Educational Research Review 42 (2024) 100584

Contents lists available at ScienceDirect

Educational Research Review

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

Review

Response to De Jong et al.'s (2023) paper "Let's talk evidence – The case for combining inquiry-based and direct instruction"

John Sweller [a,*], Lin Zhang [b], Greg Ashman [c,d], William Cobern [e], Paul A. Kirschner [f,g]

[a] School of Education, University of New South Wales, Sydney, NSW, 2052, Australia
[b] School of Education and Social Work, Providence College, Providence, RI, USA, 02918
[c] Australian Centre for the Advancement of Literacy, Australian Catholic University, North Sydney, NSW, 2060, Australia
[d] Academica University of Applied Sciences, 1017 SG, Amsterdam, the Netherlands
[e] The Mallinson Institute for Science Education, Western Michigan University, Kalamazoo, MI, USA
[f] Faculty of Educational Sciences, Open Universiteit, 6419AT, Heerlen, the Netherlands
[g] Thomas More University of Applied Sciences, 2000, Antwerp, Belgium

ARTICLE INFO

Keywords:
Inquiry learning
Explicit instruction
Randomised
Controlled trials
Cognitive load theory

ABSTRACT

De Jong et al. (2023) objected to the evidence presented by Zhang et al. (2022) to support their concerns about the unreserved acceptance and promotion of inquiry-based learning and problem solving in current policy documents related to the teaching of science. In their response, De Jong et al. (2023) reiterated their advocacy for inquiry approaches, arguing that an emphasis on a mixture of inquiry learning and explicit instruction is needed. The present article rebuts De Jong et al. (2023), in which we: 1) challenge their view of and approach to scientific methods in establishing the efficacy of different instructional approaches; 2) indicate that an underpinning theory to explain the cognitive machinery that drives inquiry-based instructional approaches is missing from their argument; and 3) address the empirical issues arising in their argument. We also highlight potential agreement with De Jong et al. (2023) on the essential role of explicit instruction and thus raise a call to the field to revise current science educational policies and standards to reflect such a role. Our agreements and disagreements advance the debate to a new focus concerning when and how inquiry-based learning and explicit instruction should be used and combined. While De Jong et al. (2023), in their theory-free paper, provided no answer to how explicit instruction and inquiry learning should be combined, we offer our suggestions based on evolutionary psychology and the expertise reversal effect from cognitive load theory.

In 2022, Zhang et al. wrote an article entitled 'There is an Evidence Crisis in Science Educational Policy'. In that article, a considerable gap was noted between what the field has learnt from educational psychology research and what can be seen in educational practice, especially as it pertains to natural sciences education. That article contrasted the findings from three different types or categories of research, namely program-based studies, correlational studies, and randomised, controlled studies. The contrast revealed a noticeable and seemingly significant gap: One specific category of research - program-based studies – has "dominated the formulation of educational standards while a large number of critical findings from randomised controlled studies and correlational studies that overwhelmingly show minimal support for the suggested policy have been marked as irrelevant and excluded". Critically,

* Corresponding author.
E-mail addresses: j.sweller@unsw.edu.au (J. Sweller), lzhang@providence.edu (L. Zhang), greg.ashman@gmail.com (G. Ashman), bill.cobern@wmich.edu (W. Cobern), paul.kirschner@ou.nl (P.A. Kirschner).

https://doi.org/10.1016/j.edurev.2023.100584
Received 20 July 2023; Received in revised form 6 December 2023; Accepted 11 December 2023
Available online 13 December 2023
1747-938X/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

it was not argued by Zhang et al. (2022) that program-based studies are irrelevant, but rather that educational standards should represent a balanced view of all the available data and such a balanced view must therefore also include findings from randomised controlled and correlational studies. Zhang et al. (2022) demonstrated how these different forms of research might inform each other and provided coherent and consistent implications of doing this for educational procedures. The apparently contrasting findings from program-based studies when compared with studies that use randomised controlled trials or correlations can perhaps be explained by the different methods involved in each type of research design. In particular, Zhang et al. (2022) argued that program-based studies do not isolate specific instructional factors or procedures of interest, such as the use of inquiry learning, from other factors that are varied between conditions and that randomised controlled trials and correlational research studies can, and often do, isolate such factors.

A different conclusion was arrived at by De Jong et al. (2023) in 'Let's talk evidence – The case for combining inquiry-based and direct instruction', their *response* to the Zhang et al. (2022) article. While attempting to refute the Zhang et al. (2022) article, De Jong et al. (2023) missed what we consider to be the key point and seemed not to acknowledge areas of apparent agreement with Zhang et al. (2022). De Jong et al. (2023) indicated, somewhat tentatively in parts, that explicit instruction combined with inquiry learning may sometimes be necessary, especially when learners are less knowledgeable in a particular area. Our view, as presented in 2006 (Kirschner, Sweller, & Clark, 2006) and reiterated in 2016 (Zhang, 2016) and 2022 (Zhang et al., 2022), is that explicit instruction is essential for novices but should be associated with an increasing emphasis on independent problem-solving practice as the knowledge of the learner increases. To the extent that De Jong et al. (2023) agree that explicit instruction can be important, we appear to have reached some level of agreement.

The driving concern of Zhang et al. (2022) was that references to explicit instruction are either missing or negligible in many standards and policy documents. The result is that for some teachers in some jurisdictions, the call to use inquiry learning has resulted in students receiving minimal explicit instruction, with an inevitable drop in knowledge acquisition. We therefore raised an urgent call to revise current standards and policy documents to reflect the essential role of explicit instruction in teaching and learning. We also note here that this call should not be controversial as all parties appear to agree that some level of explicit instruction is desirable.

In their response, De Jong et al. (2023) stated that Zhang et al. (2022) provided a definition of explicit instruction to which they claim to have responded. In fact, a definition of explicit instruction was not provided. Instead, Zhang et al. (2022) provided an example of what explicit instruction could look like. That example was not intended to rule out other potential forms of explicit instruction. It therefore seems appropriate to take this opportunity to clarify this example with a working definition. To us, the defining feature of explicit instruction is that, for novice learners, concepts are fully explained, and procedures are fully modelled before learners are asked to apply those concepts or procedures. Significantly, this working definition does not preclude the possibility of learners completing open-ended problem-solving tasks. It simply suggests that if such tasks form part of a sequence of explicit instruction, they would follow explanation and modelling by the instructor. In addition, it is important to note that this definition neither asks nor requires learners to be passive, nor does it exclude instructors from 'active teaching' methods such as asking many questions, requiring responses from all learners, and guiding student practice. Such an approach is consistent with the explicit instruction model derived from teacher effectiveness studies (Brophy & Good, 1986; Rosenshine, 2012). It is possible that through the appropriate sequencing of instructional activities, De Jong et al.'s (2023) conception of a combination of inquiry-based learning and explicit instruction in their response aligns with our definition of explicit instruction. If so, we are in full agreement. On the other hand, we cannot be sure because they did not propose a definition of explicit instruction themselves. Whatever they might propose, we would, however, still seek for such a conception of a combined inquiry-based learning and explicit instruction be clarified in standards and policy documents.

Despite our agreement with De Jong et al. (2023) on the importance of explicit instruction, our view of and approach to scientific knowledge sharply differs from theirs. These differences, and their implications, form a central part of our response. In addition, it is necessary to acknowledge and discuss the atheoretical basis on which the arguments for inquiry learning are advanced in their article.

1. Scientific methods for developing and revising theories

The history of inquiry learning in education is a history of exactly how the learning sciences, or indeed any science, should not proceed. It probably began with advocacy by philosophers such as Rousseau (1762/1979) followed by Dewey (1938) before being extended by Bruner (1961). Following Bruner's article, which used the term "discovery learning", the modern emphasis on inquiry learning in education became dominant. This historical sequence was then amplified by panic in the United States due to the Soviet Union's launch of Sputnik I in 1957 (Kirschner, 2000).

Fuelled by fears of 'losing the space race', the field of science education applied the epistemology of the scientist to the pedagogy for educating future scientists (Kirschner, 1992, 2000). Yet, these are intrinsically different activities. To paraphrase what Derek Hodson (1988) so eloquently stated 35 years ago: A scientist is doing science, a student is learning science. The subsequent rise of inquiry learning appears to have had no connection to data (e.g., Biological Sciences: An Inquiry into Life, 1963; Chemical Education Material Study, 1963; Physical Science Study Committee, 1960). It was not until after the widespread adoption and successful promotion of inquiry learning in various jurisdictions that the collection of data began.

Through the current authors' own participation in the ongoing scholarly discussion about the merits of different teaching methods, we are aware of arguments about the lack of sufficient support for schools and educators in implementing inquiry activities. It is also common to question whether teachers possess the required knowledge for delivering perfectly designed inquiry learning and implementing it with fidelity. What was not seen during the rise of inquiry learning as a teaching method but is indispensable to the development of a scientific theory, is the examination of data that do not support the use of inquiry learning. Such data should then inform the iterative revision and 'improvement' of inquiry learning. It is not scientific to advise the use of a method of instruction based only on favourable data and then focus scholarly efforts on ignoring, or generating reasons to ignore, an extensive body of unfavourable findings. Instead, advocates of inquiry learning should bear the burden of proof. It is they who must demonstrate that the methods they advocate are consistent with the totality of relevant evidence. Insisting that others who are sceptical of inquiry learning must prove it is ineffective is a reversal of the usual form of argument and widely recognised as a logical fallacy.

In their response to the original article, De Jong et al. (2023) justified an inquiry learning approach by referring solely to program-based studies and either ignored or disputed findings from randomised controlled trials and correlational studies that substantially contradicted their suggestions.

2. The irreplaceable requirement for theory in science

Scientific data are usually associated with a theory used to guide the data collection enterprise. In that respect, the use of inquiry learning was previously closely aligned with the theory of "constructivism". It was argued that students needed to learn how to construct knowledge by constructing it themselves and inquiry learning was required for that process to happen (see e.g., Richardson, 1997).

Constructivism has now been largely abandoned as an argument for specific instructional methods, with the term barely used for that purpose today. De Jong et al. (2023) do not use the term for that purpose either and make clear that in their paper they are not advocating for the use of inquiry learning as a means of teaching students how to construct knowledge or how to think but rather, they are advocating for the use of inquiry learning to facilitate knowledge acquisition. They say "Inquiry itself can be seen as the goal of learning in an instructional context, so that students learn what inquiry is, how to perform its various tasks and engage in disciplinary practices and procedures, and what the nature of science is. Alternatively, inquiry can be a means to learn about specific scientific concepts and phenomena …. In this paper, we focus on the second use of inquiry …" (p. 4).

There is a consequence to the abandonment of constructivism. De Jong et al.'s (2023) response is an advocacy for inquiry learning that, as far as we can detect, is not based on any discernible theory. Their paper argues almost entirely on empirical grounds with a heavy emphasis on meta-analyses that attempt to compile disparate studies. They provide no underpinning theory to explain the cognitive machinery that would drive inquiry learning. Such a theory-free approach has its dangers. With no guiding theory, mounting contradictory results have not been explained, as shown in De Jong et al.'s (2023) response. The same applies when they suggest explicit instruction needs to be combined with inquiry learning. Why is explicit instruction needed, why and when should inquiry learning be used, and how can they be combined? What are the principles on which we should make such instructional decisions or is the proposal that any combination of the two methods is effective? If so, why? In the absence of explanations by De Jong et al. (2023), we offer our explanations below.

There is a long-standing argument by inquiry learning supporters that students need to mimic what scientists do in discovering the unknown. Is that the instructional theory that supports inquiry learning? If so, why should methods for instructing novices recreate methods of knowledge production used by experts in the field? We have known since Chi, Feltovich, and Glaser's (1979) article that novices are not simply little experts; that "experts and novices begin their problem representations with specifiably different problem categories, and completion of the representations depends on the knowledge associated with the categories" (p. 121). Given the differences in the cognitive resources between novices and experts, why should novices be asked to acquire knowledge in the same way as experts?

Is inquiry learning being proposed because it fits a theory of motivation or a theory of conceptual development? We are sceptical about assertions that inquiry learning is necessarily more motivating than explicit teaching or that it leads to deeper understanding, but if these are the latent instructional theories adopted by advocates then they need to be made explicit. By doing so, the community of researchers can examine and test such claims. It is also important that these theoretical views should specify the mechanism through which the suggested instructional activities interact with the human cognitive machinery to lead to knowledge acquisition. Without explanations of the mechanism, these so-called theories are no more than interesting propositions. Unfortunately, such a theory, or in fact any theory, is missing in the De Jong et al.'s (2023) response.

We use cognitive load theory (Sweller, 2022; Sweller et al., 2011, 2019) to assemble and interpret data, a theory that De Jong et al. (2023) ignore. Here are some points based on the theory.

1. There are two ways in which humans have evolved to obtain information. The first is by random generation and testing during problem solving (the randomness as genesis principle) and the second is by obtaining information from other people (the borrowing and reorganising principle). These two principles, especially the second, go a long way in explaining the dominance of humans as a mammalian species. Our ability to transfer information effectively and efficiently between us is close to being unique and the huge quantity of information that we can transmit between ourselves certainly is unique. Of course, that information must first be generated, and generation requires problem solving via the randomness as genesis principle.
2. Inquiry learning uses the randomness as genesis principle while explicit instruction uses the borrowing and reorganising principle.
3. The amount of information we can acquire from others using the borrowing and reorganising principle dwarfs the information we can generate ourselves using the randomness as genesis principle. We can obtain information from others in greater quantities and in a fraction of the time it takes to generate it ourselves. Given these circumstances, what is the point of advocating for the inefficient technique of having students work out solutions to novel problems by inquiry rather than explicitly providing them?
4. Based on the above theoretical points, we would expect that showing students how to solve problems (i.e., via explicit instruction) should be superior to having them work out the solutions themselves (i.e., via inquiry learning). The worked example effect demonstrating the advantages of studying worked examples rather than solving problems has been replicated on numerous occasions by many researchers from around the world using fully randomised, controlled experiments (Sweller & Ayres, In press).

Recently, Ashman, Kalyuga, and Sweller (2020) provided clear evidence of the advantages of first providing learners with explicit instruction prior to problem solving when dealing with complex information.

Once a student has learned to solve a problem by either inquiry learning or explicit instruction, they need to both test themselves to ensure they really have learned to solve the problem and practice solving it to ensure automaticity. In other words, instead of continuously restudying the same information obtained from others, they need to solve relevant problems themselves. In cognitive load theory, this is known as the 'expertise reversal effect', and it provides answers to questions about when different instructional procedures become appropriate (Kalyuga, Ayres, Chandler, & Sweller, 2003; Sweller, 2022; Sweller et al., 2011, 2019).

According to our definition, explicit teaching requires novices to be provided with full explanations of new concepts and models of new procedures. Once they are no longer novices, activities similar to those used in episodes of inquiry learning become more appropriate, as the expertise reversal effect demonstrates. This sequence of gradual release of control from instructor to learner is part of our conception of explicit teaching but we are comfortable with others describing this as a mix of explicit teaching and inquiry learning, provided they are clear about when the different strategies are to be used and why.

The role of theory is critically important to this process and cognitive load theory (Sweller, In press) directly addresses the issue. The theory has generated a wealth of data beginning with Tuovinen and Sweller (1999) and Kalyuga, Chandler, Tuovinen, and Sweller (2001), indicating when and what type of information requires explicit instruction (Chen, Kalyuga, & Sweller, 2015; 2016a; 2016b, 2017). High element interactivity information that consists of many elements of information and needs learners to consider all relevant elements simultaneously in order to process them (Chen, Paas, & Sweller, 2023; Sweller, 2010) requires explicit instruction. Element interactivity is high when novice learners are faced with complex tasks such as learning to write an essay or manipulating mathematical equations. Such element interactivity decreases with learners' increasing expertise or if the element interactivity is naturally low because multiple elements do not interact, such as learning the translation of individual, second language words or learning the symbols of the chemical periodic table. The multiple experiments reported in the above papers overwhelmingly indicate that while high element interactivity information requires explicit instruction before problem-solving practice (Ashman et al., 2020), as element interactivity decreases, the advantage of explicit instruction decreases and may eventually reverse, resulting in the expertise reversal effect (Chen, Kalyuga, & Sweller, 2017; Kalyuga et al., 2003).

This explanation indicates why explicit instruction is essential for knowledge acquisition, and when different instructional strategies should be used or combined, as either the content or learners' expertise changes. Our recommendations are based on the relevant predictions made by cognitive load theory along with the extensive empirical base associated with the theory. If De Jong et al. (2023) disagree with cognitive load theory, they at least should provide an alternative to it rather than take a largely empirical, theory-free approach.

3. Empirical issues

There are also empirical issues arising from De Jong et al.'s (2023) response. The key problem continues to be how our field deals with available data that have led to contradictory conclusions. This section discusses conflicting empirical findings.

When discussing program-based studies, Zhang et al. (2022) questioned the possibility of attributing any learning gain to an inquiry learning approach when such studies rely on pre and post designs or compare an innovative program with 'business-*as*-usual'. To us, program-based studies are program evaluations that examine the overall effectiveness of a program or curriculum as a whole package. Since such programs and curricula usually contain a mixture of many different instructional, organisational, and/or content-related components, it is impossible to identify which specific components contribute to any learning gain found. Thus, we cannot be clear that such a learning gain is caused by the inquiry learning procedure. Many studies used to support inquiry approaches in policy and standards documents are program-based studies of this kind. Therefore, Zhang et al. (2022) firmly asserted that while program-based studies are important, relying solely on them to establish standards, such as Next Generation Science Standards (National Research Council, 2013), is inherently misleading.

In responding to Zhang et al. (2022), De Jong et al. (2023) insisted that program-based studies do support the positive effects from inquiry learning approaches. They argued that business-*as*-usual control conditions tend to involve traditional teaching methods, that teachers will be more experienced with these methods, and that students will feel more comfortable with the approaches characteristic of these conditions. Therefore, any learning gains associated with inquiry learning intervention conditions when compared to such controls are real and important.

That argument provides a clear example of the dangers in relying solely on program-based studies. The fact that teachers and students are more experienced with traditional procedures is just one of the many differences between the conditions. Certainly, we are all more comfortable with familiar procedures but how do we know, for example, that familiarity does not engender complacency and reduced effort? Other factors could equally be posited to be the cause of any learning gains. Does the use of various technologies have novelty effects on participants in an inquiry program that generates positive effects? What about the possibility that pressure from research teams' presence generates positive effects? Is it possible that there is more content or even more organised content presented to learners in some inquiry, web-enhanced learning systems? Are there possibilities that teachers are introduced to more content knowledge from workshops affiliated with the acquisition and implementation of those inquiry learning programs that could lead to increased learning by students? Properly run randomised, controlled studies that carefully ensure only one variable is altered at a time, eliminate these possible confounds.

These are not issues unique to program-based studies in educational research. The Hawthorne effect, where participants in an experiment may behave differently due to differential knowledge that they are involved in an experiment, and the placebo effect, where participants and experimenter expectations about a new intervention may produce a positive effect extrinsic to the intervention, are well known across the social and health sciences (see e.g., Merrett, 2006). Both effects seem highly relevant to program-based studies where learners and instructors will know whether they are in the experimental group or the business-*as*-usual condition but are usually far less obvious in randomised, controlled trials that vary a single factor between conditions and where both conditions will appear somewhat novel to learners.

As can be seen, the potential number of conflating factors are many, and so arguments such as the above are unresolvable. What research design is intrinsic to program-based studies that allows us to conclude with any degree of certainty that an observed learning gain within such a program is caused solely by students being taught through inquiry procedures? Our answer is that there are none and that such conclusions can only be definitively arrived at using a strict control of variables approach.

To this point, it is worth noting the wealth of research on instructing learners in the key scientific principle of controlling variables. We cannot help but notice that many researchers, through the studies they design, acknowledge the importance of teaching young learners about the need to control variables and avoid confounds. It is ironic that other researchers in the same field then place value in poorly controlled, confounded program-based studies when discussing the validity of particular, individual instructional methods.

None of the above should be seen as a blanket attack on the use of program-based studies. As emphasised by Zhang et al. (2022) such work can be valuable in providing evidence based on realistic classroom practices over the lengthy periods of time that may be required for students to understand complex concepts and procedures. They can also provide evidence concerning an entire package of instructional procedures. Our objection is to the use of such studies to advocate for specific instructional procedures such as inquiry learning when by their very nature, these studies alter multiple variables simultaneously rendering them inappropriate vehicles for determining the effectiveness or otherwise of any individual variable under consideration. Other techniques are required to provide evidence for or against the effectiveness of specific variables.

Zhang et al. (2022) also discussed findings from correlational studies that presented results that do not support inquiry learning approaches. Correlational studies investigate relationships between instructional factors and students' achievement across large populations, and if guided by theories, allow researchers to further examine if the suggested approaches have worked after being implemented in practice. Zhang et al. (2022) indicated that such studies overwhelmingly support explicit instruction. In their response, De Jong et al. (2023) argued that explicit instruction may be preferred during initial instruction but then might be enhanced by problem solving, a conclusion we entirely agree with, based on the expertise reversal effect.

It needs to be noted that just as we do not reject the use of program-based studies, we do not automatically accept the findings of correlational studies. They are not designed to establish causality and their use in the current context of student self-reports may result in biases. Notwithstanding, they may provide important indicators. Our objection is to many curriculum documents treating these studies as though they do not exist.

In responding to the results from controlled studies presented in the original article, De Jong et al. (2023) suggested that Zhang et al. (2022) handpicked those controlled studies from a small group of researchers and ignored many other studies. Thus, they challenged us by incorporating meta-analyses and reviews that synthesised studies from the literature and argued that those studies had found comparative benefits of inquiry learning over explicit instruction.

We do not dispute the results of those meta-analyses but are puzzled by De Jong et al.'s (2023) advocacy for them. These meta-analyses attempt to combine many individual studies that examine explicit and inquiry learning, with considerable variations in research designs, treatments, and contexts across those individual studies. Accordingly, the meta-analyses reflect that disparity, exactly as we may expect. These syntheses can include both strictly randomised, controlled trials and program-based studies. As Zhang et al. (2022) indicated, these two categories of studies yield different results, something the conclusions of meta-analyses often reflect. Indeed, due to the constraints that any meta-analysis must face, the authors of those meta-analyses drew their implications and conclusions with caution. For example, Furtak et al. (2012) made it clear that "In this sense, we harbour no assumptions that this or any other meta-analysis will definitively answer the questions raised about a particular instructional approach." (p. 322). We were surprised that in contrast to the caution of these review authors, De Jong et al. (2023) chose meta-analyses that because they are based on a combination of both program-based studies and randomised, controlled trials, are unable to supply any definite answers to attempt to lead this debate to a decisive conclusion.

It must be emphasised that while we favour full, randomised, controlled studies, we do not advocate that they should be the only source of information for decision-making concerning instruction. Despite their strengths, like all techniques they have their flaws. For example, they can be difficult to run in normal classrooms and tend to be of shorter duration than program-based studies. Nevertheless, because of their strengths, they should be a pivotal factor in decision-making, rather than being entirely overlooked or, worse, mistakenly equated with program-based studies. Such program-based studies may occasionally employ randomization, but they typically involve the simultaneous alteration of multiple variables, and thus can rarely be characterized as controlled.

4. Conclusion

In conclusion, while we welcome the acknowledgement of a need for explicit instruction by De Jong et al. (2023), we are bemused by their theory-free approach to inquiry learning with its limited evidence of effectiveness from randomised, controlled trials. They provide neither theory nor data to indicate when or how explicit instruction and inquiry learning should be used and combined. Meta-analyses that are primarily concerned with program-based studies rather than randomised, controlled trials that examine one factor at a time reflect exactly the issues that Zhang et al. (2022) discussed in their article. De Jong et al. (2023) strongly object to the Zhang et al. (2022) conclusions, seemingly unaware that their own conclusions must inevitably flow from Zhang et al.'s (2022) analyses.

Data without theory is likely to appear random. Cognitive load theory provides an overarching theory that can be used to both generate data and organise it. The theory indicates both how and when explicit instruction and inquiry should be used. One of the crowning glories of human evolution, our skill in rapidly communicating large amounts of information to each other, is a central principle of cognitive load theory: the borrowing and reorganising principle. Cognitive load theory is concerned with how to organise instruction to maximally make use of this principle. It is a primary skill universally used by all humans including the readers and writers in the current exchange but is considered inappropriate by De Jong et al. (2023), under some undefined conditions, for learners in educational settings. Even worse, at times and for reasons that remain unclear, that skill is to be replaced by an inefficient procedure that we otherwise only use when communication from others is unavailable. It remains a mystery to us why students, and only students, should, under unspecified conditions, be capriciously prevented from using what is arguably the most important evolved skill of our species, our unique ability to rapidly assimilate large amounts of novel information from other people.

Data availability

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

References

Ashman, G., Kalyuga, S., & Sweller, J. (2020). Problem solving or explicit instruction: Which should go first when element interactivity is high? *Educational Psychology Review, 32*, 229–247.

Brophy, J. E., & Good, T. L. (1986). Teacher behavior and student achievement. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 328–375). Macmillan.

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

Chen, O., Kalyuga, S., & Sweller, J. (2015). The worked example effect, the generation effect, and element interactivity. *Journal of Educational Psychology, 107*, 689–704.

Chen, O., Kalyuga, S., & Sweller, J. (2016a). Relations between the worked example and generation effects on immediate and delayed tests. *Learning and Instruction, 45*, 20–30.

Chen, O., Kalyuga, S., & Sweller, J. (2016b). When instructional guidance is needed. *Educational and Developmental Psychologist, 33*, 149–162.

Chen, O., Kalyuga, S., & Sweller, J. (2017). The expertise reversal effect is a variant of the more general element interactivity effect. *Educational Psychology Review, 29*, 393–405. https://doi.org/10.1007/s10648-016-9359-1

Chen, O., Paas, F., & Sweller, J. (2023). A Cognitive Load Theory approach to defining and measuring task complexity through element interactivity. *Educational Psychology Review, 35*. https://doi.org/10.1007/s10648-023-09782-w

De Jong, T., Lazonder, A., Chinn, C., Fischer, F., Gobert, J., Hmelo- Silver, C., et al. (2023). Let's talk evidence – the case for combining inquiry-based and direct instruction. *Educational Research Review*. https://doi.org/10.1016/j.edurev.2023.100536

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

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38*, 23–31. http://www.erlbaum.com. http://www.leaonline.com/loi/ep.

Kalyuga, S., Chandler, P., Tuovinen, J., & Sweller, J. (2001). When problem solving is superior to studying worked examples. *Journal of Educational Psychology, 93*, 579–588. H http://www.apa.org/journals/edu.html.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75–86.

National Research Council. (2013). *Next generation science standards: For States, by States*. The National Academies Press.

Rosenshine, B. (2012). Principles of instruction: Research-based strategies that all teachers should know. *American Educator, 36*(1), 12–39.

Rousseau, J.-J. (1762/1979). Emile: Or on education. Basic Books.. .

Sweller, J. (2023). The development of cognitive load theory: Replication crises and the incorporation of other theories can lead to theory expansion (in press) *Educational Psychology Review*.

Sweller, J. (2010). Element interactivity and intrinsic, extraneous and germane cognitive load. *Educational Psychology Review, 22*, 123–138.

Sweller, J. (2022). The role of evolutionary psychology in our understanding of human cognition: Consequences for cognitive load theory and instructional procedures. *Educational Psychology Review, 34*, 2229–2241. https://doi.org/10.1007/s10648-021-09647-0

Sweller, J., & Ayres, P. (2023). Worked examples (in press). In A. O'Donnell, N. C. Barnes, & J. Reeve (Eds.), *The Oxford Handbook of Educational Psychology*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199841332.013.17.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer.

Sweller, J., van Merriënboer, J., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*, 261–292.

Tuovinen, J. E., & Sweller, J. (1999). A comparison of cognitive load associated with discovery learning and worked examples. *Journal of Educational Psychology, 91*, 334–341. http://www.apa.org. http://www.apa.org/journals/edu.html.

Zhang, L. (2016). Is inquiry-based science teaching worth the effort? Some thoughts worth considering. *Science & Education, 25*(7), 897–915.

Zhang, L., Kirschner, P. A., Cobern, W. W., & Sweller, J. (2022). There is an evidence crisis in science educational policy. *Educational Psychology Review, 34*(2), 1157–1176. https://doi.org/10.1007/s10648-021-09646-1