A systematic review of research on laboratory work in secondary school

Niklas Gericke, Per Högström & Johan Wallin

Published online: 23 Jun 2022.

Submit your article to this journal ⬀

Article views: 12800

View related articles ⬀

View Crossmark data ⬀

Citing articles: 12 View citing articles ⬀

Routledge
Taylor & Francis Group

REVIEW ARTICLE

A systematic review of research on laboratory work in secondary school

Niklas Gericke [ID][a], Per Högström [ID][b] and Johan Wallin[c]

[a]Department of Environmental and Life Sciences, Karlstad University, Karlstad, Sweden; [b]Department of Education, Halmstad University, Halmstad, Sweden; [c]The Swedish Institute for Educational Research, Sweden

ABSTRACT

We present an integrative mixed-methods systematic review of research on laboratory work in secondary-school science education from 1996 to 2019. The aim of the study is to identify important aspects of how to successfully make use of laboratory work as a science-teaching strategy in secondary schools. By engaging teachers, our study uses a demand-driven approach where the users of evidence participate in setting the scope. Of a sample of 11,771 studies, 39 were selected for the integrative analysis. The result is structured around three theoretical frameworks to inform our understanding of what characterises laboratory work, (1) with the aim of developing students' learning of science, (2) with the aim of developing students' learning to do science (science practices), and (3) regarding the level of inquiry that facilitates aims 1 and 2. The results are discussed in the light of previous research reviews, and recommendations for future research are suggested.

Introduction

Science education is part of every secondary-school curriculum in the world. The process through which science is transformed from an academic discipline into a school subject has been termed recontextualisation (Bernstein, 1996). Hodson (1993) categorised this recontextualized curriculum of science education into three main learning goals: learning science; learning about science; and learning to do science. The first goal, learning science, involves understanding the products of science – the concepts, models and theories. The second goal, learning about science, revolves around the nature and history of science, i.e., how scientific knowledge is developed. The third goal is to learn to do science – to develop the required knowledge and skills for the practice of scientific inquiry. More recently, Hodson (2014) added the fourth learning goal of addressing socioscientific issues (SSIs), which includes developing the critical skills to confront the personal, social, economic, environmental, and moral–ethical aspects of science. In this study, we focus on learning to do science, by conducting a systematic review of the last 20 years of research regarding science education in the context of what we refer to as 'laboratory work' in school settings.

CONTACT Niklas Gericke ✉ niklas.gericke@kau.se

'Doing science' has many names in science education research, including 'practical work' (e.g., Dillon, 2008; Millar, 2004), 'laboratory work' (e.g., Freedman, 1997; Rudolph, 2019), 'hands-on work' (e.g., Caglak, 2017), 'experiments' (e.g., Arnold et al., 2014), 'science practices' (e.g., Crawford, 2014), and 'science inquiry' (e.g., Akuma & Callaghan, 2019). Due to differences in learning goals and in how 'doing science' is addressed, there exists a great overlap in the way in which each terminology is formed (Dillon, 2008). In the UK the term practical work has dominated, while laboratory work has been the dominant terminology in the US (Rudolph, 2020). We chose to use the term 'laboratory work' throughout this review because it has a long history and strong position in secondary school science education (Rudolph, 2019).

In this study, we focused on practical science activities involving students manipulating and/or observing real objects, including empirical data collection within authentic science-teaching situations (usually in the school laboratory or classroom). This criterion is in line with the definition of practical work or laboratory work used in research reviews by Dillon (2008) and Lunetta et al. (2007), and in empirical studies such as Abrahams and Reiss (2012) who use the following definition: 'students, working either individually or in small groups, are involved in manipulating and/or observing real objects and materials' (p. 1036). The reason for this focus lies in the design of this systematic review, which is based on a demand-driven approach where the users, in this case the secondary schools, set the frame for the study (Gough et al., 2019). In piloting our review, we identified students' practical science activities, including empirical data collection, as an area of significant interest, based on the reported need for up-to-date research evidence by practicing secondary-school science teachers. Hence, the secondary teachers' understanding of laboratory work, which also align with common definitions in the literature, was used to guide the choice of review questions and inclusion criteria of the systematic review (see also Method).

Inquiry, also a frequently used term in the field, can briefly be described as a student-led investigation that begins with a question (Crawford, 2014). Hence, inquiry is a rather non-specific term that can be used in many ways in relation to strategies that emphasise the student's role in the learning process. In response to this rather vague understanding of what doing science is all about, the concept of 'science practices' was introduced as a term in the 21st-century science education standards by the National Research Council (National Research Council, 2012). According to Crawford (2014), science practices describe a broader, but articulated, understanding of doing science in comparison to inquiry: 'The idea involves students going beyond experiencing inquiry by interpreting and evaluating data as evidence to developing arguments, explanations, and models' (p. 523). This definition certainly specifies the science in doing science, but also embraces aspects of how scientific knowledge is developed that are not necessarily dealt with in practical laboratory work. Therefore, we consider laboratory work to be the most appropriate concept to represent the defined phenomenon of interest in this study. As argued by Agustian and Seery (2017) 'laboratory education became a separate and distinct component of education, with the emphasis intended to teach students about how to "do science"' (p. 518). It is, nevertheless, important to point out that we include all identified studies covering our definition, regardless of the terminology used.

One of the distinguishing features of science education compared to other areas is the use of laboratory work – activities in which students manipulate and observe real objects to experience and investigate the physical world. Laboratory work is often regarded by teachers and others as central to the appeal and effectiveness of science education (Abrahams & Millar, 2008; Dillon, 2008). Within research in science education, this premise has been strongly emphasised. For example, N. G. Lederman et al. (2014) claim that: 'students will best learn scientific concepts by doing science' (p. 291). Therefore, it is important to analyse if, and under what circumstances, this assumption is supported by empirical research.

It is also of significant interest to study the use of laboratory work in relation to different learning goals, in particular the two overarching goals of learning science and learning to do science. Moreover, to our knowledge, no systematic review has been conducted from a demand perspective. Previous reviews in the field have used a supply- or knowledge-driven approach with theoretical underpinnings, rather than school science demands, as the point of departure. In these reviews, the focus is not explicitly on laboratory work, but on inquiry (Chin & Osborne, 2008; Furtak et al., 2012; Heindl, 2019; Herranen & Aksela, 2019; Lazonder & Harmsen, 2016), practical work (Akuma & Callaghan, 2019), hands-on practices (Caglak, 2017), and science practices (Halawa et al., 2020). The present study thus complements previous studies by taking a demand-driven approach to review studies on laboratory work. We conducted a systematic review of research from the last two decades regarding science education in the context of laboratory work in school settings. The aim of our study was to identify important aspects of how to successfully make use of laboratory work as a science-teaching strategy in secondary schools.

The role of the laboratory in science education

In this, and the following two sections, we describe the theoretical frameworks used to categorise and analyse our findings. We used three frameworks outlining the most important dimensions of laboratory work according to existing theory. The first framework relates to two main goals of laboratory work: to acquire knowledge about subject-specific science concepts, theories, and explanatory models, and to learn how to carry out scientific investigations; the second framework, to the different activities or science practices employed during investigative laboratory work, that is a more fine grained analysis of the learning goal to carry out scientific investigations; and the third framework, to how the teacher regulates the students' influence on the inquiry dimension of their work, which have implications both for developing students' conceptual knowledge and skills to do science.

As pointed out in the introduction, Hodson (2014) argued that 'doing science' (here equated with laboratory work) can be viewed as one learning activity or content of science education curricula that addresses the basic learning goals of science education. In the same paper, he distinguished between four specific learning goals of science education that can be identified, to various extents, in almost all science curricula (p. 2537):

- Learning science – acquiring and developing conceptual and theoretical knowledge

- Learning about science – developing an understanding of the characteristics of scientific knowledge and inquiry, the role and status of the knowledge that science generates, and the social and intellectual circumstances surrounding its origin and development
- Learning to do science – engaging in and developing expertise in scientific inquiry and problem-solving
- Learning SSIs – developing the critical skills to confront the personal, social, economic, environmental and moral–ethical aspects of science in society.

Here we can discern laboratory work as both curricular content and goal, i.e., what students are supposed to learn. According to Hodson (2014), and much of the science education research literature, doing science (laboratory work) is seen as a means of accomplishing both learning science and learning to do science (e.g., Dillon, 2008). Hodson also claims that laboratory work can facilitate the learning of some aspects of learning about science, and emphasises the interrelatedness of conceptual understanding, procedural knowledge, and investigative expertise: 'nor can one learn enough about science by restricting to doing science' (Hodson, 2014, p. 2551). Regarding the goal of learning about SSIs, other activities relating to societal applications, rather than laboratory work, are the focus. Wellington (1998) summarises three main arguments for conducting laboratory work in school science. First, the cognitive argument claims that laboratory work can improve students' understanding of science and promote their conceptual development. Second, the affective argument claims that laboratory work is motivating and exciting, and that it generates interest for science education. Third, the skills argument claims that laboratory work develops students' practical and higher-order thinking skills such as observation, prediction and inference (Wellington, 1998). Similarly, Lunetta et al. (2007) conclude that: 'the school science laboratory is a unique resource that can enhance students' interest, knowledge of science concepts and procedures.' (p. 394). In the current review we focus on the use of laboratory work to enhance student learning of science concepts and practices, whereas we consider student interest to be secondary but a reasonably important mean to support the two other aims.

Millar (2004) argues that there has always been two ways of understanding science – as a product and as a process, and in the curriculum these two aspects are intertwined (as in learning science and learning to do science). The underlying epistemological idea is an empiricist view of science and science education, where students discover the concepts while exploring the observed phenomena (Millar, 2004). This could be described as a form of 'discovery learning' based on inquiry. A second, and perhaps more dominant view in the science education community in recent years is the hypothetico-deductive one, which recognises the clear distinction between data and explanations. By observation and measurement, one can collect data on real-world objects and, as a next step, deduce some specific predictions, which can then be compared to the observed data. If the data and predictions are in agreement, confidence in a match between the explanation and the real world is strengthened.

From an educational point of view, it is important to articulate the distinction between data and explanation, and to recognise that there is no direct route from observations to valid inferences. The connection between the learning activity of laboratory work and the learning goals of learning science and learning to do science have 'gradually shifted, over

the past four decades, away from an inductive and towards a hypothetico-deductive view, the vision of a form of science education which integrates content and process has persisted; curricula and policy documents continue to portray practical activities as vehicles for developing understanding of both science content and enquiry procedure' (Millar, 2004, pp. 4–5). In line with these ideas, we use three theoretical frameworks to elucidate what research gives evidence on how to make use of laboratory work in order to achieve these two learning goals.

The science practices constituting laboratory work

The nature of laboratory work can be described as consisting of different scientific practices, i.e., a set of core knowledge and skill components that are employed when carrying out scientific investigations. The term 'science practices' was first established by the National Research Council (2012) in the United States as a way to more closely relate to what real scientists actually do (i.e., describing 'doing science'), moving away from the more narrow and simplified description of 'the scientific method' that is often raised as a straw-man proposal in school science education, i.e., a 'tendency to reduce scientific practice to a single set of procedures, such as identifying and controlling variables, classifying entities, and identifying sources of error. This tendency overemphasizes experimental investigation at the expense of other practices' (National Research Council, 2012, p. 43). In recent years, science practices have largely replaced inquiry in the literature because the latter can be translated into a solely pedagogical approach. In contrast to inquiry, the term 'science practices' keeps the focus on real science.

The National Research Council (2012) defines eight practices as essential elements of science: (1) asking questions; (2) developing and using models; (3) planning and carrying out investigations; (4) analysing and interpreting data; (5) using mathematical and computational thinking; (6) constructing explanations; (7) engaging in argumentation from evidence; (8) obtaining, evaluating, and communicating information. It is important to point out that the practices should not be seen as a linear sequence of steps; they interact and are used iteratively. Furthermore, not all practices need to be addressed in a single exercise, nor do they need to be employed in the context of laboratory work, even though this is a common way of engaging students in the practices.

In the research literature and policy documents, science practices are often reduced to a smaller number of more overriding categories, partly because secondary-school science can only mimic the work of authentic science. In this recontextualization process, the complexity must be reduced, especially when considering laboratory work as a teaching unit that should introduce students to the full spectrum of science practices. For example, Hodson (2014, p. 2542) suggests a framework of four major elements: (1) a design and planning phase; (2) a performance phase; (3) a reflection phase; (4) a recording and reporting phase. Similarly, the Swedish national curriculum (Swedish National Agency for Education, 2019) makes a comparable distinction of the different elements of laboratory work in the science classroom. In line with our demand-driven approach, we adopted a science practices framework similar to the four-item versions of Hodson (2014) and the Swedish curriculum to stay close to teacher practices: (1) ask questions and plan; (2) investigate and collect data; (3) analyse, interpret, and explain; (4) argument, document,

and communicate. Analysing the findings from empirical studies in relation to this framework enabled us to elucidate what the collection of research can tell us about how to employ scientific investigations in the context of laboratory work regarding each practice.

Level of inquiry – Teacher regulation and student influence

Our third theoretical framework relates to how the teacher regulates the students' influence on the inquiry dimension of the laboratory work, which also corresponds to our third research question. One of the most important aspects of laboratory work as identified in the literature is whether, and to what degree teachers should organise a more fixed trajectory for the students to follow during an inquiry, or if, and to what degree, the students themselves should be responsible for how to manage the science practices (Dobber et al., 2017; Lazonder & Harmsen, 2016).

According to Banchi and Bell (2008), different instructional approaches with regard to teacher regulation can be described as a continuum of four levels. At the first level, confirmatory inquiry, teachers provide students with the question to be investigated and the procedure (method) to be applied, and the results are known in advance. Hence, the students follow a fixed trajectory and have no influence on the question being asked, or on the design of the inquiry or the conclusions that should be drawn. At the second level, structured inquiry, the question and procedure are still provided by the teacher, but the students themselves generate an explanation, supported by the evidence that they have collected, i.e., the teacher does not provide the answer. At the third level, guided inquiry, the teacher provides students with only the research question, and students are tasked with designing the investigation to answer it, as well as with constructing valid explanations. Finally, at the fourth level, open inquiry, all aspects of the scientific investigation are open to the students. Here, students are asked to act like real scientists; deriving questions, designing and carrying out investigations, constructing valid inferences, and communicating their inquiry, largely with no specific guidance from the teacher.

In science education literature, the organisation of inquiry within laboratory work is a highly debated issue (Bybee, 2000), and many studies have been conducted to investigate how to best organise teaching according to this continuum (Dobber et al., 2017; Lazonder & Harmsen, 2016). The organisation of inquiry has also been recognised as important by teachers in practitioners' papers (Puttick et al., 2015). Therefore, we included the organising of inquiry within laboratory work in our review. However, we make no distinction between confirmatory inquiry and structured inquiry because the empirical literature portrays these as practically inseparable, i.e., when the focus is on posing problems and performing procedures. Moreover, in structured inquiry, teacher guidance might be included. A previous meta-analysis indicated that it is difficult to discern any differences in the effects of different kinds of guidance, structured or guided, at a more specific level (Lazonder & Harmsen, 2016). Accordingly, we used only three levels within the third framework: confirmatory inquiry, guided inquiry, and open inquiry.

To summarise, by analysing empirical studies of laboratory work in secondary-school science education, this systematic review elucidates three essential aspects related to laboratory work: the learning goals including conceptual and theoretical knowledge, knowledge and skills specific to the science practices, and teacher regulation in terms

of level of inquiry. By organising our review around the three outlined frameworks, we believe that it covers the most important discussions and controversies within science education research on how to understand and improve the use of laboratory work in secondary schools. Before going into the study, we will present what previous review studies have to say in these matters.

Previous reviews relevant to laboratory work

Previous reviews from last decades have addressed aspects that we believe are relevant to the current study. In those reviews, the focus is on inquiry (Chin & Osborne, 2008; Dobber et al., 2017; Furtak et al., 2012; Heindl, 2019; Herranen & Aksela, 2019; Lazonder & Harmsen, 2016), practical work (Akuma & Callaghan, 2019; Dillon, 2008), hands-on practices (Caglak, 2017), and science practices (Halawa et al., 2020). Four reviews on laboratory work were also found; however, only two partly overlapping reviews looked at secondary school science (Hofstein & Lunetta, 2004; Lunetta et al., 2007), while the third reviewed tertiary-level chemistry education (Agustian & Seery, 2017), and the fourth reviewed practitioner articles (Puttick et al., 2015). In the following sections, we summarise these reviews.

Effects of learning science and learning to do science

Several of the identified reviews were conducted with the aim of evaluating the effectiveness of student-centred inquiry education in comparison to traditional teaching, often equated with teacher-centred confirmatory approaches. However, Hofstein and Lunetta (2004) have precedingly concluded that laboratory investigations offer important opportunities for students to connect science concepts and theories with observations of phenomena and systems. In a meta-analysis of the effect of inquiry teaching, Furtak et al. (2012) analysed 37 studies from 22 articles, and found a moderate effect size in favour of inquiry teaching. Similarly, in a meta-analysis of 13 studies on the effectiveness of inquiry-based learning in science, Heindl (2019) showed a tendency towards improved learning outcomes as a result of inquiry-based learning in comparison to traditional approaches. The effects of inquiry-based learning were more prominent in secondary than in primary education, and the largest effect size was seen when a preparation unit prior to the inquiry-based lesson was included (Heindl, 2019). Another meta-analysis including 15 Turkish studies of hands-on science activities revealed a very large impact (Hedges' $g = 1.55$) on students' science achievement. However, the author raised the possibility of an influence of publication bias on the results (Caglak, 2017). Nevertheless, previous reviews consistently point to student-centred approaches being more beneficial for students' academic achievement than teacher-centred options. Moreover, in a recent review of 12 international studies, Halawa et al. (2020) found that teaching science in accordance with guided inquiry has positive effects on students' science learning, with respect to both conceptual knowledge and procedural skills. In the following, we look into what previous reviews reveal about the nature of teacher support when student-centred inquiry approaches are employed.

Furtak et al. (2012) found that approaches that incorporated teacher guidance had larger effect sizes than open-inquiry strategies with no teacher support. When investigating effects of different aspects of inquiry teaching, they found particularly strong effects

when the students were engaged in the epistemic domain of inquiry, as well as when the procedural, epistemic, and social domains were combined. Similarly, Lazonder and Harmsen (2016) found positive effects of guided inquiry teaching with respect to both content knowledge and inquiry skills, though the effects were twice as large for the latter, indicating the importance of clear guidance for students learning the science practices. However, no obvious differences were identified regarding the specificity of the guidance, suggesting that less specific forms of guidance lead to comparable learning outcomes as more specific guidance. Some age-related differences indicated that more specific guidance was more beneficial for younger learners (Lazonder & Harmsen, 2016).

Based on these reviews, one could claim that there are clear indications of guided inquiry being more effective than open inquiry, and that the effects exist for both learning science and learning to do science. Still, little can be said about the nature of the guidance, and its impact on these two learning goals. Some additional details can be found in a review covering inquiry teaching in general, not only in science education, by Dobber et al. (2017). In their synthesis of 32 studies, they concluded that guided inquiry was promoted as the most effective way of teaching: 'it appears beneficial for teachers to experiment with letting students direct some or all of the inquiry process themselves instead of engaging in teacher-directed inquiry' (p. 203). Furthermore, the authors stressed the difficulties in implementing such approaches. They claimed that the teachers, as well as the students, need distinct training to adapt to the demands and the role change required for inquiry teaching. Dobber et al. (2017) identified three main regulatory approaches and the related teacher roles. All three – metacognitive regulation, social regulation, and conceptual regulation – were relevant for overcoming obstacles to inquiry teaching. Accordingly, teachers can employ metacognitive regulation by improving students' thinking skills, building a culture of inquiry, facilitating guided-inquiry discourse, or familiarising the students with the nature of science. Conceptual regulation can be achieved by providing the students with detailed information on the research topic and helping them focus on their conceptual understanding. Social regulation can be affected by organising groups that mix high and low achievers and focusing on the collaboration processes. In the reviewed studies, metacognitive regulation seemed to be most prominent strategy used to facilitate inquiry-based learning approaches (Dobber et al., 2017).

Trends in studies on laboratory work

In an early review by Lunetta et al. (2007) five trends in studies on laboratory work were identified: articulating and implementing learning goals, applying learning theory organisers, developing classroom communities of inquiries, developing student understanding of the nature of science, and developing inquiry and learning empowering technologies. In the recent review by Halawa et al. (2020), which included 310 articles published between 2008 and 2017, features and trends of inquiry teaching for scientific practices were discussed, and four teaching goals were identified: cognitive, affective, epistemic, and sociocultural. Regarding these goals, which are very similar to the ones formulated by Hodson (2014), Halawa et al. (2020) reported that cognitive aspects dominated and were included in all 310 studies. Affective issues, such as interest and motivation, were addressed in 91 cases, epistemic aspects in 34 and sociocultural aspects such as SSIs in only 6. Thus, a focus on cognitive aspects, as in

our review, mirrors a trend in the recent research literature. Moreover, Halawa et al. (2020) showed that student-centred inquiry, herein denoted as guided and open inquiry, is more frequently investigated and advocated than teacher-centred strategies.

Remarkably, this imbalance in strong favour of student-centred approaches somewhat contradicts what is reported in the practitioner literature. In a review on laboratory-based instruction in biology in *The American Biology Teacher* journal, Puttick et al. (2015) reported that most identified practitioner papers focused on learning of content knowledge through confirmatory inquiry, where student engagement is limited to knowledge construction (similar to structured inquiry). Hence, there seems to be a gap between what teachers actually do and what the science education research community suggests they do, making the present demand-driven systematic review even more pertinent.

Interestingly, in the few papers presenting laboratory work with an emphasis on more student-centred approaches, Puttick et al. (2015) found that the research designs frequently covered a whole series of lessons. It is therefore reasonable to conclude that implementation of the open inquiry approach in science teaching is demanding and time-consuming.

Challenges for teaching laboratory work

In some of the previous reviews, the challenges of implementing science practices in laboratory work were analysed in more detail (Agustian & Seery, 2017; Akuma & Callaghan, 2019; Chin & Osborne, 2008; Herranen & Aksela, 2019). Akuma and Callaghan (2019) investigated the challenges facing secondary-school teachers when teaching practical work using guided and open inquiry. In their review of 29 articles, they analysed difficulties related to the implementation of the practices categorised as initiation, planning, implementation, and evaluation phases. Concerning the initiation phase, they found that if teachers kept the focus on subject-specific requirements and an emphasis on scientific content, they often perceived it as challenging to initiate practical work in line with guided or open inquiry. Regarding the planning phase, teachers often found it difficult to design student exercises for practical work, including multiple coherent phases, i.e., formulating hypotheses, identifying and setting up controls, handling equipment, and designing a reliable and valid plan for the inquiry. In the implementation phase, teachers struggled to engage and challenge the students, and also struggled with finding ways to provide support and to hold back information in order to promote authenticity. Finally, formative and summative assessments of the practical work were likewise found to be challenging for the teachers (Akuma & Callaghan, 2019). Many of these challenges were also found in an early review by Lunetta et al. (2007) indicating their persistence in laboratory teaching.

Chin and Osborne (2008) concluded, in a review of students' questions during inquiry teaching, that those questions can help students monitor their own learning, and explore and scaffold their ideas. However, the analysis also revealed that students do not automatically pose relevant questions, and that this might be a problem in open and guided inquiry. To improve students' performance in terms of both inquiry skills and content knowledge, explicit teaching of questioning skills was thus called for (Chin & Osborne, 2008).

Similarly, Herranen and Aksela (2019) reviewed 30 studies on students' role in posing questions within inquiry teaching that was categorised as successful. Two main reasons for the practice of posing questions were observed. First, questioning was regarded as a significant inquiry skill in and of itself. Second, it was considered to be a driving force for inquiry. Herranen and Aksela (2019) concluded that in the first case, the teacher's role is to teach the students to formulate scientifically meaningful questions, whereas in second case, it is to support the students' questioning skills. So, there seems to be a difference between guiding the students in a certain direction, and letting them to ask their own questions in an active learning process. They also found that inquiry skills and self-confidence among the students were important factors in their ability to pose questions. The promotion and use of epistemic questions had the greatest potential to improve the students' capabilities. Finally, the importance of structuring the questioning within the lessons using specific tools was suggested (Herranen & Aksela, 2019).

In a review of laboratory work in tertiary chemistry education, Agustian and Seery (2017) emphasised the importance of preparing students beforehand. They separated supportive and procedural information and claimed that the former is needed to give students an understanding of the whole laboratory task in general terms – a kind of theoretical framework in which the laboratory experiment operates. Procedural information, on the other hand, related to the specific details necessary for operating in the laboratory, i.e., the skills needed to handle certain equipment (Agustian & Seery, 2017). These findings amplify the importance of the planning phase, and as shown by Akuma and Callaghan (2019), this aspect seems difficult for teachers to accomplish, making it key for further investigation.

To summarise, previous reviews focusing on effects have suggested that student-centred approaches are preferable. However, it is evident that teachers find it difficult to reorient their teaching in this direction, and the present study should contribute to bridging this gap.

Aim and research questions

The aim of this study was to review and analyse the information that can be retained from primary empirical studies performed over the last two decades to answer the following:

(1) What characterises laboratory work in secondary-school science teaching with the aim of developing students' learning of science?
(2) What characterises laboratory work in secondary-school science teaching with the aim of developing students' learning to do science?
(3) How can student-centred teaching, focusing on guided and open inquiry, be implemented in the classroom?

Method

The review's overall scope emerged from the identified need for up-to-date research evidence as reported by practicing science teachers. To engage potential end-users to take part in setting the scope for a review can be described as a demand-driven model of research synthesis. The idea behind acknowledging practitioners as primary stakeholders

is to increase the relevance, and thus usefulness, of the review's questions and answers. This approach aims to summarise and synthesise empirical evidence from primary studies that can help practitioners navigate the body of literature and solve problems experienced in their everyday work. In contrast, a supply- or knowledge-driven review is first and foremost based upon the existing literature, although end-users can make use of its findings (Gough et al., 2019; Pollock et al., 2018; Weiss, 1979).

Focus groups were arranged to engage practitioners as primary stakeholders. Data from these discussions were analysed, along with quality assessments of teaching and learning in the schools, to identify key concepts and to frame areas of interest in relation to actual problems faced by teachers. In total about one hundred teachers and school leaders were invited to take part in informal group discussions to contribute their views on which questions they assess to be the most important for teaching development in schools. All participants were recruited via regional organisations at Swedish universities that are tasked to bridge the gap between research and practice.[1] This methodology appeals to design-based implementation research, where multiple stakeholders' perceptions of problems are outlined together with concerns to develop useful theory, knowledge and tools for teaching practice and also embrace the possibilities for sustained changes in teaching practice (Penuel et al., 2020). Despite laboratory work being part of a long tradition in science education, our initial analyses revealed that teachers find it challenging to navigate among the suggested goals and purposes of the inquiry activities that are employed in schools. Furthermore, teachers expressed a need for guidance on how to think about different levels of inquiry, and purposeful use of materials and equipment, and on how to support communication with and among students concerning the content of a particular laboratory exercise.

The focus of the review was then decided upon by a committee composed of practitioners, school leaders and educational researchers. Review questions, as well as detailed inclusion and exclusion criteria, were clarified by the review team as described below.

Review Design

The methodology used in this integrative mixed-methods systematic review was informed by the PRISMA statement (Moher et al., 2009). The mixed-methods approach allows for the inclusion of a variety of methodologies to understand an educational issue from different viewpoints or aspects. Approaches to combine quantitative and qualitative evidence have garnered interest within systematic reviewing in recent years, due in no small part to a growing understanding of how different research methodologies can provide a more multifaceted picture of the evidence when they are combined (e.g., Gough et al., 2017; Hong et al., 2017, 2020).

Criteria for selecting studies

Criteria for considering studies for this review included the following five aspects of interest: types of participants, learning activities, results or outcome measures, settings or context, and studies. The inclusion and exclusion criteria are summarised in Table 1 and elaborated on in the next section.

Table 1. Inclusion and exclusion criteria.

Aspect of interest	Include	Exclude
Participants	Students in grades 6–12 (about 12–19 years of age) and their teachers.	Research focused on specific groups of students, e.g., second-language learners or students with special needs.
Learning activities	Laboratory work in which students manipulate and/or observe real objects, and collect empirical data within authentic science-teaching situations. Specific interventions aimed at improving teaching and learning in the context of students' laboratory work and scientific inquiry.	
Results or outcome measures	Findings relating to students or students and their teachers. Findings that provide an understanding of students' laboratory work in an educational context aimed at the development of content knowledge and/or understandings of scientific inquiry and ability to carry out scientific investigations. Students' academic achievement in relation to their laboratory work. For studies of intervention effects, two or more groups of students must be compared. Students' (and, if available, their teachers') perceptions and views of teaching and learning experiences in relation to defined laboratory work using, e.g., interviews. Descriptions and interpretations of students' (and, if available, their teachers') engagement and reasoning in relation to laboratory work using, e.g., video observation, field notes or document analysis. The activities studied must be clearly described in terms of context, time, and participants.	Solely teacher views or perspectives. General views or perceptions of laboratory work in school. Self-reported academic achievement. Results regarding only motivational aspects, such as interest or perceived usefulness for future studies, general attitudes, and perceived self-confidence, lacking a clear link to students' development and learning.
Settings or context	Authentic situations in a school setting during regular school hours. Laboratory work within the domains of biology, physics or chemistry subject content, combinations of these domains, or general science. Studies conducted in any country.	Research in the context of special needs education or education for second-language learners. Extracurricular science activities, out-of-school activities such as visits to science centres, authentic workplaces or museums, or outdoor education.
Studies	Primary empirical studies. Studies reported in English or Scandinavian languages. Studies published 1996–2019. Peer-reviewed publications.	

This review focuses on secondary education, when subject-specific teaching normally starts for biology, physics, and chemistry, and when laboratory work becomes an integral part of the science subjects. In the Swedish school system, this corresponds to the final years of compulsory school and upper secondary school, from grade 7 (students 12–13 years old when entering) to grade 12 (students 18–19 years old). Since school systems in different countries may differ in terms of how science education is organised, e.g., in relation to grade levels, we included studies with students from grade 6, or about 12 years of age. Grade 12 was chosen as the upper limit for inclusion; higher education was beyond the scope of the review.

To be eligible, studies had to focus on laboratory work involving students manipulating and/or observing real objects and collecting empirical data within authentic science learning situations (cf., Abrahams & Reiss, 2012). Students' laboratory work could include

any type of empirical scientific investigation, such as experimentation, observation, measurements, or sampling, with or without using technical instruments or equipment. We included studies investigating students' use of digital tools as a complementary activity. However, solely computer-based activities, such as simulations or virtual laboratories, were not included. Another consequence of this criterion is that we excluded studies in which students used only observation reports (data provided to students) and, hence, did not collect the empirical data themselves.

The review was limited to teaching and learning focusing on laboratory work within the domains of biology, physics or chemistry subject content, combinations of these domains, or general science. Consequently, engineering or geoscience content was not included, even though these domains can be part of the science curriculum in some countries. To be eligible, studies had to have a clearly stated interest in students' laboratory work. The activities or exercises in which the students were involved had to be well-described. Evaluations of science curricula or complex multicomponent models, in which students' laboratory work was an inseparable part of a larger whole, were excluded. We further excluded studies in which the laboratory work could be interpreted merely as context. To be included, the study had to have been conducted in authentic situations, i.e., students taught by their own teacher during regular school hours. We included studies conducted in any country.

We included primary empirical studies that used either quantitative or qualitative study designs, or a mixed-methods approach. At the data level, we used both numerical and non-numerical data, including evaluations of effects, and reported interpretations of student and teacher perceptions as well as of interactions between individuals during teaching and learning, both between students and between students and teachers. Qualitative study designs could include, e.g., naturalistic observation concerned with understanding the process of students' meaning-making, interviews to capture students' and teachers' experiences, and analyses of documents used or created adjacent to exercises, such as teachers' written instructions or students' laboratory reports. Quantitative study designs could include intervention studies of either randomised controlled trial (RCT) or non-RCT designs. To be eligible, quantitative studies of intervention effects had to compare at least two groups of learners, and all groups had to be equal in terms of important background measures at the start of the study period. Mixed-method designs could include assessments of the effects of specific interventions combined with follow-up interviews with students and teachers to generate information on, e.g., feasibility aspects.

Only research reports published in 1996 or later, written in English or in Scandinavian languages, were considered for inclusion in the review. The rationale for the 1996 cut-off was the publication of the National Science Education Standards guidelines for science education in primary and secondary schools in the United States, as established by the National Research Council (1996). In our view, these guidelines have had an important impact on science curriculum development, teaching and educational research in many countries (see, e.g., Crawford, 2014).

Literature search

Systematic search strategies were designed and conducted by an information specialist in close collaboration with the researchers on the review team. The focus of the searches was broad and comprehensive. The following databases were searched: The Danish National Research Database, ERIC, Education Source, Oria.no, PsycINFO, Scopus, and SwePub. These databases were selected for their focus on educational research, but also to include the broader social and psychological sciences, as well as Scandinavian research. The search strategy consisted of different combinations of the following search terms: *laboratory, laboratory work, hands-on, practical work, practical skill, practical activity, practical lesson, experiment, investigation, demonstration, i(e)nquiry, classroom activity, active learning, argumentation, hypotheses, reasoning, problem-based, process-based, problem-solving, planning, questioning, science, scientific, physics, chemistry, biology, classroom, teach, learn, student, school, practice, instruction, education, skill, didactic* and *pedagogical*. The English search terms were supplemented in the Scandinavian databases with Swedish, Danish, and Norwegian translations.[2]

Literature searches were conducted in December 2018 and February 2019 in two main steps. First, a broad search was conducted in the ERIC database with search terms related to laboratory work (block A) and science (block B) in an educational school context. For a study to be captured, search terms from both blocks needed to occur in either the title, abstract, or keywords. Based on the results of the first ERIC search, the strategy was slightly modified for the remaining databases, i.e., a search term needed to occur in both the title and abstract.

Study selection and screening process

Rayyan QCRI software (Ouzzani et al., 2016) was used by the review team to enable independent and masked screening of abstracts and full-text publications. Publications retrieved from database searches were imported into Rayyan for screening. Rayyan is a mobile and web-based application that facilitates collaboration between reviewers involved in screening and study selection in a systematic review.

Level one: screening based on abstract information
After removing duplicates, the retrieved list of publications was first subjected to a crude exclusion of irrelevant publications based on title and abstract information. This first level of screening was conducted by two reviewers in an unmasked fashion. In case of uncertainty, the study remained included until the next screening step. In a second step, the remaining abstracts were subjected to masked screening by another two reviewers. Full-text versions of all studies selected by at least one of these reviewers were obtained.

Level two: screening based on full-text information
All studies that passed the abstract-screening phase were assessed as full texts for inclusion independently by two reviewers. When a reviewer excluded a study based on full-text information, the reason for the exclusion was documented. At the end of the screening process, any disagreement was resolved by discussion of the review team.

Level three: quality assessment

The final step of the study selection process involved assessment of potential methodological limitations. All remaining studies from the screening process were assessed using a generic critical appraisal tool addressing the following eight areas of interest:

(1) Research objectives: the objectives, and their relation to previous research and theory, had to be well-described.
(2) Research questions: the questions had to be well defined and answerable.
(3) Data collection: the strategy had to be described and justified.
(4) Data analysis: the analysis had to be described and justified.
(5) Findings and interpretations: these had to be clearly linked to the research objectives and questions, and convincing in relation to the performed data analysis.
(6) Claims: these had to provide a clear answer to the research question and be supported by sufficient evidence, and generalisability had to be clarified.
(7) Reporting quality: reporting had to be comprehensible, well-structured, and coherent.
(8) Ethics: where relevant, ethical considerations had to be clearly reported and in line with good research practice.

The quality assessments were documented by the reviewer and discussed by the review team. To be included in the review, studies had to be relevant to the review questions and of sufficient methodological quality according to the assessment and in relation to the review questions. Hence, the assessment considered not only generic aspects of study implementation and reporting quality, but also potential limitations in relation to our review questions.

Data-Extraction process

When the studies had been selected, relevant information was extracted from each using a coding sheet. Coded information included both descriptive study characteristics and study findings as guided by the review questions and inclusion criteria. Tentative themes were identified to obtain a first, preliminary arrangement of the studies and their findings and to prepare for synthesis. Regardless of whether the information was quantitative or qualitative, all coding had to focus on key concepts as well as succinct summaries of the study findings. In addition to informing the synthesis, extracted information was used to map the selected studies aimed at describing the research field.

Synthesis and assessment of confidence

The current integrative synthesis was guided by both framework and thematic synthesis approaches. Conceptual frameworks were used by the review team primarily to understand and arrange the study findings, rather than to develop new modified frameworks. The frameworks were used to categorise findings of individual studies, with the purpose of developing higher-order review findings, i.e., analytical output from the synthesis

based on findings from primary studies. To assess our evidence claims, we were inspired and guided by the CERQual approach (Confidence in the Evidence from Reviews of Qualitative Research) developed by Lewin et al. (2015).

The review questions served as an overarching starting point for arranging the evidence. In addition to these, we used three conceptual frameworks as outlined in Background: (1) distinguishing the basic learning goals of science education – learning science and learning to do science (Hodson, 1996, 2014); (2) the four-item condensed version of science practices – ask questions and plan; investigate and collect data; analyse, interpret, and explain; argument, document, and communicate (cf., National Research Council, 2012); (3) categorisation of teachers' regulation of three levels of inquiry – confirmatory, guided, and open (cf., Banchi & Bell, 2008).

The synthesis took the form of three main stages that overlapped to some degree: the arrangement of studies with regard to their fit according to selected frameworks; the organisation of condensed summaries of the study findings within the different aspects outlined in these frameworks; and the linking of individual study findings to develop review findings. Since many studies concerned more than one aspect, the same study could contribute to several aspects. As we were guided by our review questions and the frameworks, the extracted study findings did not necessarily always correspond to the authors' core messages. The approach enabled us to discover findings from the included studies that were relevant to the different components of the frameworks. During this process, we had to be vigilant of the risk of misinterpretation by re-examining all individual study findings to check for consistency of interpretation.

The third stage of the synthesis work was done iteratively, by repeatedly and in a cyclical process considering tentative review findings in relation to individual study findings. Synthesis meetings were alternated with rereadings of the studies. The purpose of the meetings was to test, and if necessary, revise, tentative review findings by creating additional abstractions or reformulations. Thus, review findings were the final analytical outputs constructed by the review team, which emerged by relating and linking individual study findings to each other.

Assessment of confidence in the review findings was guided by a modified CERQual approach, based on judgements made for each of the four components: methodological limitations of the collection of studies contributing to a review finding; relevance to the review questions of the studies contributing to a review finding; coherence of the review finding; and adequacy of data supporting a review finding. Tentative review findings were constructed by the review team by linking individual study findings to each other. Every tentative review finding was then assessed in relation to the studies contributing to this finding in order to make sure: (1) there were no major concerns about the design or conduct of the collection of studies supporting the review finding; (2) the review finding was applicable to the context specified in the review questions; (3) the review finding had a clear and cogent connection to individual study findings; (4) the richness or quantity of data supporting the review finding was sufficient. Analysis continued until final review findings were formulated and agreed upon by the review team as reasonable representations of the collective evidence (cf., Lewin et al., 2015). However, it should be noted that we did not grade the overall assessments of confidence on any scale. Our synthesis was focused on constructing valid review findings that were clearly supported by the underlying evidence and likely to be of importance to practitioners as primary stakeholders.
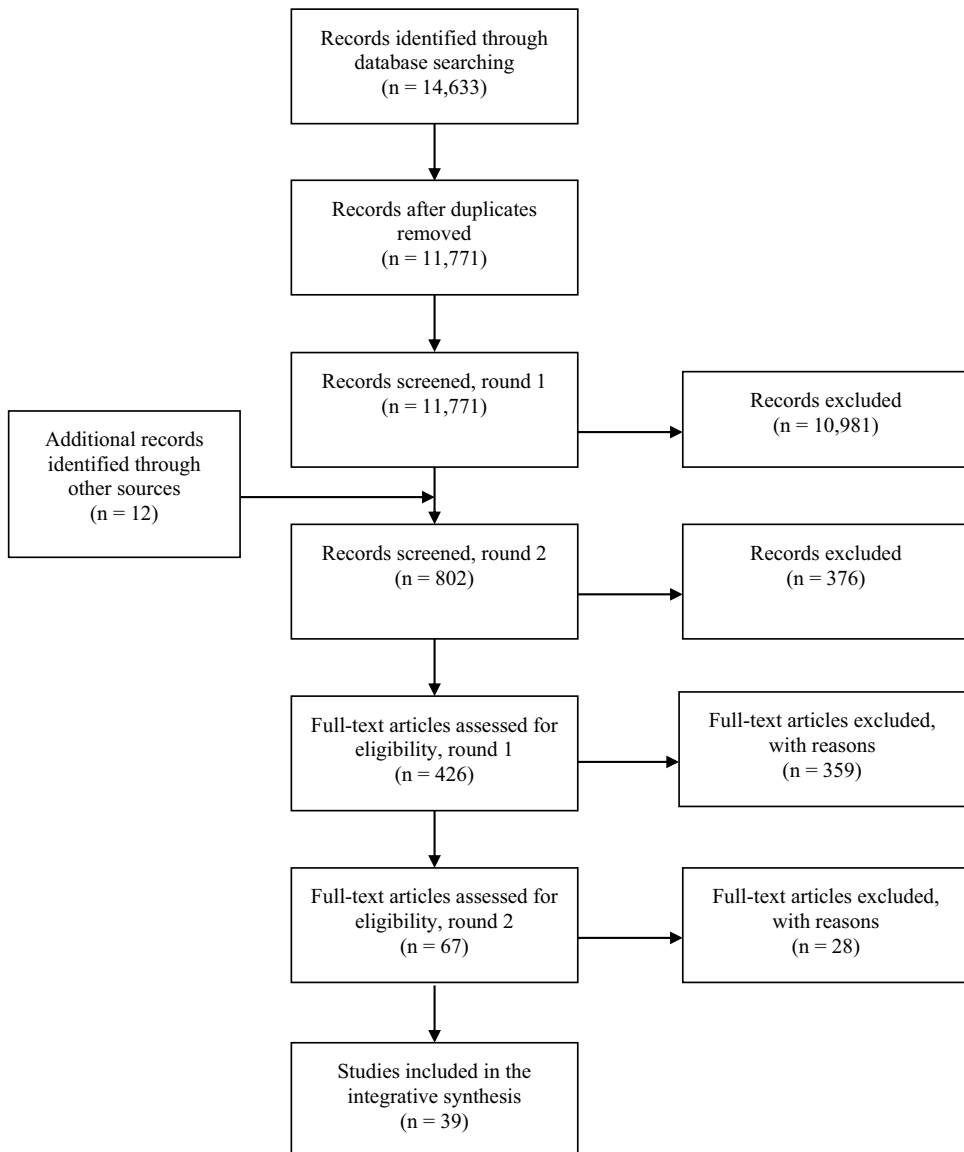
Figure 1. Flowchart of the study selection process

Results

The results of the literature searches and study selection process are summarised in a flowchart (Figure 1). After removal of duplicates, 11,771 studies were screened for inclusion in the review. The study selection process resulted in a final sample of 39 studies. In this final sample, the distribution was relatively even between grades 6–9 and 10–12, respectively. The included studies were conducted in Asia, Europe, North America, and Oceania, with 10 out of 39 conducted in the United States. Regarding school subjects, the distribution was relatively even between biology, physics, and

chemistry. The issues of learning science and learning to do science in relation to students' laboratory work were addressed in 17 and 29 of the included studies, respectively.[3] Five studies explicitly addressed the level of inquiry by comparing at least two different approaches or levels.

Review findings

In this section, we present the review findings structured within the broad domains of learning science, learning to do science, and levels of inquiry. Within each domain, we relate our results to the previously outlined theoretical frameworks to present the review findings in a meaningful way. Individual study findings, and how these correspond to the selected frameworks, are presented in Tables 2 and 3.

Learning science

The learning of science concerns students' development of knowledge about subject-specific science concepts, theories, and explanatory models. It is evident that laboratory work can have a positive effect on students' development of conceptual knowledge. Laboratory work can function as a lever for learning, although students may encounter the same concepts and experience similar challenges through more theoretical learning tasks, such as the creation of concept maps (Freedman, 1997; Hamza & Wickman, 2013; Högström et al., 2010; Lazarowitz & Naim, 2014). However, the way in which the teacher designs the exercises is crucial for success. Observations suggest that students' laboratory work is typically designed as confirmatory inquiry. If an exercise is short-lasting and the subject content is limited, such a design can be successful, but it carries the risk of students overlooking the established explanatory models and feeling less motivated, even as they carry out the work as instructed (Abrahams & Millar, 2008; Abrahams & Reiss, 2012; Högström et al., 2010; Schmidt-Borcherding et al., 2013).

It is challenging for students to develop a conceptual understanding based on laboratory work, and the amount and type of teacher guidance is crucial. Students need teachers' support during, as well as before and after the laboratory work. The extent and difficulty of the subject content, as well as the amount of teaching time available, also play a role. Several studies have evaluated the effectiveness of different levels of inquiry on students' development of subject knowledge (Blanchard et al., 2010; Cheng et al., 2018; Hand et al., 2004; Hodges et al., 2018; Schmidt-Borcherding et al., 2013; Seda Cetin et al., 2018; Sesen & Tarhan, 2013; Strimaitis et al., 2017; Wolf & Fraser, 2008). Overall, it appears that guided inquiry can offer students better learning opportunities than confirmatory inquiry. The guided inquiry can promote students' conceptual understanding as the teacher guides them forward, but in a way that encourages them to regularly reflect on the relevance for the subject content and learning goals. Suggested features of guided inquiry in the literature present a clear strategy by which aspects of the inquiry should be more open to students' creativity, where the teacher answers students' questions with guiding counter-questions instead of giving them direct answers (Blanchard et al., 2010; Cheng et al., 2018; Sesen & Tarhan, 2013; Strimaitis et al., 2017). However, for the approach to be successful, adherence to the strategy seems critical; if the guided inquiry is not carried out as intended, it may be less effective than the confirmatory inquiry. Evaluation of teachers' enactment of assigned

Table 2. Basic information of research studies (N = 39) on science teaching including practical investigations with selected study findings.

Author(s) (year)	Methodology and design	Science discipline(s)	Grade level (age)	n [b]	Study findings [c]
Abrahams & Millar (2008)	QL	Chemistry, Biology, Physics	Key stages 3 and 4 (11–16 y)		The laboratory approach that emerged most clearly was confirmatory inquiry. The students could not, on their own, perceive the subject content related to the laboratory work they were doing.
Abrahams & Reiss (2012)	QL	Chemistry, Biology, Physics	Primary and secondary (4–17 y)	857	Students lacked guiding support from teachers in trying to link doing science with subject content knowledge. The teachers' purposes were not clear.
Allen (2011)	MM	Physics	8 (12–13 y)	52	Students tended to ignore observations that did not match their own expectations.
Andersson & Enghag (2017)	QL	Physics	Upper secondary [a]	20	The students' ways of talking to each other differed depending on the phase of their investigation, and the different types of conversations could fulfill different functions. The students' interpretations and analysis of results were characterized by both disputative and exploratory conversations.
Arnold et al. (2014)	MM	Chemistry	12 (16–19 y)	96	The students had difficulty designing sophisticated control variable experiments individually. The students would have benefited from teacher guidance on how reasoning about methodological issues can be translated into action. The importance of planning the documentation was emphasized.
Blanchard et al. (2010)	CA	Chemistry, Biology, Physics	Middle and High [a]	1700	The students who consistently participated in guided inquiry performed better than those who performed confirmatory inquiry. The students who participated in guided inquiry performed significantly better on questions about methods and procedures than the students who performed confirmatory inquiry.
Carter et al. (1999)	QL/WS	Physics	9 [a]	26	The students' access to everyday experiences was of great importance, even though not always relevant to the investigation. The students' conversations when carrying out their investigations drew attention to their different roles.
Cheng et al. (2018)	QED	Chemistry, Physics	8 [a]	126	The students who were allowed to work according to problem solving performed, on average, significantly better than those who were allowed to do confirmatory inquiry.
Chin & Kayalvizhi (2002)	QN	Chemistry, Biology, Physics	6 (11–12 y)	39	Both repeated practice and collaboration can facilitate students' abilities to formulate questions. A typology is suggested for researchable questions that are suitable in a school context.
Ding & Harskamp (2011)	RCT	Chemistry	11	96	The students in the peer tutoring group and in the collaborative learning group performed, on average, better than students who worked individually.
Fadzil & Saat (2017)	QL	Chemistry, Biology, Physics	Lower secondary [a]	10	The students had difficulties practicing what they had learned theoretically. The students' understanding of and skills in handling lab equipment in an appropriate way seemed to develop hierarchically.

(Continued)

Table 2. (Continued).

Author(s) (year)	Methodology and design	Science discipline(s)	Grade level (age)	n	Study findings [c]
Freedman (1997)	CA	Physics	9 [a]	270	Science teaching that included practical investigations resulted in students achieving, on average, significantly better knowledge test scores. Science teaching that included practical investigations correlated positively with the students' attitudes to physics and to their self-perceived ability.
Garcia-Mila et al. (2011)	MM	Biology	6 (11–12 y)	34	The students who demonstrated knowledge of how variables can be controlled were most successful in planning and performing their investigations. Both the extent and the content of the students' notes seemed to impact their ability to be successful in their investigations.
Hamza & Wickman (2013)	QL	Chemistry	10 (16–17 y)	22	The students generated questions and explanations about the chemistry content when interacting with laboratory materials and equipment.
Hand et al. (2004)	QED	Biology	7 [a]	105	The student groups who used the SWH during their inquiry performed better than the control group. Support from the SWH seemed to create conditions for open inquiry, with greater opportunities for students to take responsibility, which may have contributed to students' ability to understand the relevant scientific concepts.
Haslam & Hamilton (2010)	MM/WS	Physics	Secondary (14–15 y)	96	By gaining access to mutually complementary sources of information, it became less strenuous for the students to understand the instructions of how the inquiry should be performed and how the equipment works.
Hodges et al. (2018)	QED	Chemistry	10 [a]	578	The student group that used the digital learning resource as a complement within the practical investigation performed better than the control group. The possibility for students to visually experience an explanatory model in a dynamic way, in close connection with an investigation, created conditions for their understanding of how observations could be explained.
Hofstein et al. (2004)	QN	Chemistry	11–12 [a]	[b]	A prolonged period of teaching creates conditions for the students to develop their ability to ask scientifically qualified questions. The students' ability to formulate researchable questions was not tied to either a specific inquiry or a specific science content.
Hofstein et al. (2005)	QED	Chemistry	11–12 (17–18 y)	111	The students who had been taught how to formulate researchable questions asked more questions, and a significant part of these questions were more advanced compared to the control group. A prolonged period of teaching, where the students have been trained gradually in their ability to do inquiry, seemed to create conditions for the students to develop their understanding of how scientific research can be conducted.

(Continued)

Table 2. (Continued).

Author(s) (year)	Methodology and design	Science discipline(s)	Grade level (age)	n	Study findings [c]
Högström et al. (2010)	QL	Chemistry	8 (13–14 y)	25	The teacher's purpose with the laboratory work was only perceived to a small extent by the students. During inquiry, the students shifted their focus from understanding the subject content knowledge to handling of the equipment. The teacher's talk and actions were important for the students' interpretation of what they were supposed to learn.
Kanari & Millar (2004)	QL	Chemistry	Primary (10–14 y)	60	The students hypothesized that variables covariate and, in particular, that variables covariate positively. Where covariations appeared, students drew the proper conclusions more easily.
Katchevich et al. (2013)	QED	Chemistry	11–12 [a]	116	Qualitative arguments emerged when the inquiries created unexpected results or gave rise to uncertainties. Open inquiry gave rise to more arguments in students' lab reports. Open inquiry improved students' ability to create arguments related to scientific facts.
Kind et al. (2011)	QL	Chemistry, Physics	8 (12–14 y)	67	The students rarely argued about the scientific content during laboratory work.
Kipnis & Hofstein (2008)	MM	Chemistry	11–12 [a]	[b]	The students' individual planning and formulation of questions stimulated group conversations on a metacognitive level. The students' data collection during inquiry emphasized the importance of their ability to control, monitor and reconsider their experimental design. The students' attempts to interpret and understand collected data revealed misconceptions.
Lazarowitz & Naim (2014)	RCT	Biology	9 (14–15 y)	669	The students who created their own scientific models performed, on average better, especially with respect to basic scientific knowledge.
Lundin & Lindahl (2014)	QL	Biology, Physics	7 and 9 (13 and 15 y)	[b]	The relevance of the teacher's purpose could be jeopardized if contradictory purposes appeared.
Marcum-Dietrich & Ford (2002)	RCT	Chemistry	10 [a]	103	The students who used probeware handled the equipment properly, but had difficulty understanding the purpose of the included procedures. The inquiry approach did not contribute to students' use of a scientific language. The way in which digital measuring equipment was used did not contribute to a developed scientific language.
McRobbie & Thomas (2000)	MM	Chemistry	11 (15–16 y)	21	The students' laboratory experiences reflected the teacher's attitude and teaching. The inquiry did not provide possibilities for the students to find uses for the digital equipment.

(Continued)

Table 2. (Continued).

Author(s) (year)	Methodology and design	Science discipline(s)	Grade level (age)	n	Study findings [c]
Munn et al. (2017)	QED	Biology	9–12 [a]	259	The students tended to stick to stereotypical perceptions of what research is, and the database activity offered new perspectives on scientific research. The students perceived that they performed scientific research to a greater extent within the database activity than in the traditional laboratory work, in terms of analyzing and interpreting data.
Peker & Wallace (2011)	MM	Biology	10 [a]	16	The students' explanations rarely included generalizations. Written documentation showed mismatches in students' experiences. The SWH did not provide students with sufficient support to link own results to established scientific models. Teacher support must accompany SWH teaching methods.
Roth et al. (1997)	QL	Physics	12 (16–18 y)	24	The students constructed different interpretations from identical laboratory investigations. The differences between teachers' and students' backgrounds and experiences raised difficulties.
Schmidt-Borcherding et al. (2013)	RCT	Physics	8 (13–14 y)	173	The guided inquiry seemed to represent a balanced alternative for organizing the laboratory work. The guided inquiry promoted the students' building of conceptual knowledge and their motivation.
Schwichow et al. (2016)	RCT	Physics	8 (12–15 y)	161	The students built conceptual knowledge similarly. The students' conceptual knowledge of variable control experiments was similar. The practical inquiry and the paper-and-pencil inquiry provided similar results.
Seda Cetin et al. (2018)	QED	Chemistry	9 (15–16 y)	60	The students who were taught on the basis of ADI models performed, on average, better than students in the control group. ADI facilitated the students' understanding of subject-specific concepts and changed the students' perceptions of how to produce scientific reports. ADI promoted guided inquiry.
Sesen & Tarhan (2013)	RCT	Chemistry	Secondary (17 y)	62	The students who were taught using a guided inquiry approach performed, on average, better than those in the control group. Guided inquiry was considered useful for learning science and generated positive attitudes.
Strimaitis et al. (2017)	QED	Biology	9–10 [a]	402	The students who used a guided inquiry approach showed greater improvements in conceptual knowledge.
Tan (2008)	QL	Biology	Secondary (16–17 y)	36	The teacher and the textbook, as carriers of knowledge, and the students' own results from observations need to be balanced.
Toplis (2007)	MM	Chemistry, Physics	10–11[a]	17	Even though students were able to identify deviating data, this rarely affected their practical investigations.

Table 2. (Continued).

Author(s) (year)	Methodology and design	Science discipline(s)	Grade level (age)	n	Study findings [c]
Wolf & Fraser (2008)	QED	Physics	7 [a]	165	No mean difference between open inquiry and confirmatory inquiry could be found in knowledge tests. The students who performed open inquiry developed, over time, their skills in handling the equipment. The students who performed open inquiry needed initial support and guidance from teachers, and gradually became confident in finding own possible solutions.

Note. n = number of students in study; RCT = randomised controlled trial; QED = quasi-experimental; QED = quasi-experimental, between-subjects design; WS = quasi-experimental, within–subjects design; QL = qualitative; QN = quantitative; MM = mixed methods; CA = correlation.
[a]Students' age not provided. Grade level may differ between countries of origin.
[b]Sample size not provided.
[c]Students' refers to individual participants who are undergoing education; the synonymous term 'pupil' was used in some research studies.

Table 3. Research studies (N = 39) on science teaching, including practical investigations, mapped within frameworks.

Author(s), year	Framework 1 (n = 39) (RQ1 & RQ2, basic science education goals)		Framework 2 (n = 29) (RQ2, scientific investigations consisting of science practices [a])				Framework 3 (n = 5) (RQ3, compared levels of inquiry)
	Learning science (n = 17)	Learning to do science (n = 29)	Ask questions and plan (n = 9)	Investigate and collect data (n = 14)	Analyze, interpret, and explain (n = 13)	Argument, document and communicate (n = 7)	
Abrahams & Millar (2008)	x, T						
Abrahams & Reiss (2012)	x, T						
Allen (2011)		x					
Andersson & Enghag (2017)		x		x	x		
Arnold et al. (2014)	x, ACK	x	x	x		x	
Blanchard et al. (2010)	x, ACK	x	x				x, GI, OI
Carter et al. (1999)		x		x			
Cheng et al. (2018)	x, ACK						
Chin & Kayalvizhi (2002)		x	x				
Ding & Harskamp (2011)	x, GCO						
Fadzil & Saat (2017)		x		x			
Freedman (1997)	x, AA						
Garcia-Mila et al. (2011)		x	x			x	
Hamza & Wickman (2013)	x, AA						
Hand et al. (2004)	x, STT	x				x	
Haslam & Hamilton (2010)		x		x			
Hodges et al. (2018)	x, STT	x			x		
Hofstein et al. (2004)		x	x				
Hofstein et al. (2005)		x	x				x, CI, OI
Högström et al. (2010)	x, T	x		x			
Kanari & Millar (2004)		x	x				
Katchevich et al. (2013)		x			x	x	x, CI, OI
Kind et al. (2011)		x		x	x		
Kipnis & Hofstein (2008)		x		x	x		
Lazarowitz & Naim (2014)	x, AA		x				
Lundin & Lindahl (2014)		x		x			
Marcum-Dietrich & Ford (2002)		x		x	x	x	

(Continued)

Table 3. (Continued).

Author(s), year	Framework 1 (n = 39) (RQ1 & RQ2, basic science education goals)		Framework 2 (n = 29) (RQ2, scientific investigations consisting of science practices [a])				Framework 3 (n = 5) (RQ3, compared levels of inquiry)
	Learning science (n = 17)	Learning to do science (n = 29)	Ask questions and plan (n = 9)	Investigate and collect data (n = 14)	Analyze, interpret, and explain (n = 13)	Argument, document and communicate (n = 7)	
McRobbie & Thomas (2000)		x		x			
Munn et al. (2017)		x		x	x		
Peker & Wallace (2011)		x			x	x	
Roth et al. (1997)		x			x		
Schmidt-Borcherding et al. (2013)	x, ACK						
Schwichow et al. (2016)	x, AA	x	x	x			
Seda Cetin et al. (2018)	x, STT	x				x	x, CI, GI
Sesen & Tarhan (2013)	x, ACK						
Strimaitis et al. (2017)	x, ACK						
Tan (2008)		x			x		
Toplis (2007)		x			x		
Wolf & Fraser (2008)	x, ACK	x	x	x			x, CI, OI

Note. RQ = research questions in this study; x = study finding within specific framework; T = teaching; ACK = acquire conceptual knowledge; GCO = group or class organisation; AA = alternative activities; STT = specific teaching tools; GI = guided inquiry; CI = confirmatory inquiry; OI = open inquiry.
[a]Practices are related to the context of the lab work as outlined in this study

pedagogical strategy using video data revealed that poor adherence to a guided inquiry instruction is related to worse learning outcomes than would otherwise be expected (Blanchard et al., 2010).

In the case of open inquiry, it is more difficult for students to succeed when the aim of the laboratory activities is to acquire conceptual knowledge. The risk of failure might be too high when students are put in a position where adequate guidance is lacking. Open inquiry requires students to not only understand relevant concepts and theories, but also manage the many aspects of investigative methodology and procedures (Schmidt-Borcherding et al., 2013; Wolf & Fraser, 2008). Too much student responsibility may also be perceived as unclear by some learners, can increase the risk of an unfair workload within groups, and may be differentially effective, favouring male over female students (Wolf & Fraser, 2008). In general, it seems to be more effective to frame students' laboratory work as collaborative or peer-tutoring rather than as individual learning situations (Ding & Harskamp, 2011).

A key factor for students to develop subject knowledge, within the framework of laboratory work, is that they manage to link the collected data to established scientific knowledge (Hand et al., 2004; Hodges et al., 2018; Seda Cetin et al., 2018). The use of instructional approaches or tools, such as Argument-Driven Inquiry (ADI) or Science Writing Heuristic (SWH), which draw attention to the different aspects of a scientific inquiry and integrate writing about content into the teaching activities, increase students' academic achievement (Hand et al., 2004; Seda Cetin et al., 2018). The activity of writing may elicit the importance of the structure of scientific arguments. If writing is integrated into the inquiry process rather than being done after the laboratory work, students directly practice how to express a scientific claim by combining their own data with established knowledge. Moreover, content-specific modelling tools, such as digital visualisations that let students experience a model for chemical reactions, may be effective for students' knowledge acquisition (Hodges et al., 2018).

Learning to do science

To learn to do science within the context of laboratory work concerns both skills and knowledge of methodology, i.e., knowledge of how to carry out investigations and why they should be conducted in certain ways. We used a condensed version of the National Research Council (2012) framework of science practices, as outlined in the Background and Methods sections, to structure the review findings. However, in doing science, it should be noted that the practices interact and are used iteratively; consequently, there is some degree of overlap.

Ask questions and plan. The formulation of scientifically sound questions that can be answered using available resources is a demanding task for students. Therefore, it seems important to focus explicitly on question formulation and introduce students to different categories of questions, which in turn can be answered with different types of investigations. Examples of such question categories are: comparative, explanatory, descriptive, and predictive (Chin & Kayalvizhi, 2002). Better questions can emerge in a learning environment in which students can reason together about different suggestions (Hofstein et al., 2005, 2005; Kipnis & Hofstein, 2008; Wolf & Fraser, 2008).

In the case of experiments, the teaching needs to address the control-of-variables strategy (CVS), i.e., how and why variables are controlled to obtain reliable and valid results (Arnold et al., 2014; Garcia-Mila et al., 2011). Good planning includes finding ways to isolate the variables of interest, as well as considering possible confounders or sources of error that may impact a result (Arnold et al., 2014). Students who demonstrate a good understanding of CVS in their planning are also well placed to conduct experiments successfully (Garcia-Mila et al., 2011). CVS may be taught as part of the students' laboratory work or as part of a theory lesson (Schwichow et al., 2016). Notably, it takes time to build students' capability to formulate appropriate questions and design experiments, and repeated activities combining theory and practice are called for (Arnold et al., 2014; Chin & Kayalvizhi, 2002; Hofstein et al., 2005, 2004; Kipnis & Hofstein, 2008; Wolf & Fraser, 2008).

When students are asked to formulate a hypotheses or are encouraged to make predictions, they tend to prefer correlations that are clear-cut and positive (Kanari & Millar, 2004). Hence, teaching could start with variables that are positively correlated and introduce negative correlation and complex patterns as the students progress in their education.

Investigate and collect data. It may be difficult for students to correctly handle laboratory equipment, to understand why it is designed in a particular way, and to utilise different approaches in their investigation (Arnold et al., 2014; Fadzil & Saat, 2017; Haslam & Hamilton, 2010; McRobbie & Thomas, 2000; Schwichow et al., 2016; Wolf & Fraser, 2008). For example, it seems to be challenging for students to apply theoretical knowledge on how to use specific equipment, or how to design an investigation, in practice (Arnold et al., 2014; Fadzil & Saat, 2017). Moreover, techniques or procedures used in the laboratory may refer more or less to task-specific skills (Schwichow et al., 2016; Wolf & Fraser, 2008). Instructions that refer to the principles behind equipment design and function may enhance the students' ability to use it correctly. The use of printed instructions containing integrated text and illustrations may help students focus on procedures and content, rather than on understanding the instructions (Fadzil & Saat, 2017; Haslam & Hamilton, 2010).

Different types of technology can be used in the laboratory to support the task at hand. Data-collection tools, such as probeware, may be used to improve measurement accuracy and to create graphical representations of the data (Marcum-Dietrich & Ford, 2002; McRobbie & Thomas, 2000). Databases, as repositories of organised data, can be used as a complementing resource for investigation (Munn et al., 2017). However, to be of benefit, it is important that teachers clarify the relevance of the technology in use and carefully monitor students' ideas about inquiry-based work. There is a risk that the adoption of technology will result in a focus on the equipment itself or challenge students' preconceptions about what laboratory work is all about (Marcum-Dietrich & Ford, 2002; McRobbie & Thomas, 2000; Munn et al., 2017).

Student communication while working in the laboratory can be related to their understanding of the content, and to the atmosphere they create through conversations and interactions with the content (Andersson & Enghag, 2017; Carter et al., 1999; Kind et al., 2011; Kipnis & Hofstein, 2008). During laboratory work, students tend to focus their conversations and arguments on what is perceived as manageable, e.g., how to use specific equipment or to carry out a certain procedure. According to Mercer's (1995) talk-type categories, cumulative talk, in which students noncritically and positively build on

each other's ideas, becomes prominent during laboratory work and may help students manage the procedures (Andersson & Enghag, 2017). Nevertheless, individuals' knowledge and experience are important for the groups' ability to find solutions, for example, by affecting student roles or the ability to evaluate ideas (Carter et al., 1999; Kipnis & Hofstein, 2008). How the work is distributed among group members may affect the workload of individual students, as well as the group's opportunities to appropriately solve a task, and thus to learn (Carter et al., 1999; Wolf & Fraser, 2008). An example is when the students themselves appoint someone in a group as the 'expert', regardless of the actual relevance of this individual's knowledge or previous experience (Carter et al., 1999).

During the data-collection phase, students are given the opportunity to practice their ability to monitor and evaluate ideas. When students are encouraged to find creative solutions, their suggestions may require teacher authorisation for safety reasons (Kipnis & Hofstein, 2008; Wolf & Fraser, 2008). However, too much focus on safety procedures can obscure other aspects of the laboratory work (Högström et al., 2010; Lundin & Lindahl, 2014).

Analyse, interpret, and explain. In general, generated data must be processed and linked to explanatory models to be conclusive and meaningful. Students, who do not yet know much about the scientific phenomena they will be investigating, need guidance to meet the requirement to strive for objectivity and to avoid being entrapped by confirmation bias. Students may have a tendency to search for, interpret and favour information in a way that confirms their beliefs or expectations, in particular if the collected data are perceived as indistinct or difficult to interpret (Allen, 2011; Kanari & Millar, 2004; Kind et al., 2011; Toplis, 2007).

Students use different strategies to identify patterns in their data. If the patterns are relatively clear-cut, students manage to interpret the data, for example, by focusing on trends or differences in measurements. Although, it is easier for students to interpret simple positive correlations, it seems important to expose them to more complex data patterns, uncertainties, and unexpected outcomes to stimulate reasoning and critical thinking (Allen, 2011; Arnold et al., 2014; Katchevich et al., 2013; Kind et al., 2011; Toplis, 2007).

For an investigation to be comprehensible and meaningful, it is important to understand the link between the collected data and established scientific knowledge. Our findings emphasise the students' difficulty in making valid generalisations with evidence from their own observations to create meaning (Högström et al., 2010; Kind et al., 2011; Kipnis & Hofstein, 2008; Peker & Wallace, 2011; Roth et al., 1997; Tan, 2008). One way of facilitating students' understanding of these links is to emphasise the importance of creating valid scientific arguments and to establish a learning environment that stimulates students' construction of explanations (Katchevich et al., 2013; Kind et al., 2011; Marcum-Dietrich & Ford, 2002).

Students' opportunities to make meaning of the activities may be strengthened by facilitating the data collection and processing (Marcum-Dietrich & Ford, 2002), or by organising discussions as the laboratory work progresses (Kind et al., 2011). It seems to be important to maintain an argumentative discourse, characterised by collaboration, exploration and reasoning, to prevent students from sticking to their original positions (Andersson & Enghag, 2017; Katchevich et al., 2013; Kipnis & Hofstein, 2008). Inquiry-based activities that are well-balanced in terms of level of difficulty in relation to the specific context, such as the

students' prior knowledge of the phenomenon they are to investigate and time available, seem to provide the best conditions (Katchevich et al., 2013; Kipnis & Hofstein, 2008). Students' argumentative conversations may become sparse when interpretation of the data is either too simple (Katchevich et al., 2013) or too complex (Kind et al., 2011).

Opportunities for students to explore scientific explanatory models while conducting an investigation may be promoted by providing additional resources, such as digital visualisation tools (Hodges et al., 2018) or databases containing large datasets (Munn et al., 2017).

The teacher's responses to students' questions during laboratory work and inquiry are important for how students perceive the work's value. When teachers relate to the knowledge that the students are acquiring, and point out the fact that observations are rarely perfect, the students' understanding is enhanced. Without support from the teacher, the students might stop at simple descriptions of the phenomenon in question, for example, ignoring discrepancies in the data. Students do not easily discover the facts on their own (Högström et al., 2010; Roth et al., 1997; Tan, 2008).

Argument, document, and communicate. Documentation is a central practice of laboratory work. It is not only about collating and displaying data; it is also central to keeping records during the work and reporting investigations. Careful documentation is important as a memory and organisational tool during the laboratory work (Arnold et al., 2014; Garcia-Mila et al., 2011), and as a thinking tool to clarify and justify scientific arguments (Hand et al., 2004; Seda Cetin et al., 2018).

Writing about content, and the ability to construct written arguments ought to be made explicit in students' laboratory reports. The laboratory report may be viewed as a specific genre of scientific text that summarises what has been done and how the retrieved results can be interpreted. Thus, the laboratory report is a documentation of the author's reasoning and thinking in relation to the assignment, which can be critically reviewed by others. Traditionally, laboratory reports are written after the completion of the practical work, but writing may also be integrated into the exercise itself. In general, students find it difficult to express themselves scientifically and it seems to be demanding for students to deal with the transition from spoken to written form of communication (Arnold et al., 2014; Hand et al., 2004; Katchevich et al., 2013; Marcum-Dietrich & Ford, 2002; Peker & Wallace, 2011; Seda Cetin et al., 2018).

Instructional tools and support materials, such as the SWH (Hand et al., 2004; Peker & Wallace, 2011) and ADI (Seda Cetin et al., 2018), can help both learners and teachers organise the work, and can draw students' attention to reflecting on the content as they go forward. Besides aiming to promote structured reasoning and thinking, these tools incorporate a writing-to-learn approach. Integration of writing can help students focus on the scientific content during their work, thus making it understandable and meaningful. The adoption of such instructional approaches may stimulate students to negotiate and critically assess scientific arguments and explanations, but teacher guidance still seems important for students to grasp the links between their own data and established knowledge (Katchevich et al., 2013; Marcum-Dietrich & Ford, 2002; Peker & Wallace, 2011). Giving students the task of communicating their findings to others, rather than just the teacher, may improve the quality of their written conclusions and scientific explanations (Hand et al., 2004).

Teaching approaches related to levels of inquiry

Scientific inquiry is often described as a dynamic and creative process characterised by change. For students to be introduced to the work of real scientists and to develop expertise in doing science, a shift from teacher-centred to more student-centred inquiry has been called for. Clearly, there is potential in letting students take greater responsibility for their laboratory work (Blanchard et al., 2010; Hofstein et al., 2005; Katchevich et al., 2013; Seda Cetin et al., 2018; Strimaitis et al., 2017), although our findings point to a number of issues that need to be addressed for these approaches to be successful. First, a clear strategy is needed to determine what aspects of the laboratory work students should take greater responsibility for, and how, and adherence to the strategy of choice seems important (Blanchard et al., 2010). Second, it may be time-consuming for students to acquire the prerequisite knowledge and skills needed to engage in more high-level scientific inquiry. Repeated activities that emphasise the methodologies and practices of scientific inquiry can gradually build the students' ability to reason in a scientific context (Hofstein et al., 2005; Katchevich et al., 2013). Third, students in higher grades may generally have better opportunities to engage in high-level scientific inquiry, although the level of difficulty in relation to the specific context is crucial for success, regardless of grade level. Although only one of the included studies directly analysed possible age-related differences (Blanchard et al., 2010), most studies indicated improved learning opportunities with greater student responsibility in upper secondary school (Hofstein et al., 2005; Katchevich et al., 2013; Seda Cetin et al., 2018; Strimaitis et al., 2017).

Regarding some of the frameworks of inquiry levels available in the literature (see, e.g., Banchi & Bell, 2008; Blanchard et al., 2010), the idea of gradually increasing student responsibility in a given sequence may be viewed as a starting point. As outlined in the Background section, the logic behind the different instructional approaches is generally described as gradually increasing or decreasing the number of key activities that are open to students or given by the teacher, respectively. According to these models, students start by focusing on an interpretation of the results as the key activity, while data-collection methods and the question for inquiry are given by the teacher. As a next step, data collection and interpretation of the results are open to the students, but the question is given by the teacher. As a final step, students formulate the question as well, thereby taking responsibility for all key aspects of the investigation. Our review findings suggest that a more varied way of viewing students' progress in relation to becoming acquainted with the core ideas of scientific practices may be preferable. Simply put, our findings support the idea of approaching science practices, including laboratory work, as a matrix rather than a sequence as the students progress through science education. For example, it is feasible to design laboratory exercises that emphasise students' ability to ask questions and plan investigations in combination with adequate guidance concerning the other key aspects.

Discussion

In this section, we elaborate on our findings in relation to the three research questions and previous research reviews outlined in the Background section.

Laboratory work as a way of learning science

If the aim of a laboratory activity is to promote students' science learning or conceptual knowledge, it is evident that one of the most important characteristics is the linking of collected data and inferences to established scientific knowledge. Therefore, guided inquiry should be the first choice, but if time is sparse, confirmatory inquiry seems to be more appropriate than open inquiry. This, we would argue, is an important finding, because most previous reviews have indicated that teacher-centred approaches are generally less effective (e.g., Caglak, 2017; Dobber et al., 2017; Furtak et al., 2012; Heindl, 2019). Considering the aim of teaching, it may be important to nuance this conclusion of a limited effectiveness of teacher-centred strategies, particularly when taking into account the contextual limitations frequently experienced by teachers, such as schedule, students' prior knowledge, and time constraints (Puttick et al., 2015). Nevertheless, in the absence of these limitations, guided inquiry seems to be most suitable due to the unbeatable opportunities for formative feedback from the teacher. The teacher's support appears to be essential for students before, during and after an exercise, as implied in previous reviews (Agustian & Seery, 2017; Akuma & Callaghan, 2019; Lunetta et al., 2007).

What, then, should the teacher's guidance look like? This issue is recognised as one of the main obstacles to implementing student-centred laboratory work according to Akuma and Callaghan (2019), especially if the teaching aim is the learning of science content. As shown in this review, an important feature of adequate guidance is that teachers utilise guiding counter-questions in response to the students' ideas in a way that encourages them to regularly reflect on relevance for the subject content during the inquiry process. This finding may cast new light on the issue of the nature of teacher guidance. Lazonder and Harmsen (2016), who investigated types of guidance that differed in specificity, could not discern any significant differences in effectiveness, with the exception of a small advantage for more specific guidance with younger students. However, considerable variation was observed between the estimated effects in relation to the framework used (Lazonder & Harmsen, 2016), and it should be noted that none of the studies included in that review met our inclusion criteria. Here, there is definitely a need for more studies and further development of theoretical frameworks for classifying teacher guidance. On the other hand, the results of this study may give teachers opportunities to investigate ways to implement the portrayed ideas. As such, the demand-driven perspective is still warranted.

Moreover, our results indicate that the integration of student writing into the laboratory work, instead of only as a reporting task at the end, is a powerful strategy to help students maintain their focus on the science content and overall aim of their work. Such writing-to-learn approaches emphasise the ability to present scientific arguments containing established knowledge supported by the collected data. In this way, it might be possible to overcome the many barriers of student-centred laboratory work as noted by Akuma and Callaghan (2019). To overcome such hurdles, teachers may be encouraged to implement the available tools for writing and arguing outlined in this review. These tools can function as support materials to guide the students and catalyse linkage of the generated data with existing knowledge, thereby creating a continuous empowering discussion on relevant concepts rather than on how to manage practicalities, which

otherwise appears to be common in laboratory work. Moreover, social regulation in the form of collaborative or peer-tutoring exercises is acknowledged in our review, similar to what Dobber et al. (2017) stressed as important, as a way to give structure and facilitate discussions around science content during laboratory work.

Laboratory work as a way of learning to do science

In the Results, we outline the findings as specific advice on what to think about for teachers implementing laboratory work in school. Our review findings highlight opportunities as well as challenges, and are structured according to science practices as constituent elements of laboratory inquiry: (1) ask questions and plan; (2) investigate and collect data; (3) analyse, interpret, and explain; and (4) argument, document, and communicate. Even though the science practices can be addressed in a more or less stepwise manner during a certain exercise, our findings support the notion that science practices interact within the inquiry process. To achieve the goal of learning to do science, it seems important to emphasise this reciprocity. Our findings also support the notion that 'engaging in scientific inquiry requires coordination both of knowledge and skill simultaneously' (National Research Council, 2012, p. 41).

Regarding the first science practice in our framework, posing questions is a fundamental skill to develop according to the included studies. As a prerequisite for the ability to formulate sound questions, students need to have an initial understanding of what type of questions are realistic and reasonable in relation to what is achievable, and several categories of investigable questions are proposed. This finding implies that activities that are solely aimed at posing questions should also be addressed prior to students' laboratory work, as mirrored in previous reviews focusing on question posing (Chin & Osborne, 2008; Herranen & Aksela, 2019). In line with Herranen and Aksela (2019), questioning can be seen as an essential skill or as embedded in the learning process, emphasising the importance of teachers' acknowledgement of the function of the scientific question, and the mutual connection between the question and the design of the investigation. Concerning the laboratory experiment, our study highlights the importance of how and why to control variables. The concept of isolating and controlling the variables that are being investigated seems to be particularly challenging for students and this issue, which has not been addressed much in previous reviews, warrants further attention. This finding also emphasises that not only skill, but also knowledge is required in the planning of a scientific investigation.

Regarding the second science practice in our framework, to carry out an investigation and collect data, our findings once again show that students need to be prepared both theoretically and practically to be able to handle the procedures and various instruments. This was pinpointed in a review of laboratory work in tertiary education (Agustian & Seery, 2017), but not specifically mentioned in reviews of secondary education. Our findings highlight students' difficulty in transferring acquired theoretical knowledge to actionable practice in the laboratory and point out that the learning of certain procedures may be task-specific. This implies that teachers need to support students in the process of knowledge transfer, and to carefully reflect upon which laboratory procedures to introduce to their students. While Akuma and Callaghan (2019) acknowledged that teachers find it difficult to challenge and regulate

students' data collection during laboratory work, our findings suggest that the use of technology, such as probeware, may facilitate this, and thus help them overcome certain practical challenges. Moreover, and perhaps surprisingly, we show that students' use of cumulative talk, where the students build positively and noncritically on each other's ideas, might be functional during the implementation phase of the inquiry.

Concerning the third science practice of our framework, to analyse, interpret, and explain, our review findings clearly show that data must be processed and linked to theoretical explanatory models to be meaningful, even when an exercise is aimed at promoting inquiry skills rather than conceptual learning. Hence, the scientific phenomenon and its underlying theory need to be thoroughly introduced, and be continuously present in the minds of the students. This observation may contribute to explaining why meta-analyses have not been able to clearly separate students' acquisition of content knowledge and inquiry skills in terms of effectiveness (e.g., Furtak et al., 2012; Halawa et al., 2020), although the review by Lazonder and Harmsen (2016) is somewhat conflicting in this regard. The ambiguity may relate to the inherently difficult task of differentiating between students' understanding of scientific inquiry and subject-specific knowledge. As pointed out by J. Lederman et al. (2019), the research base on how to assess the knowledge required for inquiry in a valid and reliable way is still small. As might be expected, our findings show that students find it easier to interpret simple positive correlations when conducting experiments. That is why it is necessary to adjust the complexity of the inquiry in the laboratory work to the student group. In terms of progression, it appears reasonable to introduce the principles underlying the CVS using clear-cut examples and gradually increase complexity as students become more experienced.

Regarding the fourth science practice, to formulate scientific arguments, document, and communicate, our review findings emphasise the importance of getting students to pay attention to these aspects during their laboratory work, and not just in a reporting phase at the end of the investigation. By using such a strategy, the criterion of linking theory and practice can be better accommodated, not least by continually reminding the students of how scientific arguments are constructed in terms of both content and structure. Moreover, the explicit activity of student writing seems essential in this regard. Writing tools, as previously mentioned, can help teachers in the process of formative and summative assessments as students move forward within a certain exercise. Furthermore, progression of assessment strategies has been recognised as a barrier for learning (Akuma & Callaghan, 2019; Hofstein & Lunetta, 2004; Lunetta et al., 2007), and the use of specific tools could help address progress throughout the students' education.

Laboratory work and levels of inquiry

As already noted, our work reinforces the finding of several other reviews that guided inquiry is frequently successful as an instructional strategy (e.g., Furtak et al., 2012; Lazonder & Harmsen, 2016). However, through our work, this claim is nuanced by confirming the many challenges associated with the more student-centred approaches in terms of how guidance is employed, and its fidelity aspects. Put simply, there is always

a tradeoff between contextual constraints and choice of inquiry level in laboratory work. If, for example, time and other resources are limited, it will be difficult to maintain student-centred approaches as planned throughout the inquiry process, making them less effective than expected. Furthermore, we highlight the notion that the choice of inquiry level is dependent upon the learning goal of an activity. When the aim of the laboratory work is to demonstrate a certain scientific phenomenon or introduce students to a specific procedure, more teacher-centred approaches can be used. If the goal is to promote knowledge and skills, student engagement appears vital, which was previously also pointed out in a review by Dillon (2008).

Teaching according to open inquiry requires that students have a reasonably developed understanding of how to pose scientifically sound questions, handle laboratory equipment, manage complex data patterns – not always with positive outcomes – and interpretation bias, and document their work (see also, Chin & Osborne, 2008; Herranen & Aksela, 2019). Accordingly, open inquiry is a complex teaching approach that requires a great deal of time to enact and to develop, preferably using conscious feed-forward strategies that will enable the students' gradual engagement. Of course, it can be a challenge for teachers to take responsibility for design and data interpretation based on students' questions, but we believe that the alternative is worse.

Limitations of the study

The setting of inclusion criteria for a systematic review necessitates a tradeoff between the review's breadth and depth (Gough et al., 2017). As guided by the identified needs of teachers, our sample of included studies was limited to laboratory work where the students themselves collected the empirical data. Laboratory work appears to be an important strategy for secondary-school science teachers to allow students to explore scientific phenomena, as well as science practices. However, it is evident that many aspects of scientific inquiry can be addressed outside the context of students' own interaction with the physical world, e.g., by using observation reports such as available literature or databases, or by employing virtual laboratories. Moreover, the theories behind science methodologies can be addressed in the theoretical classroom instruction. Focusing solely on learning outcomes would probably have resulted in a somewhat different sample of studies and, in turn, additional nuances of science practices beyond the context of laboratory work as defined in this review.

Our mixed-methods approach, combining quantitative and qualitative primary studies, enabled a richer account and a contextual understanding of laboratory work as a learning strategy. Primary studies may be heterogeneous, also with respect to the theoretical perspectives that frame the research interest and analyses (Hong et al., 2020). To integrate such a diversity of evidence delimits the possibilities of making precise generalisations regarding expected learning outcomes in new contexts. Accordingly, our review findings are constructed based on realistic interpretation and familiarity with theory and practice. As guided by the CERQual analytical tool (Lewin et al., 2015), we are confident that our evidence claims are reliable and valid, and that the outlined practical implications are recognisable by secondary-school science teachers.

Regardless of whether quantitative or qualitative methodologies were used in the primary studies, we acknowledge the risk of dissemination or publication bias, i.e., that research showing empty or negative results may to a greater extent remain unpublished. The general problem of bias may also include any conscious or unconscious influence distorting individual study findings which cannot be detected through critical appraisal of the studies. As a consequence, there is a risk that the compilation of data on effectiveness is overrated, and that the understanding of phenomena may be distorted or incomplete (Thornton & Lee, 2000). To address such issues, our review takes into account inter-rater reliability as researchers performed the data-extraction process independently.

Conclusion

This systematic review provides new and important knowledge about laboratory work employed to promote students' learning of science, as well as their learning to do science and levels of inquiry. Our review findings are presented in a way that directly translates and recontextualizes science education research into teachers' practice. As we cast new light on several challenges announced in recent reviews, our findings are relevant not only for teachers but also for future studies in the field of science education research. Based on our findings, we would suggest that research should focus on the following areas:

- *The connection between level of inquiry and learning goals in laboratory work*. Often in the literature, open inquiry is propagated. Our review shows a more complex situation, and the level of openness needs to be adjusted to the intended learning goals. Here, additional studies are needed exploring what teacher guidance is required in relation to different learning goals, and how appropriate teacher guidance, corresponding to these goals, could be developed? So far, research indicates no single fix for all contexts exists.
- *The connection between science knowledge and skills within laboratory work*. Our review indicates a reciprocity between students' development of content knowledge and inquiry skills. There are no straight answers from research clarifying this connection, and more studies are therefore called for.
- *The issue of transfer and learning progression in laboratory work*. In studies of laboratory work, the specific lab investigated is often not problematised albeit the results are generalised. Our review shows that students' experiences and understanding of various laboratory work, of which experiments are part, differ, and the students have difficulties to transfer their understanding from one exercise to the next. Therefore, we need further studies on how and when to teach different kinds of laboratory work. Particular attention should be paid to the issue of progressively introducing control-of-variables in experiments.
- *The issue of assessment in laboratory work.* As pointed out in the background of this paper, reviews have over the last two decades repeatedly identified assessment of laboratory work as a major challenge for the teachers, and the studies included in our review do not contribute with solutions. Therefore, further studies investigating assessment practices of laboratory work, besides traditional lab reports, are called for.

We hope that fellow researchers will explore, develop and further investigate our understanding of these areas in future empirical research.

Notes

1. It should be noted that focus groups were arranged as part of a separate survey with the aim to identify practitioners' need for up-to-date research more generally.
2. Note that no studies reported in Scandinavian languages were included in the final sample.
3. Note that there is overlap between these two categories of learning goals.

Acknowledgments

Disclosure statement

The study is a further development of a report for teachers (The Swedish Institute for Educational Research, 2020). No potential conflict of interest was reported by the author(s).

Notes on contributors

Niklas Gericke is professor in science education and Director of the research centre SMEER (Science, Mathematics and Engineering Education Research) at Karlstad University, Universitetsgatan 2, SE-65188 Karlstad, Sweden; email: niklas.gericke@kau.se. He is also visiting professor at NTNU in Trondheim, Norway. His main research interests are biology education and sustainability education. Much of his work relates to how the disciplinary knowledge is transformed into school knowledge and what impact these transformations have on teachers' work and students' understanding.

Per Högström is a senior lecturer in science and technology education and researcher at Halmstad University, Kristian IV:s väg 3, SE-30250 Halmstad, Sweden; email: per.hogstrom@hh.se. His main research interests are science education and collaborative approaches to education. Much of his work relates to digitalisation and science practices and how roles, relations and learning are reframed within teacher education and in schools.

Johan Wallin is a senior researcher and programme manager of systematic reviews at the Swedish Institute for Educational Research, Hemvärnsgatan 15, SE-17121 Solna, Sweden; email: johan.wallin@skolfi.se. The Swedish Institute for Educational Research is a government agency tasked with providing practitioners and other school professionals with the best available evidence. Johan Wallin's main research interests are development of methods of systematic reviews in educational research and evidence-informed decision making in schools.

ORCID

Niklas Gericke 🄳 http://orcid.org/0000-0001-8735-2102
Per Högström 🄳 http://orcid.org/0000-0002-2407-9433

Author contributions

All three authors have made equal contributions to the conception of the ideas in this paper and have been involved in drafting or revising it critically for important intellectual content.

References

References marked with an asterisk (*) indicate studies included in the review.

*Abrahams, I., & Millar, R. (2008). Does practical work really work? A study of the effectiveness of practical work as a teaching and learning method in school science. *International Journal of Science Education*, *30*(14), 1945–1969. doi. https://doi.org/10.1080/09500690701749305

*Abrahams, I., & Reiss, M. J. (2012). Practical work: Its effectiveness in primary and secondary schools in England. *Journal of Research in Science Teaching*, *49*(8), 1035–1055. https://doi.org/10.1002/tea.21036

*Allen, M. (2011). Theory-led confirmation bias and experimental persona. *Research in Science & Technological Education*, *29*(1), 107–127. https://doi.org/10.1080/02635143.2010.539973

*Andersson, J., & Enghag, M. (2017). The relation between students' communicative moves during laboratory work in physics and outcomes of their actions. *International Journal of Science Education*, *39*(2), 158–180. https://doi.org/10.1080/09500693.2016.1270478

*Arnold, J. C., Kremer, K., & Mayer, J. (2014). Understanding students' experiments—What kind of support do they need in inquiry tasks? *International Journal of Science Education*, *36*(16), 2719–2749. https://doi.org/10.1080/09500693.2014.930209

*Blanchard, M. R., Southerland, S. A., Osborne, J. W., Sampson, V. D., Annetta, L. A., & Granger, E. M. (2010). Is inquiry possible in light of accountability?: A quantitative comparison of the relative effectiveness of guided inquiry and verification laboratory instruction. *Science Education*, *94*(4), 577–616. https://doi.org/10.1002/sce.20390

*Carter, G., Westbrook, S. L., & Thompkins, C. D. (1999). Examining science tools as mediators of students' learning about circuits. *Journal of Research in Science Teaching*, *36*(1), 89–105. https://doi.org/10.1002/(SICI)1098-2736(199901)36:1<89::AID-TEA6>3.0.CO;2-7

*Cheng, S. C., She, H. C., & Huang, L. Y. (2018). The Impact of problem-solving instruction on middle school students' physical science learning: Interplays of knowledge, reasoning, and problem solving. *Eurasia Journal of Mathematics, Science and Technology Education*, *14*(3), 731–743. https://doi.org/10.12973/ejmste/80902

*Chin, C., & Kayalvizhi, G. (2002). Posing problems for open investigations: What questions do pupils ask? *Research in Science & Technological Education*, *20*(2), 269–287. https://doi.org/10.1080/0263514022000030499

*Ding, N., & Harskamp, E. G. (2011). Collaboration and peer tutoring in chemistry laboratory education. *International Journal of Science Education*, *33*(6), 839–863. https://doi.org/10.1080/09500693.2010.498842

*Fadzil, H. M., & Saat, R. M. (2017). Exploring students' acquisition of manipulative skills during science practical work. *Eurasia Journal of Mathematics, Science and Technology Education*, *13*(8), 4591–4607. https://doi.org/10.12973/eurasia.2017.00953a

*Freedman, M. P. (1997). Relationship among laboratory instruction, attitude toward science, and achievement in science knowledge. *Journal of Research in Science Teaching*, *34*(4), 343–357. https://doi.org/10.1002/(SICI)1098-2736(199704)34:4<343::AID-TEA5>3.0.CO;2-R

*Garcia-Mila, M., Andersen, C., & Rojo, N. E. (2011). Elementary students' laboratory record keeping during scientific inquiry. *International Journal of Science Education*, *33*(7), 915–942. https://doi.org/10.1080/09500693.2010.480986

*Hamza, K. M., & Wickman, P.-O. (2013). Student engagement with artefacts and scientific ideas in a laboratory and a concept-mapping activity. *International Journal of Science Education*, *35*(13), 2254–2277. https://doi.org/10.1080/09500693.2012.743696

*Hand, B., Wallace, C. W., & Yang, E.-M. (2004). Using a Science Writing Heuristic to enhance learning outcomes from laboratory activities in seventh-grade science: Quantitative and qualitative aspects. *International Journal of Science Education*, *26*(2), 131–149. https://doi.org/10.1080/0950069032000070252

*Haslam, C. Y., & Hamilton, R. J. (2010). Investigating the use of integrated instructions to reduce the cognitive load associated with doing practical work in secondary school science. *International Journal of Science Education*, *32*(13), 1715–1737. https://doi.org/10.1080/09500690903183741

*Hodges, G. W., Wang, L. L., Lee, J., Cohen, A., & Jang, Y. (2018). An exploratory study of blending the virtual world and the laboratory experience in secondary chemistry classrooms. *Computers & Education*, *122*, 179–193. https://doi.org/10.1016/j.compedu.2018.03.003

*Hofstein, A., Shore, R., & Kipnis, M. (2004). Providing high school chemistry students with opportunities to develop learning skills in an inquiry-type laboratory: A case study. *International Journal of Science Education*, *26*(1), 47–62. https://doi.org/10.1080/0950069032000070342

*Hofstein, A., Navon, O., Kipnis, M., & Mamlok-Naaman, R. (2005). Developing students' ability to ask more and better questions resulting from inquiry-type chemistry laboratories. *Journal of Research in Science Teaching*, *42*(7), 791–806. https://doi.org/10.1002/tea.20072

*Högström, P., Ottander, C., & Benckert, S. (2010). Lab work and learning in secondary school chemistry: The importance of teacher and student interaction. *Research in Science Education*, *40*(4), 505–523. https://doi.org/10.1007/s11165-009-9131-3

*Kanari, Z., & Millar, R. (2004). Reasoning from data: How students collect and interpret data in science investigations. *Journal of Research in Science Teaching*, *41*(7), 748–769. https://doi.org/10.1002/tea.20020

*Katchevich, D., Hofstein, A., & Mamlok-Naaman, R. (2013). Argumentation in the chemistry laboratory: Inquiry and confirmatory experiments. *Research in Science Education*, *43*(1), 317–345. https://doi.org/10.1007/s11165-011-9267-9

*Kind, P. M., Kind, V., Hofstein, A., & Wilson, J. (2011). Peer argumentation in the school science laboratory–Exploring effects of task features. *International Journal of Science Education*, *33*(18), 2527–2558. https://doi.org/10.1080/09500693.2010.550952

*Kipnis, M., & Hofstein, A. (2008). The inquiry laboratory as a source for development of metacognitive skills. *International Journal of Science and Mathematics Education*, *6*(3), 601–627. https://doi.org/10.1007/s10763-007-9066-y

*Lazarowitz, R., & Naim, R. (2014). Learning the cell structures with three-dimensional models: Students' achievement by methods, type of school and questions' cognitive level. *Journal of Science Education and Technology*, *22*(4), 500–508. https://doi.org/10.1007/s10956-012-9409-5

*Lundin, M., & Lindahl, M. G. (2014). Negotiating the relevance of laboratory work: Safety, procedures and accuracy brought to the fore in science education. *NorDiNa: Nordic Studies in Science Education*, *10*(1), 32–45. https://doi.org/10.5617/nordina.638

*Marcum-Dietrich, N. I., & Ford, D. J. (2002). The place for the computer is in the laboratory: An investigation of the effect of computer probeware on student learning. *Journal of Computers in Mathematics & Science Teaching*, *21*(4), 361–379. https://www.learntechlib.org/primary/p/9527/

*McRobbie, C. J., & Thomas, G. P. (2000). Epistemological and contextual issues in the use of microcomputer-based laboratories in a year 11 chemistry classroom. *Journal of Computers in Mathematics & Science Teaching*, *19*(2), 137–160. https://www.learntechlib.org/primary/p/8067/

*Munn, M., Knuth, R., Van Horne, K., Shouse, A. W., & Levias, S. (2017). How do you like your science, wet or dry? How two lab experiences influence student understanding of science concepts and perceptions of authentic scientific practice. *CBE – Life Sciences Education*, *16*(2), 1–16. https://doi.org/10.1187/cbe.16-04-0158

*Peker, D., & Wallace, C. S. (2011). Characterizing high school students' written explanations in biology laboratories. *Research in Science Education*, *41*(2), 169–191. https://doi.org/10.1007/s11165-009-9151-z

*Roth, W.-M., McRobbie, C. J., Lucas, K. B., & Boutonné, S. (1997). The local production of order in traditional science laboratories: A phenomenological analysis. *Learning and Instruction*, *7*(2), 107–136. https://doi.org/10.1016/S0959-4752(96)00013-8

*Schmidt-Borcherding, F., Hänze, M., Wodzinski, R., & Rincke, K. (2013). Inquiring scaffolds in laboratory tasks: An instance of a "worked laboratory guide effect"? *European Journal of Psychology of Education*, *28*(4), 1381–1395. https://doi.org/10.1007/s10212-013-0171-8

*Schwichow, M., Zimmerman, C., Croker, S., & Härtig, H. (2016). What students learn from hands-on activities. *Journal of Research in Science Teaching*, *53*(7), 980–1002. https://doi.org/10.1002/tea.21320

*Seda Cetin, P., Eymur, G., Southerland, S. A., Walker, J., & Whittington, K. (2018). Exploring the effectiveness of engagement in a broad range of disciplinary practices on learning of Turkish high-school chemistry students. *International Journal of Science Education*, *40*(5), 473–497. https://doi.org/10.1080/09500693.2018.1432914

*Sesen, B. A., & Tarhan, L. (2013). Inquiry-based laboratory activities in electrochemistry: High school students' achievements and attitudes. *Research in Science Education*, *43*(1), 413–435. https://doi.org/10.1007/s11165-011-9275-9

*Strimaitis, A. M., Southerland, S. A., Sampson, V., Enderle, P., & Grooms, J. (2017). Promoting equitable biology lab instruction by engaging all students in a broad range of science practices: An exploratory study. *School Science and Mathematics*, *117*(3), 92–103. https://doi.org/10.1111/ssm.12212

*Tan, A.-L. (2008). Tensions in the biology laboratory: What are they? *International Journal of Science Education*, *30*(12), 1661–1676. https://doi.org/10.1080/09500690701564621

*Toplis, R. (2007). Evaluating science investigations at ages 14–16: Dealing with anomalous results. *International Journal of Science Education*, *29*(2), 127–150. https://doi.org/10.1080/09500690500498278

*Wolf, S. J., & Fraser, B. J. (2008). Learning environment, attitudes and achievement among middle-school science students using inquiry-based laboratory activities. *Research in Science Education*, *38*(3), 321–341. https://doi.org/10.1007/s11165-007-9052-y

Agustian, H., & Seery, M. (2017). Reasserting the role of pre-laboratory activities in university chemistry laboratories: A proposed framework for their design. *Chemistry Education Research and Practice*, *18*(4), 518–532 https://doi.org/10.1039/C7RP00140A

Akuma, F. V., & Callaghan, R. (2019). A systematic review characterizing and clarifying intrinsic teaching challenges linked to inquiry-based practical work. *Journal of Research in Science Teaching*, *56*(5), 619–648. https://doi.org/10.1002/tea.21516

Banchi, H., & Bell, R. (2008). The many levels of inquiry. *Science and Children*, *46*(2), 26–29. https://www.jstor.org/stable/43174976

Bernstein, B. (1996). *Pedagogy, symbolic control, and identity: Theory, research, critique*. Taylor & Francis.

Bybee, R. W. (2000). Teaching science as inquiry. In J. Minstrell & E. H. van Zee (Eds.), *Inquiring into inquiry learning and teaching in science* (pp. 1–27). American Association for the Advancement of Science.

Caglak, S. (2017). Does hands-on science practices make an impact on achievement in science? A meta-analysis. *Journal of Education in Science Environment and Health*, *3*(1), 69–87. https://doi.org/10.21891/jeseh.275708

Chin, C., & Osborne, J. (2008). Students' questions: A potential resource for teaching and learning science. *Studies in Science Education*, *44*(1), 1–39. https://doi.org/10.1080/03057260701828101

Crawford, B. (2014). From inquiry to scientific practices in the science classroom. In N. G. Lederman & S. K. Abell (Eds.), *Handbook of research on science education* (Vol. 2, pp. 515–541). Routledge.

Dillon, J. (2008). A review of the research on practical work in school science. King's College.

Dobber, M., Zwart, R., Tanis, M., & van Oers, B. (2017). Literature review: The role of the teacher in inquiry-based education. *Educational Research Review*, *22*, 194–214. https://doi.org/10.1016/j.edurev.2017.09.002

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

Gough, D., Oliver, S., & Thomas, J. (2017). *An introduction to systematic reviews*. SAGE.

Gough, D., Thomas, J., & Oliver, S. (2019). Clarifying differences between reviews within evidence ecosystems. *Systematic Reviews*, *8*(170), 1–15. https://doi.org/10.1186/s13643-019-1089-2

Halawa, S., Hsu, Y.-S., Zhang, W.-X., Kuo, Y.-R., & Wu, J.-Y. (2020). Features and trends of teaching strategies for scientific practices from a review of 2008–2017 articles. *International Journal of Science Education*, *42*(7), 1183–1206. https://doi.org/10.1080/09500693.2020.1752415

Heindl, M. (2019). Inquiry-based learning and the pre-requisite for its use in science at school: A meta-analysis. *Journal of Pedagogical Research*, *3*(2), 52–61. https://doi.org/10.33902/JPR.2019254160

Herranen, J., & Aksela, M. (2019). Student-question-based inquiry in science education. *Studies in Science Education*, *55*(1), 1–36. https://doi.org/10.1080/03057267.2019.1658059

Hodson, D. (1993). Re-thinking old ways: Towards a more critical approach to practical work in school science. *Studies in Science Education*, *22*(1), 85–142. https://doi.org/10.1080/03057269308560022

Hodson, D. (1996). Laboratory work as scientific method: Three decades of confusion and distortion. *Journal of Curriculum Studies*, *28*(2), 115–135. https://doi.org/10.1080/0022027980280201

Hodson, D. (2014). Learning science, learning about science, doing science: Different goals demand different learning methods. *International Journal of Science Education*, *36*(15), 2534–2553. https://doi.org/10.1080/09500693.2014.899722

Hofstein, A., & Lunetta, V. N. (2004). The laboratory in science education: Foundations for the twenty-first century. *Science Education*, *88*(1), 28–54. https://doi.org/10.1002/sce.10106

Hong, Q. N., Pluye, P., Bujold, M., & Wassef, M. (2017). Convergent and sequential synthesis designs: Implications for conducting and reporting systematic reviews of qualitative and quantitative evidence. *Systematic Reviews*, *6*(61). https://doi.org/10.1186/s13643-017-0454-2

Hong, Q. N., Rees, R., Sutcliffe, K., & Thomas, J. (2020). Variations of mixed methods reviews approaches: A case study. *Research Synthesis Methods*, *11*(6), 795–811. https://doi.org/10.1002/jrsm.1437

Lazonder, A. W., & Harmsen, R. (2016). Meta-analysis of inquiry-based learning: Effects of guidance. *Review of Educational Research*, *86*(3), 681–718. https://doi.org/10.3102/0034654315627366

Lederman, N. G., Antink, A., & Bartos, S. (2014). Nature of science, scientific inquiry, and socio-scientific issues arising from genetics: A pathway to developing a scientifically literate citizenry. *Science & Education*, *23*(2), 285–302. https://doi.org/10.1007/s11191-012-9503-3

Lederman, J., Lederman, N., Bartels, S., Jimenez, J., Akubo, M., Aly, S., Bao, C., Blanquet, E., Blonder, R., Bologna Soares de Andrade, M., Buntting, C., Cakir, M., EL-Deghaidy, H., ElZorkani, A., Gaigher, E., Guo, S., Hakanen, A., Hamed Al-Lal, S., Han-Tosunoglu, C., & Zhou, Q. (2019). An international collaborative investigation of beginning seventh grade students' understandings of scientific inquiry: Establishing a baseline. *Journal of Research in Science Teaching*, *56*(4), 486–515. https://doi.org/10.1002/tea.21512

Lewin, S., Glenton, C., Munthe-Kaas, H., Carlsen, B., Colvin, C. J., Gülmezoglu, M., Noyes, J., Booth, A., Garside, R., & Rashidian, A. (2015). Using qualitative evidence in decision making for health and social interventions: An approach to assess confidence in findings from qualitative evidence syntheses (GRADE-CERQual). *PLoS Medicine*, *12*(10), e1001895. https://doi.org/10.1371/journal.pmed.1001895

Lunetta, V. N., Hofstein, A., & Clough, M. P. (2007). Teaching and learning in the school science laboratory. An analysis of research, theory, and practice. In S. K. A. In & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 393–431). Erlbaum.

Mercer, N. (1995). *The guided construction of knowledge: Talk amongst teachers and learners*. Multilingual Matters Ltd.

Millar, R. (2004). The role of practical work in the teaching and learning of science. *High School Science Laboratories: Role and Vision]* ([Paper prepared for the committee). National Academy of Sciences.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, T. P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

National Research Council. (1996). *National science education standards*. The National Academies Press.

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. The National Academies Press. https://www.nap.edu/read/13165/chapter/1

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—A web and mobile app for systematic reviews. *Systematic Reviews*, *5*(1), 210. https://doi.org/10.1186/s13643-016-0384-4

Penuel, W. R., Riedy, R., Barber, M. S., Peurach, D. J., LeBoeuf, W. A., & Clark, T. (2020). Principles of collaborative education research with stakeholders: Toward requirements for a new research and development infrastructure. *Review of Educational Research*, *90*(5), 627–674. https://doi.org/10.3102/0034654320938126

Pollock, A., Campbell, P., Struthers, C., Synnot, A., Nunn, J., Hill, S., Goodare, H., Morris, J., Watts, C., & Morley, R. (2018). Stakeholder involvement in systematic reviews: A scoping review. *Systematic Reviews*, *7*(208), 1–26. https://doi.org/10.1186/s13643-018-0852-0

Puttick, G., Drayton, B., & Cohen, E. (2015). A study of the literature on lab-based instruction in biology. *The American Biology Teacher*, *77*(1), 12–18. https://doi.org/10.1525/abt.2015.77.1.3

Rudolph, J. L. (2019). *How we teach science: What's changed, and why it matters*. Harvard University Press.

Rudolph, J. L. (2020). The lost moral purpose of science education. *Science Education*, *104*(5), 895–906. https://doi.org/10.1002/sce.21590

The Swedish Institute for Educational Research. (2020). *Laborationer i naturvetenskapsundervisningen. Systematisk översikt 2020:01. [Laboratory work in science education. A systematic review]*. Skolforskningsinstitutet.

Swedish National Agency for Education. (2019). *Läroplan för grundskolan, förskoleklassen och fritidshemmet 2011 Reviderad 2019 [Curriculum for the compulsory school, preschool class and school-age educare 2011: Revised 2019]* (6th ed.). https://www.skolverket.se/getFile?file=4206

Thornton, A., & Lee, P. (2000). Publication bias in meta-analysis: Its causes and consequences. *Journal of Clinical Epidemiology*, *53*(2), 207–216. https://doi.org/10.1016/S0895-4356(99)00161-4

Weiss, C. H. (1979). The many meanings of research utilization. *Public Administration Review*, *39*(5), 426–431. https://doi.org/10.2307/3109916

Wellington, J. (Ed.). (1998). *Practical work in school science. which way now?* Routledge.