative tests results for this first analysis, however, we cannot say anything definite about the mechanisms involved in the learning process or if the actual energy conceptions of participants really have changed. Therefore, we are currently qualitatively analysing the video data of the learning process. Such analysis is necessary to understand the actual *process* of conceptual change on an appropriate scale (Kapon and diSessa 2012). Also, this might be useful to validate the quantitative results in terms of identification of actual learning situation, where understanding might be especially fostered or inhibited according to the intervention groups.

Appendix A: exemplary script

Experiment 1: wind turbine

Wind is a phenomena we are all familiar with—in a positive as well as in a negative regard. From a scientific perspective, wind is a compensation stream of air particles between areas of high air pressure to areas of low air pressure. When the air particles are moving in the stream, they possess *kinetic energy*. Mankind used this circumstance already hundreds of years ago, for example to grind grain in windmills.

Nowadays, wind is often used to produce electricity with the help of wind turbines. Thereby, the movement energy of the air particles (*kinetic energy*) is transformed into electric energy. The energy conversion efficiency of this energy transformation heavily depends on *number form and angle* of the rotor blade. If there is only weak wind, the rotor blades have to be turned into the wind in a manner that they transform the *kinetic energy* into a rotation as efficient as possible. When the wind is too strong and the rotor spins too fast, the turbine even can be damaged. In this case, the blades are adjusted in a way that they transform less energy into rotation. Sometimes, in order not to be damaged, the wind turbines even have to be turned off. Thus, depending on current wind velocity, the amount of electricity produced constantly varies. The contemporary electricity grids sometimes even are incapable of transferring the huge amounts of energy available on really windy days. (source: https://de.wikipedia.org/wiki/Windkraftanlage)
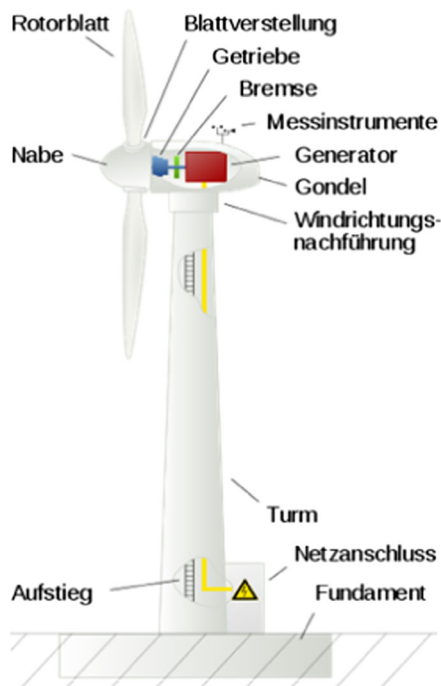
Material

Wind generator, model of a wind turbine, volt and ampere meter, energy meter, cable, bulb, angle protractor

Setup
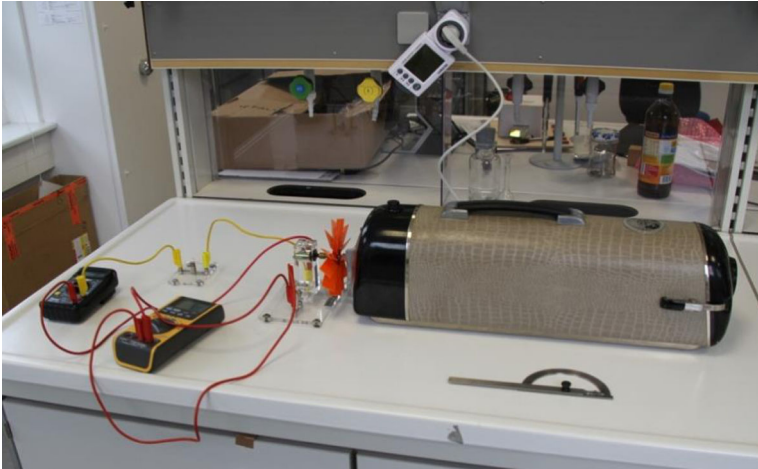
Realisation Turn on the wind generator at maximum power ($p = 400$ W) adjusted to the rotor for 30 s and measure the current and voltage. Vary the angle of the rotor blades for each trial and write down your results.

Trial	Angle (°)	Voltage (U in V)	Current (I in A)
1	5		
2	45		
3	90		

Analysis

1. Give three examples for the use of wind energy from everyday life.
2. In which region of Germany have the most wind turbines?
3. Describe the energy transformations occurring in this experiment. Start at the power plug supplying the wind generator with electricity and end up with the bulb.
   (Example: in a flashlight, the chemical energy of the battery is transformed into electric energy. This electric energy in turn is transformed into thermal and radiation energy by the bulb.
4. Calculate the energy conversion efficiency (ECE/η_i) of the experiment for each of the three angles. Explain the differences and why the energy conversion efficiency of the experiment is so low in total.

Trial	Angle (°)	Voltage (U in V)	Current (I in A)	E_{Effort} (J)	E_{Use} (J)	ECE/η
1	5			12000		
2	45			12000		
3	90			12000		

5. In 2009, scientists of Harvard University reckoned that completely exploited worldwide, wind energy potential would cover the 40-fold of the world's energy demand. Evaluate if it would be useful to switch the world's energy supply completely to wind energy.

Appendix B: text complexity in experimental script

Text complexity in experimental script

Context	Flesch-index	Sentences	Words
Wind turbine	50	14	217
Photosynthesis	45	24	273
Fuels	53	21	266
Power-to-gas	44	18	242
Steam engine	49	17	261
Water power plant	44	18	228
Solar power plant	42	15	237
Text variable	Mean homogeneous	Mean heterogeneous	Level of confidence
Flesch-index	$M_{\text{ho}} = 46.25$	$M_{\text{he}} = 48.00$	$p = 0.56$
Sentences	$M_{\text{ho}} = 16.00$	$M_{\text{he}} = 19.25$	$p = 0.23$
Words	$M_{\text{ho}} = 235.75$	$M_{\text{he}} = 249.50$	$p = 0.42$

Appendix C: exemplary items

Example 1: item group: wind turbine/energy aspect: forms/complexity: facts

Which form of energy can be assigned to the wind that is used to produce electricity in wind turbines?

Example 2: item group: homogeneous/energy aspect: transformations/complexity: description of processes

Solar thermal elements and photovoltaic elements each transform the energy of the sun into another form of energy. Which of the following statements is true?

– Solar thermal elements transform the energy of the sun into electric energy and photo-voltaic elements transform the energy of the sun into thermal energy.

- Solar thermal elements transform the energy of the sun into thermal energy and photo-voltaic elements transform the energy of the sun into kinetic energy.
- Solar thermal elements transform the energy of the sun into kinetic energy and photovoltaic elements transform the energy of the sun into electric energy.
- Solar thermal elements transform the energy of the sun into electric thermal and photo-voltaic elements transform the energy of the sun into electric energy.
- Both transform the energy of the sun into electric energy.

Example 3: item group: heterogeneous/energy aspect: transformations/complexity: description of processes

During the reaction of a fuel with oxygen, energy is released. Which statement is true?

- The energy contained in the fuel is mainly transformed into chemical and thermal energy.
- The energy contained in the fuel is mainly transformed into kinetic and thermal energy.
- The energy contained in the fuel is mainly transformed into light and heat.
- The energy contained in the fuel is mainly transformed into potential and thermal energy.
- The energy contained in the fuel is mainly transformed into thermal and electric energy.

Example 4: item group: contextualised transfer/energy aspect: forms/complexity: facts

An energy drink contains water, sugar, carbonic acid, caffeine and guarin. Which of these ingredients contains the most energy?

Example 5 (ECI 2005): item group: abstract transfer/energy aspect: forms/complexity: Fact

A steel ball is placed at position A on the curved, hard surface shown below. The surface is fixed to the table on which it sits so that it does not move. The ball is held at rest at position A and then is released. It rolls smoothly along the surface. Which choice below most nearly describes the greatest height attained by the ball on the other side of the curve?
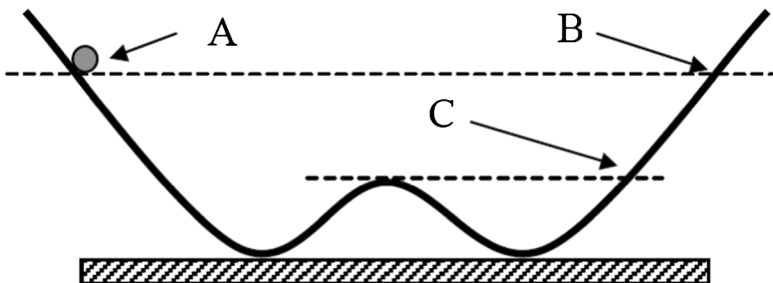
- It will not get over the hill in the middle.

- Significantly below position C
- Almost to position C
- Almost to position B
- Slightly higher than position B

References

AAAS. (2007). *Atlas of science literacy—volume 2: Project 2061*. Washington, DC: American Association for the Advancement of Science.

Aikenhead, G. S. (1994). What is STS teaching? In J. Solomon & G. S. Aikenhead (Eds.), *STS education: international perspectives on reform*. New York: Teachers College Press.

Amin, T. G. (2009). Conceptual metaphor meets conceptual change. *Human Development, 52*(3), 165–197. doi:10.1159/000213891 .

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn?: a taxonomy for far transfer. *Psychological Bulletin, 128*(4), 612–637. doi:10.1037//0033-2909.128.4.612 .

Bennett, J., & Lubben, F. (2006). Context-based chemistry: the salters approach. *International Journal of Science Education, 28*(9), 999–1015.

Bennett, J., Hogarth, S., & Lubben, F. (2004). *A systematic review of the effects of context-based and Science-Technology-Society (STS) approaches in the teaching of secondary science. Review.* York: Department of Educational Studies, University of York.

Bennett, J., Lubben, F., & Hogarth, S. (2007). Bringing science to life: a synthesis of the research evidence on the effects of context-based and STS approaches to science teaching. *Science Education, 91*, 347–370.

Bernholt, S., & Parchmann, I. (2011). Assessing the complexity of students' knowledge in chemistry. *Chemistry Education Research and Practice, 12*, 167–173.

Bliss, J., & Ogborn, J. (1985). Children's choices of uses of energy. *European Journal of Science Education, 7*, 195–203.

Bodzin, A. (2012). Investigating urban eighth-grade students' knowledge of energy resources. *International Journal of Science Education, 34*(8), 1255–1275. doi:10.1080/09500693.2012.661483 .

Boyes, E., & Stanissstreet, M. (1990). Pupils' ideas concerning energy sources. *International Journal of Science Education, 12*(5), 513–529.

Brewe, E. (2011). Energy as a substancelike quantity that flows: theoretical considerations and pedagogical consequences. *Physical ReviewSpecial Topics - Physics Education Research, 7*(2), 1–14.

Brown, D. E. (2014). Students' conceptions as dynamically emergent structures. *Science & Education, 23*(7), 1463–1483.

Brown, D., & Hammer, D. (2008). Conceptual change in physics. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 127–155). New York: Routledge.

Chen, R. F., Eisenkraft, A., Fortus, D., Krajcik, J. S., Neumann, K., Nordine, J., & Scheff, A. (Eds.). (2014). *Teaching and learning of energy in K-12 education*. New York: Springer.

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

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297–334.

Dawson, C. (2014). Towards a conceptual profile: rethinking conceptual mediation in the light of recent cognitive and neuroscientific findings. *Research in Science Education, 44*(3), 389–414. doi:10.1007/s11165-013-9388-4 .

diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction, 10*(2 & 3), 105–122.

diSessa, A. A. (2008). A bird's-eye view of the "Pieces" vs. "Coherence" controversy: (from the "Pieces" Side of the Fence). In S. Vosniadou (Ed.), *Educational psychology handbook series. International handbook of research on conceptual change* (pp. 35–60). New York: Routledge.

diSessa, A. A., & Sherin, B. L. (1998). What changes in conceptual change? *International Journal of Science Education, 20*(10), 1155–1192.

diSessa, A. A., & Wagner, J. F. (2005). What coordination has to say about transfer. In J. P. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 121–154). Greenwich: Information Age Publishing.

Dori, Y. J., & Sasson, I. (2013). A three-attribute transfer skills framework: part I: establishing the model and its relation to chemical education. *Chemistry Education Research and Practice, 14*, 363–375.

Driver, R., & Millar, R. (1986). *Energy matters: proceedings of an invited conference: teaching about energy within the secondary science curriculum*. Leeds: University of Leeds, Center for Studies in Science and Mathematics Education.

Driver, R., & Warrington, L. (1985). Students' use of the principle of energy conservation in problem situations. *Physics Education, 20*, 171–176.

Duit, R. (1981a). Students' notions about the energy concept—before and after physics instruction. In W. Jung, H. Pfundt, & C. von Rhoeneck (Eds.), *Proceedings of the international workshop on "Problems Concerning Students" Representation of Physics and Chemistry Knowledge* (pp. 268–319). Ludwigsburg: Paedagogische Hochschule.

Duit, R. (1981b). Understanding energy as a conserved quantity—remarks on the article by R. U. Sexl. *European Journal of Science Education, 3*, 291–301.

Duit, R. (1986). *Der Energiebegriff im Physikunterricht. IPN / Institut für die Pädagogik der Naturwissenschaften* (Vol. 100). Kiel: IPN.

Duit, R. (1987). Should energy be illustrated as something quasi-material? *International Journal of Science Education, 9*(2), 139–145.

Duit, R. (2010). Kontextorientierter Physikunterricht: PiKo-Brief Nr. 5. In R. Duit (Ed.), *PiKo-Briefe. Der fachdidaktische Forschungsstand kurzgefasst* (pp. 1–4). IPN - Leibniz Institute for Science and Mathematics Education: Kiel.

Duranti, A., & Goodwin, C. (1992). *Rethinking context: language as an interactive phenomenon*. Cambridge: Cambridge University Press.

ECI (2005). Energy Concept Inventory. Version 05.04.26. http://energyeducation.eku.edu/sites/energyeducation.eku.edu/files/EnergyConceptInventory.pdf. Accessed 26 April 2017.

Euler, M., Brückmann, M., Del Borrello, M. (2011). *Energiebildungsstudie Modul 1: Bestandsaufnahme des intendierten Curriculums und Handlungsempfehlungen*. Kiel. Retrieved from http://www.ebs.ipn.uni-kiel.de.

Fechner, S. (2009). *Effects of context-oriented learning on student interest and achievement in chemistry education. Studien zum Physik- und Chemielernen* (Vol. 95). Berlin: Logos.

Feynman, R. P. (1970). *Feynman lectures on physics*. Boston: Addison Wesley Longman.

Finkelstein, N. (2005). Learning physics in context: a study of student learning about electricity and magnetism. *International Journal of Science Education, 27*(10), 1187–1209.

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology, 32*(3), 221–233.

Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 199–241). Cambridge: Cambridge University Press.

Gilbert, J. (2006). On the nature of "Context" in chemical education. *International Journal of Science Education, 28*(9), 957–976. doi:10.1080/09500690600702470 .

Gilbert, J. K., Bulte, A. M. W., & Pilot, A. (2011). Concept development and transfer in context-based science education. *International Journal of Science Education, 33*(6), 817–837.

Glemnitz, I. (2007). *Vertikale Vernetzung im Chemieunterricht. Ein Vergleich von traditionellem Unterricht mit Unterricht nach Chemie im Kontext*. Berlin: Logos.

Glynn, S., & Koballa, T. R. (2005). The contextual teaching and learning instructional approach. In R. E. Yager (Ed.), *Exemplary science. Best practices in professional development* (pp. 75–84). Arlington: NSTA press.

Goldstone, R. L., & Sakamoto, Y. (2003). The transfer of abstract principles governing complex adaptive systems. *Cognitive Psychology, 46*(4), 414–466.

Gräber, W. (1995). Schülerinteressen und deren Berücksichtigung im STS-Unterricht: Ergebnisse einer empirischen Studie zum Chemieunterricht. *Empirische Pädagogik, 9*(2), 221–238.

Greeno, J. (2009). A theory bite on contextualizing, framing, and positioning: a companion to son and goldstone. *Cognition and Instruction, 27*(3), 269–275. doi:10.1080/07370000903014386 .

Haladyna, T. M. (2004). *Developing and validationg multiple-choice test items*. New Jersey: Lawrence Erlbaum Associate Publishers.

Harrer, B. W., Flood, V. J., Wittmann, M. C. (2013). Productive resources in students' ideas about energy: An alternative analysis of Watts' original interview transcripts. *Physical Review Special Topics - Physics Education Research, 9*(2). doi:10.1103/PhysRevSTPER.9.023101 .

Hartig, J., Klieme, E., & Leutner, D. (2008). *Assessment of competencies in educational settings: state of the art and future prospects*. Cambridge: Hogrefe.

Heller, K. A., & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 5.-12. Klassen (KFT 5-12+ R)*. Göttingen: Beltz Testgesellschaft.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher, 30*(3), 141–166.

Kaminski, J. A., Sloutsky, V. M., Heckler, A. F. (2005). Relevant Concreteness and its Effects on Learning and Transfer. In B. Bara, Barsalou, L., & M. Bucciarelli (Eds.), *Proceedings of the XXVII Annual Conference of the Cognitive Science Society* (pp. 1090–1095).

Kaper, W. H., & Goedhart, M. J. (2002). 'Forms of energy', an intermediary language on the road to thermodynamics? *International Journal of Science Education, 24*(1), 81–96.

Kapon, S., & diSessa, A. A. (2012). Reasoning through instructional analogies. *Cognition and Instruction, 30*(3), 261–310. doi:10.1080/07370008.2012.689385 .

Keller, J. M. (1987). Development and use of the ARCS model of motivational design. *Journal of Instructional Development, 10*(3), 2–10.

Klostermann, M. (2012). Vorkurse als Schnittstelle zwischen Schule und Universität [Preparatory courses to bridge school and university]. In S. Bernholt (Ed.), *Inquiry-based learning - Forschendes Lernen* (pp. 224–226). Kiel: IPN.

Kölbach, E., & Sumfleth, E. (2013). Analyzing influences of context-oriented learning while learning with worked-examples in chemistry education. *Zeitschrift für Didaktik der Naturwissenschaften, 19*, 159–188.

Koponen, I. T., & Huttunen, L. (2013). Concept development in learning physics: the case of electric current and voltage revisited. *Science & Education, 22*(9), 2227–2254.

Krapp, A. (1999). Interest, motivation and learning: an educational-psychological perspective. *European Journal of Psychology of Education, 14*(1), 23–40.

Kruger, C. (1990). Some primary teachers' ideas about energy. *Physics Education, 25*(2), 86–91.

Lancor, R. A. (2014). Using student-generated analogies to investigate conceptions of energy: a multidisciplinary study. *International Journal of Science Education, 36*(1), 1–23. doi:10.1080/09500693.2012.714512 .

Lee, H.-S., & Liu, O. L. (2010). Assessing learning progression of energy concepts across middle school grades: the knowledge integration perspective. *Science Education, 94*(4), 665–688. doi:10.1002/sce.20382 .

Liu, X., & McKeough, A. (2005). Developmental growth in students' concept of energy: analysis of selected items from the TIMSS database. *Journal of Research in Science Teaching, 42*(5), 493–517. doi:10.1002/tea.20060 .

Lye, H., Fry, M., & Hart, C. (2001). What does it mean to teach physics 'in context'? *Australian Science Teachers Journal, 48*(1), 16–22.

Marton, F. (2006). Sameness and difference in transfer. *The Journal of the Learning Sciences, 15*(4), 499–535.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data. A model comparison perspective* (Vol. 2). Mahwah: Lawrence Erlbaum Associates.

McCullough, L. (2004). Gender, context, and physics assessment. *Journal of International Women's Studies, 5*(4), 20–30.

Muckenfuß, H. (1995). *Lernen im sinnstiftenden Kontext*. Berlin: Cornelsen.

Murphy, P., Lunn, S., & Jones, L. (2006). The impact of authentic learning on students' engagement with physics. *The Curriculum Journal, 17*, 229–246.

National Research Council. (2011). *A framework for K-12 science education: practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies Press.

Neumann, K., Viering, T., Boone, W. J., & Fischer, H. E. (2013). Towards a learning progression of energy. *Journal of Research in Science Teaching, 50*(2), 162–188. doi:10.1002/tea.21061 .

Nicholls, G., & Ogborn, J. (1993). Dimensions of children's conceptions of energy. *International Journal of Science Education, 15*(1), 73–81. doi:10.1080/0950069930150106 .

Nordine, J., Krajcik, J., & Fortus, D. (2011). Transforming energy instruction in middle school to support integrated understanding and future learning. *Science Education, 95*(4), 670–699.

Özdemir, G., & Clark, D. (2007). An overview of conceptual change theories. *Eurasia Journal of Mathematics Science and Technology Education, 3*(4), 351–361.

Papadouris, N., & Constantinou, C. P. (2016). Investigating middle school students' ability to develop energy as a framework for analyzing simple physical phenomena. *Journal of Research in Science Teaching, 53*(1), 119–145.

Parchmann, I., Gräsel, C., Baer, A., Nentwig, P., Demuth, R., Ralle, B., & Group, t. C. P. (2006). "Chemie im Kontext": a symbiotic implementation of a context-based teaching and learning approach. *International Journal of Science Education, 28*(9), 1041–1062.

Prins, G. T., Bulte, A. M. W., & Pilot, A. (2016). An activity-based instructional framework for transforming authentic modeling practices into meaningful contexts for learning in science education. *Science Education. 100*(6), 1092–1123.

Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: rethinking mechanism of a neglected phenomenon. *Educational Psychologist, 24*(2), 113–142.

Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland [KMK]. (2004). *Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss. Beschluss vom 16.12.2004*. München: Luchterhand.

Sjøberg, S., & Schreiner, C. (2010). *The ROSE project - An overview and key findings*. University of Oslo.

Solomon, J. (1983). Learning about energy: how pupils think in two domains. *European Journal of Science Education, 5*(1), 49–59.

Solomon, J. (1985). Teaching the conservation of energy. *Physics Education, 20*(2), 165–170.

Stuckey, M., Hofstein, A., Mamlok-Naaman, R., & Eilks, I. (2013). The meaning of 'relevance' in science education and its implications for the science curriculum. *Studies in Science Education, 49*(1), 1–34.

Swackhamer, G. (2005). Cognitive resources for understanding energy. Retrieved September 2, 2016, from http://modeling.asu.edu/modeling/CognitiveResources-Energy.pdf.

Taasoobshirazi, G., & Carr, M. (2008). A review and critique of context-based physics instruction and assessment. *Educational Research Review, 3*(2), 155–167.

Taasoobshirazi, G., & Sinatra, G. M. (2011). A structural equation model of conceptual change in physics. *Journal of Research in Science Teaching, 48*(8), 901–918.

Thaden-Koch, T. C., Dufresne, R. J., & Mestre, J. P. (2006). Coordination of knowledge in judging animated motion. *Physical Review Special Topics - Physics Education Research, 2*(2), 1–11.

Trumper, R. (1990). Being constructive: an alternative approach to the teaching of the energy concept—part one. *International Journal of Science Education, 12*(4), 343–354.

Trumper, R. (1993). Children's energy concepts: a cross-age study. *International Journal of Science Education, 15*(2), 139–148.

Trumper, R. (1998). A longitudinal study of physics students' conceptions on energy in pre-service training for high school teachers. *Journal of Science Education and Technology, 7*(4), 311–318.

van Oers, B. (1998). From context to contextualizing. *Learning and Instruction, 8*(6), 473–488.

van Vorst, H. (2013). *Kontextmerkmale und ihr Einfluss auf das Schülerinteresse im Fach Chemie. Studien zum Physik- und Chemielernen* (Vol. 145). Berlin: Logos.

van Vorst, H., Dorschu, A., Fechner, S., Kauertz, A., Krabbe, H., & Sumfleth, E. (2014). Charakterisierung und Strukturierung von Kontexten im naturwissenschaftlichen Unterricht – Vorschlag einer theoretischen Modellierung. *Zeitschrift für Didaktik der Naturwissenschaften*. doi:10.1007/s40573-014-0021-5 .

Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: a study of conceptual change in childhood. *Cognitive Psychology, 24*, 535–585.

Vosniadou, S., & Ioannides, C. (2001). Designing learning environment for teaching mechanics. In D. Psillos, P. Kariotoglou, V. Tselfes, G. Bisdikian, G. Fassoulopoulos, E. Hatzikraniotis, & M. Kallery (Eds.), *Proceedings of the third international conference on science education research in the knowledge based society* (Vol. 1, pp. 105–111). Thessaloniki: Aristotle University of Thessaloniki.

Vosniadou, S., Vamvakoussi, X., & Skopeliti, I. (2008). The framework theory approach to the problem of conceptual change. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 3–31). New York: Routledge.

Warren, J. W. (1982). The nature of energy. *European Journal of Science Education, 4*(3), 295–297.

Watts, D. M. (1983). Some alternative views of energy. *Physics Education, 18*, 213–217.

Wierstra, R. F. A., & Wubbels, T. (1994). Student perception and appraisal of the learning environment: core concepts in the evaluation of the PLON physics curriculum. *Studies in Educational Evaluation, 20*, 437–455.

Winther, A. A., & Volk, T. L. (1994). Comparing achievement of inner-city high school students in traditional versus STS-based chemistry courses. *Journal of Chemical Education, 71*(6), 501–505.

Yager, R. E., & Weld, J. D. (1999). Scope, sequence and coordination: the Iowa Project, a national reform effort in the USA. *International Journal of Science Education, 21*(2), 169–194.

Zembylas, M. (2005). Three perspectives on linking the cognitive and the emotional in science learning: conceptual change, socio-constructivism and poststructuralism. *Studies in Science Education, 41*(1), 91–115. doi:10.1080/03057260508560215 .