Educational Research Review 44 (2024) 100623

Contents lists available at ScienceDirect

Educational Research Review

journal homepage: www.elsevier.com/locate/edurev

Review

Beyond inquiry or direct instruction: Pressing issues for designing impactful science learning opportunities

Ton de Jong [a,*], Ard W. Lazonder [b], Clark A. Chinn [c], Frank Fischer [d], Janice Gobert [e], Cindy E. Hmelo-Silver [f], Ken R. Koedinger [g], Joseph S. Krajcik [h], Eleni A. Kyza [i], Marcia C. Linn [j], Margus Pedaste [k], Katharina Scheiter [l], Zacharias C. Zacharia [m]

[a] *Department of Instructional Technology, University of Twente, the Netherlands*
[b] *Behavioral Science Institute, Radboud University, the Netherlands*
[c] *Graduate School of Education, Rutgers University, USA*
[d] *Department of Psychology, Ludwig-Maximilians-Universität München, Germany*
[e] *Graduate School of Education, Rutgers University, USA & Apprendis LCC, USA*
[f] *Center for Research on Learning and Technology, Indiana University, USA*
[g] *Human-Computer Interaction Institute, Carnegie Mellon University, Pittsburgh, PA, USA*
[h] *College of Education, Michigan State University, East Lansing, MI, USA*
[i] *Media, Cognition and Learning Research Group, Department of Communication and Internet Studies, Cyprus University of Technology, Cyprus*
[j] *University of California, Berkeley Graduate School of Education, USA*
[k] *Institute of Education, University of Tartu, Estonia*
[l] *Department of Educational Science, University of Potsdam, Germany*
[m] *Research in Science and Technology Education Group, University of Cyprus, Cyprus*

ARTICLE INFO

Keywords:
Inquiry-based instruction
Direct instruction
Instructional design
Evidence-based instruction

ABSTRACT

We recently published a paper in this journal (de Jong et al., 2023) that presented an overview of the literature on learning in science domains through direct instruction and guided inquiry-based learning. This paper was, in part, a response to Zhang et al. (2022) who argued that the evidence firmly supported the superiority of direct instruction over inquiry learning. Sweller et al. (2024) recently replied by repeating this claim and also argued that we had ignored evidence against our position, questioned our analysis of the evidence, and claimed that direct instruction (unlike inquiry learning) is grounded in a strong theory. In this rebuttal we start by reemphasizing the conclusion from our previous paper: adequate instruction always involves different strategies, which should be thoughtfully selected based on contextual factors. Next, we demonstrate that inquiry-based learning is firmly rooted in both cognitive and socio-cultural theories of learning and conclude from recent literature that Sweller et al.'s belief that direct instruction is overall more effective than inquiry learning is not supported by the data from empirical studies.

In 1983, the American physicist and educator Fred Reif made the following compelling observation: "Physicists often claim that the great beauty of physics […] is that there is relatively little to remember. By contrast, many novice students complain that there is so much to remember. Both are probably right" (Reif, 1983, p. 47). This quote illustrates that the knowledge base of domain experts and

* Corresponding author.
E-mail address: a.j.m.dejong@utwente.nl (T. de Jong).

https://doi.org/10.1016/j.edurev.2024.100623
Received 15 April 2024; Received in revised form 23 May 2024; Accepted 23 July 2024
Available online 25 July 2024

novices is organized differently. In our view, the main task for education is to help students reach a level of conceptual understanding in which knowledge is abstracted and organized in a more expert-like way.

We recently published a paper in this journal (de Jong et al., 2023) that presented an overview of the literature on learning in science domains through direct instruction and guided inquiry-based learning. This paper was, in part, a response to Zhang et al. (2022) who argued that the evidence firmly supported the superiority of direct instruction over inquiry learning, considering three types of research—controlled studies, program-based studies, and correlational work. Based on a comprehensive review of the literature, we concluded that research using each of these three types of studies converged to consistently demonstrate that guided inquiry learning outperforms direct instruction for gaining advanced (more expert-like) conceptual knowledge. Nevertheless, recognizing the nuances and complexity the many pedagogical approaches to teaching science, we identified a set of potential factors that might influence the effectiveness of specific instructional methods and made a plea for investigating the conditions under which combinations of inquiry learning and direct instruction would be favorable for learning.

A response to our paper has recently been published in this journal (Sweller et al., 2024). This response repeated Zhang et al.'s (2022) earlier claims about what they believe the evidence shows. Sweller et al. (2024) also argued that we had ignored evidence against our position, questioned our analysis of the evidence, and claimed that one reason for favoring direct instruction over inquiry learning is that direct instruction (unlike inquiry learning) is grounded in a strong theory. The Editor of this journal invited us to write a rebuttal to their response.

We start this rebuttal by reemphasizing the conclusion from our previous paper: adequate instruction always involves different strategies, which should be thoughtfully selected based on contextual factors. This stance is related to the observation that student activity is needed to acquire knowledge—as is structure, support, and sometimes explicitness. Next, we attend to the theoretical undergirding of inquiry-based learning from both cognitive and socio-cultural perspectives. After that, we revisit the literature and, again, conclude that Sweller et al.'s belief that direct instruction is more effective overall than inquiry learning is not supported by the data from empirical studies. Following a response to several methodological issues raised by Sweller et al. we end with a concluding remark.

1. Moving from a horse race to what is best for students

Although our debate with Sweller et al. might seem to be about which of the two instructional methods is best, we do not think this is the main issue. Our intention has at least been to find out which instructional events lead to productive learning under what circumstances, to include these events in instructional material, and to give teachers the tools to implement them in the classroom. Bearing this goal in mind, we would like to reiterate that both direct instruction and inquiry learning can play a role in effective instructional methods, depending on the topic, the instructional goal, the student, the teacher, the instructional design, and similar situational characteristics (see also, for example, Clements et al., 2023).

In our previous paper (de Jong et al., 2023), we argued for the importance of determining under which conditions direct instruction and guided inquiry can be effective. In this rebuttal, we emphasize further that most real learning environments involve a complex mixture of direct instruction and student activity, including exploration and inquiry. We believe that the only reasonable position in this debate is to acknowledge that there will be a complex mixture of telling students things and affording them opportunities to invent things themselves. But whatever specific procedures are involved, it will often be valuable for students—even novices—to engage in exploration and invention. Indeed, we would argue that if one carefully scrutinizes guided-inquiry curricula in science, such as the Open Science curricula[1] that have been designed to meet the Next Generation Science Standards in the USA, one finds that these curricula contain a rich mixture of telling students certain things and having them figure out other things. Other examples are project-based learning programs that have a clear inquiry kernel but also include other instructional methods (Krajcik et al., 2022; Schneider et al., 2022).

This orchestration of instructional methods should also be reflected in the role that professional teachers play. As Hattie (2009, p. 261) concluded after revisiting many instructional methods, "… experienced experts possess pedagogical content knowledge that is more flexibly and innovatively employed in instruction; they are more able to improvise and to alter instruction in response to contextual features of the classroom situation ….". This quotation emphasizes that teachers are crucial for successfully implementing instructional strategies in the classroom, even if delivered in a (scaffolded) digital way (see also, for example, Dobber et al., 2017). We agree that a thorough training of teachers in understanding both the theoretical background of inquiry-based learning and the mechanisms to implement these theoretical notions in their classrooms is conditional to its success (Strat et al., 2023). Teachers need to be well-prepared for inquiry learning to implement this instructional method to enhance student learning. We agree that also policy documents should emphasize this nuance.

We acknowledge that the distinction between direct instruction and guided inquiry-based learning is not always straightforward because student activity plays a role in both instructional methods. Sweller et al. (2024) state that the Zhang et al. (2022) paper included no definition of direct instruction. Although it may seem ironic that such an explicit definition was missing in a paper on direct instruction, it left us with no option but to infer what Zhang et al. meant by direct instruction on the basis of the examples they provided. This seemed to entail a rather narrow view that direct instruction is simply telling students about the domain or demonstrating content. In our 2023 paper, we emphasized that direct instruction, following the original approach by Engelmann (1980), can

[1] See https://www.openscied.org/.

have a very active learning component. To us, the key difference between the two methods, as indicated above, is that in direct instruction, there is a full explanation before students start their investigations, whereas in inquiry-based learning, students have to generate (part) of the content themselves (Chinn & Duncan, 2021). Two very nicely worked out and concrete examples of this distinction can be found in a paper by Schuster et al. (2018). One of these examples concerns the physics topic of Newton's second law of motion. In the direct instruction condition, the initiative is with the instructor presenting Newton's law to the students by explaining the relations between the elements of the law both conceptually and in a formula. Students can ask questions to the instructor and the instructor can pose questions to the students to check their understanding. After that, students test and verify Newton's law in a practical situation using a wheeled skateboard. In the inquiry condition, students start by considering the question which type of motions may happen after imposing a force on an object. Directly after that students conduct experiments with the skateboard accompanied by instructor guidance. It is important to note that both approaches in this example include an active component for the students. For a full description of this example, and an additional one on the physics topic of density, see Schuster et al. (2018, pp. 393–394).

2. Theoretical underpinnings of inquiry-based learning

A central argument in Sweller et al.'s reaction is that we did not present a theoretical basis for inquiry-based instruction, whereas direct instruction is solidly rooted in cognitive load theory (CLT). This disparity, they argue, entails that their direct instruction position is superior due to its grounding in theory. However, it is incorrect that inquiry-based instruction lacks theoretical grounding. Both cognitive and sociocultural theories have been developed to explain the successes of inquiry-based learning, as we detail below.

Cognitive theories explaining the success of inquiry-based learning emphasize the active integration of knowledge (Linn & Eylon, 2011). This involves emphases pertinent to CLT including supporting learners to distinguish among ideas (Linn et al., 2023), engage in 'generative learning activities' (Fiorella & Mayer, 2016; Mayer, 2024), and undertake schema (re)construction (Rumelhart, 1980). For example, Fiorella and Mayer (2022, p. 339) elaborate: "Generative learning involves actively making sense of the learning material by engaging in activities for organizing the material and integrating it with one's existing knowledge". This process of generative knowledge integration involves recognizing when an approach does not function adequately including identifying a 'cognitive conflict' (Limón, 2001). This features a process of reconstruction, for example, when students see an experimental result that is not in line with their existing knowledge and draws on generative learning activities such as summarizing, explaining, comparing, predicting, visualizing, organizing, integrating, and inferencing (Brod, 2021; Fiorella, 2023). To succeed, generative learning activities benefit from respecting the many partial, incomplete, or experiential perspectives the student brings to the lesson; guiding students to discover additional perspectives; promoting collaborative efforts to distinguish among these perspective using evidence; and encouraging students to sort out their perspectives into a coherent account of the phenomenon (Linn et al., 2023).

Eysink et al. (2009) provided a cognitive explanation for the advantages of inquiry-based learning. Eysink et al. compared four different instructional methods in terms of learning effectiveness: hypermedia learning (Tell me how it works), observational learning (Show me how it works), inquiry learning (Let me investigate how it works), and explanation-based learning (Let me explain how it works). The domain was probability theory, and a total of 613 secondary education students participated in this study. All four instructional methods were computer-delivered to rule out possibly confounding teacher effects. Results on a posttest measuring different types of knowledge (corrected for pretest scores) showed that students in the explanation-based condition outperformed those in all other three conditions. However, students in the inquiry-based learning condition performed better than those in the hypermedia and observational learning conditions, and outperformed students from all other conditions on far transfer problems. These results indicate that, overall, the two methods that have students generate content themselves were more successful than the two conditions in which direct instruction of all content was used. Students in the hypermedia condition, though, learned less but did learn more efficiently. A follow-up study (Eysink & de Jong, 2012) using the same four learning environments had students ($N = 40$) think aloud while learning. The analysis of the think-aloud data produced two noteworthy findings. First, students in the direct instruction conditions (hypermedia and observational learning) were far less active in terms of learning processes than students in the two generative conditions. Second, in the latter two conditions the relative number of elaborative processes was significantly higher. This means that students in the generative conditions were more involved in schema construction activities (relating and integrating) than students in the two direct instruction conditions.

In addition to this cognitive perspective, inquiry-based learning has also been grounded in sociocultural theory dating back to the ideas of Rousseau (see for example, Dimopoulou & Gasparatou, 2024), Dewey and Vygotsky (see for example, Glassman, 2001). In contrast to the cognitive perspective, sociocultural perspectives assume that knowledge is developed through participation in social, cultural, and historical practices (Danish & Gresalfi, 2018). A sociocultural perspective on learning recognizes the importance of providing learners with opportunities to engage in authentic, meaningful, and collaborative experiences (Lave & Wenger, 1991; Rogoff, 1994; Vygotsky, 1978). Learning is conceptualized as changes in patterns of participation in social activity; thus, engagement in these activities is essential for learning to occur and be demonstrated. This perspective assumes that learners are active agents who construct their knowledge and understanding through engaging in meaningful tasks (e.g., problems, driving questions) in which they are investigating phenomena by exploring, questioning, using different sources of evidence, and reflecting on their activity (Dewey, 1938; Hmelo-Silver, 2004; Piaget, 1952). The social context provides opportunities for sharing ideas, discussing, debating, and collaborating with peers, teachers, and other sources of information, and feedback (Collins & Kapur, 2014). Moreover, engaging in meaningful tasks in inquiry-based learning supports student motivation and engagement (Renninger et al., 2018). An important aspect of inquiry-based learning is the scaffolding provided along with the various tools and artifacts that mediate learning and activity (Collins & Kapur, 2014; Gobert et al., 2023; Hmelo-Silver et al., 2007; Zacharia et al., 2015).

All in all, we conclude that inquiry-based learning is rooted in solid theories that explain learning and, hence, guide the design of instructional interventions. That said, the cognitive efforts of the students learning through inquiry-based instruction should, of course, stay within the limits of their cognitive capacities (see also, Fiorella, 2023). This is nicely illustrated by a study by Perez et al. (2017) that studied inquiry-based learning with simulations of electrical circuits. In this study, low-prior-knowledge students who were more successful in learning chose to explore simple circuits first, whereas this pattern was not observed in similar students who did not learn successfully. Also, the successful students introduced pauses in their learning, which was not observed in the less successful learners. Thus, students who were successful managed their cognitive load successfully. This study illustrates that inquiry-based learning and CLT are in no way opposed and can also be reconciled without having to fall back on direct instruction.

Sweller et al. (2024) argued that CLT provides the grounding for direct instruction, but constraints on working memory do not entail that direct instruction is the only instructional road to follow. CLT only indicates that, for learning to be productive, students' experienced cognitive load should stay within their cognitive capacities. This condition can be met in guided inquiry-based learning as well as in direct instruction. One approach to doing so is to examine how scaffolds can help manage cognitive load, as inquiry-based theorists have done. A meta-analysis by Belland et al. (2017) found a moderate effect of scaffolding on conceptual learning outcomes with STEM pedagogical interventions other than direct instruction. A recent meta-analysis by Dai et al. (2024) suggests that artificial intelligence can play a role in making scaffolding more effective. This is also why we emphasized that scaffolds are important in inquiry-based learning and why we perceive a need for additional research investigating how inquiry-based learning and direct instruction can be combined to promote learning.

Pointing to CLT as the sole theory that helps explain phenomena is too restrictive, and there is more and other theoretical work, as we have shown above, that is useful. Given the strong preponderance of evidence for the efficacy of guided inquiry over direct instruction in many circumstances, it follows that CLT theorists should be assiduously working out how to revise their theories to explain these effects instead of denying that the effects exist. The work of inquiry-learning theorists on scaffolding suggests one likely viable direction for such research.

3. Revisiting the literature

Sweller et al. (2024) asserted that the meta-analyses we presented in our 2023 review solely included program-based studies and, hence, that we ignored the results of randomized control trials and correlational studies. This critique simply does not hold. The conclusions in our previous paper (de Jong et al., 2023) were based on a comprehensive overview of the literature that included program-based studies, controlled studies, and correlational studies as these were the categories distinguished by Zhang et al. (2022). Importantly, we did not base our conclusions on program-based studies alone as is implied by (Sweller et al., 2024). Rather, we sought to highlight that this particular type of study may provide valuable insights, in contrast to what was stated by Zhang et al. (2022). In addition, program-based studies may very well adhere to the standards of randomized controlled studies (see e.g., Krajcik et al., 2022; Schneider et al., 2022). Overall, program-based and controlled studies alike demonstrate a general advantage for guided inquiry-based learning over direct instruction for conceptual learning. Correlational studies likewise point to a positive relation between student performance and (guided) inquiry-based instruction.

When focusing on controlled studies, some meta-analyses distinguished levels of control of the included studies, and the most rigorously controlled studies certainly count as experimentally-controlled comparisons. A synthesis of these well-controlled studies pointed to clear advantages of inquiry-based instruction over direct instruction. For example, 42 of the studies in the review by Minner et al. (2010) involved comparisons between conditions with various degrees of inquiry saturation. Of these, 55% showed a significant advantage of inquiry-based learning, and only 1 (2%) showed an advantage for a less inquiry-focused condition. And Furtak et al. (2012) showed in an analysis of 37 experimental and quasi-experimental studies that there was an overall positive effect of inquiry-based methods over more direct-instruction-based methods. We reported all of these results in our 2023 paper.

In light of this substantial body of evidence, we fail to see how the limited set of controlled studies cited by Zhang et al. (2022) provides clear-cut support for their position. First, several controlled studies cited by Zhang et al. (2022) compared direct instruction to *unguided or minimally guided* inquiry, in which the inquiry is neither scaffolded nor guided (these include Klahr & Nigam, 2004; Matlen & Klahr, 2013). But few, if any, contemporary proponents of inquiry learning advocate *unguided* inquiry, which is widely recognized to be ineffective (Hmelo-Silver et al., 2007). Thus, these studies say nothing about what policy should be adopted by science educators apart from not using unguided inquiry.

Furthermore, not all papers Zhang et al. (2022) cited as evidence for direct instruction support these authors' position. For example, Zhang et al. (2022) cited Mayer (2004) to support their claim about the superiority of "direct, explicit instruction when comparing it with exploration-based learning" (p. 1165), but Mayer (2004) discussed many studies showing that guided inquiry is superior to unguided inquiry or pure discovery. Mayer also highlighted two studies that directly pitted guided inquiry against direct instruction; in each of these comparisons, students in a guided inquiry condition performed better. Indeed, Mayer (2004) provided a cognitive explanation for the benefits of guided inquiry or guided discovery: "Guided discovery is effective because it helps students meet two important criteria for active learning—(a) activating or constructing appropriate knowledge to be used for making sense of new incoming information and (b) integrating new incoming information with an appropriate knowledge base." (p. 15).

All in all, we think that the research literature favors the use of guided inquiry-based learning over direct instruction for acquiring deep conceptual knowledge. Exemplary is the work by Alfieri et al. (2011), already reported in our 2023 paper, who performed a meta-analysis of 164 studies and found that assisted inquiry is more effective than explicit instruction involving "explicit teaching of strategies, procedures, concepts, or rules in the form of formal lectures, models, demonstrations, and so forth and/or structured problem solving" (p. 5). This meta-analysis included numerous controlled studies of the exact sort that Zhang et al. (2022) prefer.

And the evidence keeps accruing. Skene et al. (2022) presented a meta-analysis of the use of guided play in early childhood education. These authors concluded that guided play is more effective than direct instruction for promoting executive functioning and numeracy (early math skills and shape knowledge), with effect sizes of $g = 0.40$, $g = 0.24$, and $g = 0.63$. The two methods did not differ in learning outcomes for literacy or socioemotional outcomes. For program-based studies, two recent large scale evaluations of project-based learning programs showed an advantage of this method over traditional ways of teaching. Schneider et al. (2022) compared the performance of students following project-based learning programs on chemistry and physics with students following traditional instruction. The project-based learning programs centered around real-world questions (such as "Why is table salt safe to eat but the substances that forms it are explosive or toxic when separated?") and students performed experiments, created models, explained and predicted phenomena etc. This randomized-control evaluation involving more than 4000 students, showed an advantage of the project-based condition over traditional teaching of close to 0.2 standard deviation on a summative science assessment developed by an independent third party. Krajcik et al. (2022) compared the achievements of students learning in project-based arrangements on physics and biology topics with students following 'business as usual' courses. This study involved more than 2000 students and used a randomized control set-up. Again, students in the project-based conditions outperformed the students in the traditional condition, now with a difference of a little over 0.2 standard deviation on a summative science assessment which was developed by an independent third party. Regarding correlational studies, Wan et al. (in press) re-analyzed the data of PISA 2015 from eight Western and East Asian regions and found an overall positive relationship between inquiry-based learning (as experienced by students) and science achievement, although in their analysis some aspects of inquiry (in particular, investigation) had a negative contribution. Teig (2024) combined the Norwegian data from PISA 2015 with log files from field trials, and reported that "… students who were exposed to particular inquiry-based teaching and learning practices, such as engaging in a class debate about investigations, were more likely to be in a profile with higher inquiry performance" (p. 221). It also became clear from the log file data that students definitely need guidance for specific inquiry activities such as applying the CVS (Control of Variables Strategy) in multivariable situations, coordinating theory and evidence, and constructing scientific explanations. Similar work by Gobert et al. (2024b) used AI algorithms on log file data to assess a wide range of inquiry competencies and support students' learning of these in real time via a digital agent. Also, a dashboard, again driven by AI-based analyses of student log files can be used to guide teachers' inquiry instruction (Gerard et al., 2020; Gobert et al., 2024a; Gobert et al., 2023; Wiley et al., 2023). This work gives very detailed information to inform the design of (digital) guidance in inquiry learning. In short, these recent studies corroborate our earlier analysis and at the same time underscore our plea for more research into the design of adequate instructional guidance (see also, Zacharia et al., 2015).

4. Using meta-analyses and program-based studies

The previous section showed that we base most of our conclusions on meta-analyses. Sweller et al. (2024) discuss several general objections to this type of research, which seem to resonate with some of the arguments by Renkl (2022) against the straightforward use of meta-analyses by teachers. Renkl (2022) argued for combining the results of meta-analyses with insights from learning theories to shape instruction in the actual classroom. We do agree with Renkl (2022) that meta-analytical findings may not readily apply in the complex classroom situation teachers face in daily practice and that contextual factors have a determinate influence on the effectiveness of instructional methods. The challenge of generalizability of controlled studies, often conducted in laboratory settings, to complex classrooms also applies to the controlled studies—often conducted in laboratory settings—valued by Zhang et al. (2022) and Sweller et al. (2024). This is why we argue for more research on the role of context in direct instruction and inquiry-based learning.

We also agree with Renkl (2022) that theories or general principles of learning can be an important information source for teachers, and inquiry-based learning is firmly rooted in such theoretical notions (see Section 2 of this paper). Still, meta-analyses do provide robust scientific evidence and protect researchers and practitioners from drawing conclusions based on single studies with potentially confounding features. Meta-analyses are syntheses of evidence taken from a large number of studies that fulfil a number of pre-specified inclusion criteria, in many cases attend to publication bias[2] and account for heterogeneity in results by including moderator analyses. Meta-analytical findings thus provide a closer estimation of true effect sizes than single studies do. It is unclear why Sweller et al. (2024) dismiss them and prefer to base their conclusions on individual studies for which the criteria for selection remain opaque. Moreover, meta-analytical methodologies have been further developed in recent years and, beyond providing an overall average effect only, often entail important information on the contextual conditions under which effects are more or less likely to occur based on moderator analyses.

Instead of systematically reviewing the literature, Zhang et al. (2022) seem to have based their conclusions on a selected series of studies, which provides less robust evidence. How do readers of their paper know that the studies have not been cherry-picked? The only alternative to using existing meta-analyses is to conduct a new one, with transparent a priori guidelines for including all relevant studies. In the absence of such a comprehensive review of studies, any set of papers used as evidence carries the risk of biased selection. As an example of such bias, Schuster et al. (2018) presented a study that they characterized as "randomized controlled" (p. 389); this study (see also above) showed that direct instruction and inquiry-based learning led to the same results. Yet this study was not mentioned in the Zhang et al. (2022) paper, despite the fact that one of the co-authors of this paper (Cobern) was also co-author of this study. The existing meta-analyses are the clearest source of available evidence, and they show that guided inquiry-based learning

[2] The meta-analyses cited in de Jong et al. (2023) showed no signs of publication bias. Orwin's fail-safe-*N* indicated that between 58 and 1119 additional studies with a mean effect size of zero would be needed to reduce the aggregate effect size to the trivial level of 0.10. This corresponds to an extension of the number of included studies by a factor 2.00 to 8.33.

yields better learning outcomes than direct instruction. Citing a few studies with different results does not accurately reflect the overall body of evidence.

Sweller et al. (2024) also objected again against using program-based studies. They cited a paper by Merrett (2006) to argue that program-based studies may suffer from a "Hawthorne effect." This citation is not in their reference list, but assuming it is Merrett (2006), this short 'research note' merely explained the Hawthorne effect and indicated that we should be attentive to it—and that this effect applies to randomized controlled studies as well as to other types of research. There is, however, older and more solid work (e.g., Adair et al., 1989; Cook, 1967) that failed to find any evidence of the Hawthorne effect in educational research. Besides, there is no reason why a possible Hawthorne effect would affect program-based studies more than controlled laboratory studies.

5. Concluding remarks

In 2008, Fred Reif wrote a seminal work on how to use cognitive science for the design of instruction that helps students acquire deep knowledge instead of superficial knowledge (Reif, 2008). We add to that by considering the sociocultural grounding for inquiry learning. Inquiry-based instruction that helps students to discover and invent, restructure their knowledge, and build a well-organized knowledge base, therefore, is a solid basis for designing instruction. Taking this as a starting point, we also tried to move away from a discussion of the superiority of instructional methods and focus on possible ways for these methods to reinforce each other. We also sought to better understand how and under which conditions different orchestrations of instructional methods and strategies support students' learning of disciplinary knowledge, skills, and practices. We explicitly presented our ideas as avenues for future research that would help advance the field of instructional design.

Declaration of competing interest

- No funding was received to assist with the preparation of this manuscript.
- The authors have no relevant financial or non-financial interests to disclose.
- The authors have no competing interests to declare that are relevant to the content of this article.
- Janice Gobert is CEO of Apprendis LLC that commercializes Inq-ITS that is mentioned in the article.
- All other authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Data availability

No data was used for the research described in the article.

References

Adair, J. G., Sharpe, D., & Huynh, C.-L. (1989). Hawthorne control procedures in educational experiments: A reconsideration of their use and effectiveness. *Review of Educational Research, 59*(2), 215–228. https://doi.org/10.3102/00346543059002215

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

Belland, B. R., Walker, A. E., Kim, N. J., & Lefler, M. (2017). Synthesizing results from empirical research on computer-based scaffolding in STEM education. *Review of Educational Research, 87*(2), 309–344. https://doi.org/10.3102/0034654316670999

Brod, G. (2021). Generative learning: Which strategies for what age? *Educational Psychology Review, 33*(4), 1295–1318. https://doi.org/10.1007/s10648-020-09571-9

Chinn, C. A., & Duncan, R. G. (2021). Inquiry and learning. In R. G. Duncan, & C. A. Chinn (Eds.), *International handbook of inquiry and learning* (pp. 1–14). Routledge.

Clements, D. H., Lizcano, R., & Sarama, J. (2023). Research and pedagogies for early math. *Education Sciences, 13*(8), 839. https://doi.org/10.3390/educsci13080839

Collins, A., & Kapur, M. (2014). Cognitive apprenticeship. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (2 ed., pp. 109–127). Cambridge University Press. https://doi.org/10.1017/CBO9781139519526.008.

Cook, D. L. (1967). *The impact of the Hawthorne effect in experimental design in educational research.* Columbus: US Department of Health. *Education and Welfare*, Research OoEBo.

Dai, C.-P., Ke, F., Pan, Y., Moon, J., & Liu, Z. (2024). Effects of artificial intelligence-powered virtual agents on learning outcomes in computer-based simulations: A meta-analysis. *Educational Psychology Review, 36*(1), 31. https://doi.org/10.1007/s10648-024-09855-4

Danish, J. A., & Gresalfi, M. (2018). Cognitive and social perspectives on learning. In F. Fischer, C. E. Hmelo-Silver, & S. R. Goldman (Eds.), *International handbook of the learning sciences* (pp. 34–43). https://doi.org/10.4324/9781315617572

de Jong, T., Lazonder, A. W., Chinn, C. A., Fischer, F., Gobert, J., Hmelo-Silver, C. E., Koedinger, K. R., Krajcik, J. S., Kyza, E. A., Linn, M. C., Pedaste, M., Scheiter, K., & Zacharia, Z. C. (2023). Let's talk evidence – the case for combining inquiry-based and direct instruction. *Educational Research Review, 39*, Article 100536. https://doi.org/10.1016/j.edurev.2023.100536

Dewey, J. (1938). *Experience and education.* Macmillan Company.

Dimopoulou, G., & Gasparatou, R. (2024). Emile's inquiry-based science education. *Journal of Philosophy of Education, 58*, 58–71. https://doi.org/10.1093/jopedu/qhae012

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

Engelmann, S. (1980). *Direct instruction educational technology.* Publications.

Eysink, T. H. S., & de Jong, T. (2012). Does instructional approach matter? How elaboration plays a crucial role in multimedia learning. *The Journal of the Learning Sciences, 21*(4), 583–625. https://doi.org/10.1080/10508406.2011.611776

Eysink, T. H. S., de Jong, T., Berthold, K., Kolloffel, B., Opfermann, M., & Wouters, P. (2009). Learner performance in multimedia learning arrangements: An analysis across instructional approaches. *American Educational Research Journal, 46*(4), 1107–1149. https://doi.org/10.3102/0002831209340235

Fiorella, L. (2023). Making sense of generative learning. *Educational Psychology Review, 35*(2), 50. https://doi.org/10.1007/s10648-023-09769-7

Fiorella, L., & Mayer, R. E. (2016). Eight ways to promote generative learning. *Educational Psychology Review, 28*(4), 717–741. https://doi.org/10.1007/s10648-015-9348-9
Fiorella, L., & Mayer, R. E. (2022). The generative activity principle in multimedia learning. In R. E. Mayer, & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd ed., pp. 339–350). Cambridge University Press.
Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching. *Review of Educational Research, 82*(3), 300–329. https://doi.org/10.3102/0034654312457206
Gerard, L., Wiley, K., Bradford, A., Chen, J. K., Lim-Breitbart, J., & Linn, M. C. (2020). Impact of a teacher action planner that captures student ideas on teacher customization decisions. In M. Gresalfi, & I. S. Horn (Eds.), *The interdisciplinarity of the learning sciences* (Vol. 4, pp. 2077–2084). International Society of the Learning Sciences.
Glassman, M. (2001). Dewey and Vygotsky: Society, experience, and inquiry in educational practice. *Educational Researcher, 30*(4), 3–14. https://doi.org/10.3102/0013189x030004003
Gobert, J. D., Dickler, R., & Adair, A. (2024a). Using an AI-based dashboard to help teachers support students' learning progressions for science practices. In H. Jin, D. Yan, & J. J. Krajcik (Eds.), *Handbook of research on science learning progression*. Taylor & Francis Group.
Gobert, J. D., Li, H., Dickler, R., & Lott, C. (2024b). Can AI-based scaffolding promote students' robust learning of authentic science practices? In X. Zhai, & J. Krajcik (Eds.), *Uses of artificial intelligence in STEM education*. Oxford University Press.
Gobert, J. D., Sao Pedro, M. A., & Betts, C. G. (2023). An AI-based teacher dashboard to support students' inquiry: Design principles, features, and technological specifications. In N. G. Lederman, D. L. Zeidler, & J. S. Lederman (Eds.), *Handbook of research on science education* (Vol. 3, pp. 1011–1044). Routledge. https://doi.org/10.4324/9780367855758.
Hattie, J. A. C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.
Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? *Educational Psychology Review, 16*(3), 235–266. https://doi.org/10.1023/B:EDPR.0000034022.16470.f3
Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to kirschner, sweller, and clark (2006). *Educational Psychologist, 42*(2), 99–107. https://doi.org/10.1080/00461520701263368
Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. *Psychological Science, 15*(10), 661–668. https://doi.org/10.1111/j.0956-7976.2004.00737.x
Krajcik, J., Schneidera, B., Miller, E., Chena, I.-C., Bradforda, L., Bartza, K., Bakera, Q., Palinscar, A., Peek-Browna, D., & Coderea, S. (2022). *Assessing the effect of project-based learning on science learning in elementary schools*. Michigan State University.
Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.
Limón, M. (2001). On the cognitive conflict as an instructional strategy for conceptual change: A critical appraisal. *Learning and Instruction, 11*(4–5), 357–380. https://doi.org/10.1016/S0959-4752(00)00037-2
Linn, M. C., Donnelly-Hermosillo, D., & Gerard, L. (2023). Synergies between learning technologies and learning sciences: Promoting equitable secondary school teaching. In N. G. Lederman, D. L. Zeidler, & J. S. Lederman (Eds.), *Handbook of research on science education* (pp. 447–498). Routledge. https://doi.org/10.4324/9780367855758.
Linn, M. C., & Eylon, B. S. (2011). *Science learning and instruction: Taking advantage of Technology to promote knowledge integration*. Routledge.
Matlen, B. J., & Klahr, D. (2013). Sequential effects of high and low instructional guidance on children's acquisition of experimentation skills: Is it all in the timing? *Instructional Science, 41*(3), 621–634. https://doi.org/10.1007/s11251-012-9248-z
Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? *American Psychologist, 59*(1), 14–19. https://doi.org/10.1037/0003-066X.59.1.14
Mayer, R. E. (2024). The past, present, and future of the cognitive theory of multimedia learning. *Educational Psychology Review, 36*(1), 8. https://doi.org/10.1007/s10648-023-09842-1
Merrett, F. (2006). Reflections on the Hawthorne effect. *Educational Psychology, 26*(1), 143–146. https://doi.org/10.1080/01443410500341080
Minner, D. D., Levy, A. J., & Century, J. (2010). Inquiry-based science instruction - what is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of Research in Science Teaching, 47*(4), 474–496. https://doi.org/10.1002/tea.20347
Perez, S., Massey-Allard, J., Butler, D., Ives, J., Bonn, D., Yee, N., & Roll, I. (2017). Identifying productive inquiry in virtual labs using sequence mining. In E. André, R. Baker, X. Hu, M. M. T. Rodrigo, & B. du Boulay (Eds.), *Artificial intelligence in education: 18th international conference, AIED 2017, wuhan, China, june 28 – july 1, 2017, proceedings* (pp. 287–298). Springer International Publishing. https://doi.org/10.1007/978-3-319-61425-0_24.
Piaget, J. (1952). *The origins of intelligence in children*. International Universities Pres.
Reif, F. (1983). Understanding and teaching problem solving in physics. In *Research on physics education: Proceedings of the first international workshop* (pp. 15–53). Centre National de la Recherche Scientifique.
Reif, F. (2008). *Applying cognitive science to education: Thinking and learning in scientific and other complex domains*. MIT press.
Renkl, A. (2022). Meta-analyses as a privileged information source for informing teachers' practice? *Zeitschrift für Padagogische Psychologie, 36*(4), 217–231. https://doi.org/10.1024/1010-0652/a000345
Renninger, K. A., Ren, Y., & Kern, H. M. (2018). Motivation, engagement, and interest. In F. Fischer, C. E. Hmelo-Silver, & S. R. Goldman (Eds.), *International handbook of the learning sciences* (pp. 116–127). https://doi.org/10.4324/9781315617572
Rogoff, B. (1994). Developing understanding of the idea of communities of learners. *Mind, Culture and Activity, 1*(4), 209–229. https://doi.org/10.1080/10749039409524673
Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension: Perspectives from cognitive psychology, linguistics, artificial intelligence, and education* (pp. 38–58). Lawrence Erlbaum Associates.
Schneider, B., Krajcik, J., Lavonen, J., Salmela-Aro, K., Klager, C., Bradford, L., Chen, I., Baker, Q., Touitou, I., Peek-Brown, D., Dezendorf, R. M., Maestrales, S., & Bartz, K. (2022). Improving science achievement—is it possible? Evaluating the efficacy of a high school chemistry and physics project-based learning intervention. *Educational Researcher, 51*(2), 109–121. https://doi.org/10.3102/0013189x211067742
Schuster, D., Cobern, W. W., Adams, B. A. J., Undreiu, A., & Pleasants, B. (2018). Learning of core disciplinary ideas: Efficacy comparison of two contrasting modes of science instruction. *Research in Science Education, 48*(2), 389–435. https://doi.org/10.1007/s11165-016-9573-3
Skene, K., O'Farrelly, C. M., Byrne, E. M., Kirby, N., Stevens, E. C., & Ramchandani, P. G. (2022). Can guidance during play enhance children's learning and development in educational contexts? A systematic review and meta-analysis. *Child Development, 93*(4), 1162–1180. https://doi.org/10.1111/cdev.13730
Strat, T. T. S., Henriksen, E. K., & Jegstad, K. M. (2023). Inquiry-based science education in science teacher education: A systematic review. *Studies in Science Education, 60*(2), 191–249. https://doi.org/10.1080/03057267.2023.2207148.
Sweller, J., Zhang, L., Ashman, G., Cobern, W., & Kirschner, P. A. (2024). Response to de Jong et al.'s (2023) paper "let's talk evidence – The case for combining inquiry-based and direct instruction". *Educational Research Review, 42*, Article 100584. https://doi.org/10.1016/j.edurev.2023.100584
Teig, N. (2024). Uncovering student strategies for solving scientific inquiry tasks: Insights from student process data in PISA. *Research in Science Education, 54*, 205–224. https://doi.org/10.1007/s11165-023-10134-5
Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University press.
Wan, Z. H., Zhan, Y., & Zhang, Y. (in press). Positive or negative? The effects of scientific inquiry on science achievement via attitudes toward science. Science Education. https://doi.org/10.1002/sce.21825.

Wiley, K., Gerard, L., Bradford, A., & Linn, M. C. (2023). Teaching with technology: Empowering teachers and promoting equity in science. In A. M. O'Donnell, J. Reeve, & N. Barnes (Eds.), *Oxford handbook of educational psychology*. Oxford University. https://doi.org/10.1093/oxfordhb/9780199841332.013.52.

Zacharia, Z. C., Manoli, C., Xenofontos, N., de Jong, T., Pedaste, M., van Riesen, S. A. N., Kamp, E., Mäeots, M., Siiman, L. A., & Tsourlidaki, E. (2015). Identifying potential types of guidance for supporting student inquiry in using virtual and remote labs: A literature review. *Educational Technology Research & Development, 63*, 257–302. https://doi.org/10.1007/s11423-015-9370-0

Zhang, L., Kirschner, P. A., Cobern, W. W., & Sweller, J. (2022). There is an evidence crisis in science educational policy. *Educational Psychology Review, 34*, 1157–1176. https://doi.org/10.1007/s10648-021-09646-1