Discovery learning: zombie, phoenix, or elephant?

Arthur Bakker[1] (ID)

Abstract Discovery learning continues to be a topic of heated debate. It has been called a zombie, and this special issue raises the question whether it may be a phoenix arising from the ashes to which the topic was burnt. However, in this commentary I propose it is more like an elephant—a huge topic approached by many people who address different aspects. What is needed in the discussion about discovery learning and related approaches, I argue, is *sublation*: the kind of lifting up from the one-dimensional discussion between two extremes (minimal guidance vs. instruction) that puts an end to the everlasting tug of war by integrating justified concerns from both opposite positions. I evaluate how the different contributions to the special issue help to sublate the discussion about discovery learning to a higher level. In particular, the case study presented by Trninic illustrates how strong guidance and repetition may be needed for the discovery of something that cannot be told. I further suggest scaffolding, inferentialism, and design research as potential theoretical and methodological ways forward.

Few topics in the educational and learning sciences have led to such heated debate as discovery learning. The wonderful side of the debate is manifold. First, it shows that the topic is at the heart of what education and learning are about. Second, it has inspired many scholars over decades to write poetically and use strong rhetorical techniques to persuade

---

Commentary for the special issue on discovery learning guest-edited by Dor Abrahamson and Manu Kapur.

✉ Arthur Bakker
a.bakker4@uu.nl

[1] Freudenthal Institute, Utrecht University, Princetonplein 5, 3584 CC Utrecht, The Netherlands

others of particular views. The dark side of this debate is that a trivial version of discovery learning—the minimally guided extreme has recently dominated the more nuanced versions. To counter this problem, I start with a short history of nearly one thing (paraphrasing Bryson 2003), namely discovery learning and related ideas. I then introduce the three figures in my commentary—zombie, phoenix, and elephant—before I get to the question of how to make progress, with the help of the contributions to this special issue, on the topic of discovery learning, and lift the discussion and research about it to a higher level. I offer three suggestions for doing so: Focusing on scaffolding, using design research, and taking an inferentialist perspective.

A short history of nearly one thing

Already in the 1950s and 1960s, discovery learning was widely debated, but influences go back as mentioned by several contributors to this special issue (Roll et al. 2018; Abrahamson and Kapur 2018) as far as Plato's dialogue between Socrates and Meno, Rousseau (1762/1979), Dewey (1938), and progressive education movements (Montessori, Waldorf, Jena plan, Kees Boeke, etc.). Progressive education reacted to excesses of rote learning and several attempts were made to realize curricula that gave students room to make discoveries and understand the reasons behind rules and procedures (Beberman and Meserve 1956a, b; Davis 1960; Suchman 1964). Due to the emergence of these reform movements (Ellis and Berry 2005), one of the debates was about the role of discovery as opposed to instruction. Stanley (1949) already ruled out the ultra-progressive view that discovery would be without the need to structure the learning process, as in an Easter-egg hunt (p. 455). For those not accustomed with Easter practices: In some countries, parents hide chocolate eggs in the garden (or at home, but this is less fun) for children to discover them. Stanley further wrote that "the issue becomes, then, not instruction versus discovery, since both are essential, but a consideration of the relative importance to be accorded each in the educative process" (p. 457).

The Russians' 1957 launching of the first satellite to orbit Earth, Sputnik 1, caused much consternation in the rest of the world. In the USA, one of the many aftermaths was fierce discussion about the quality of American education. Was the USSR's technological achievement possible due to its high-quality education? The National Science Foundation became interested in science curriculum projects to close the so-called "Missile Gap" between the USSR and USA (Bruner 2006). The National Academy of Science organized the Woods Hole conference with the same intention and asked Jerome Bruner to preside over it; other famous participants were Skinner, Inhelder, and Cronbach. It fell to the chairman to write a report on this conference, *The Process of Education* (Bruner 1960a, b), on American educational problems and potential solutions. Interestingly, the first translation was in Russian. Bruner would soon afterward write *The Act of Discovery* (1961).

Discovery as a goal of learning or a means of teaching

At this stage, it seems that the primary appeal of discovery was discovery as the *goal* of science, education, or learning and teaching—much in line with the idea of autonomy as the ability to think freely and independently, which is often considered an important goal of education (Bakhurst 2011; Biesta 2015). Discovery, however, was also seen as a *means* of teaching and learning. Bruner (1960b) wrote that educators often thought in terms of reward and punishment, much in line with behaviorism. He noted that so much less was

known about interest and curiosity and he hoped that intrinsic reward could be fostered by the design of new curricula. Bruner (1961) did not want to restrict the term discovery to the act of finding something unknown to humankind, but rather wanted to "include all forms of obtaining knowledge for oneself by the use of one's own mind" (p. 22). This does not mean that students would have to be left to the own devices: "Discovery, like surprise, favors the well prepared mind" (p. 22). And Bruner (1960a) also conceded that "one cannot wait forever for discovery" or "leave the curriculum completely open" (p. 613).

Where Bruner seemed much in favor of students making their own discoveries, Ausubel (1961, 1963) accused him of conflating the *goal* with the *method* of discovery learning. Ausubel was further critical of the "mystique" (1963, p. 139) and "deification" (1963, p. 140) of the discovery method, but also of its opposite (rote learning). He demystified nine propositions about learning by discovery such as "all real knowledge is self-discovered" (p. 144).

As soon as a method is proposed, the question of effectiveness arises. One of Ausubel's concerns was that the discovery method would not be time and cost effective. In the 1960s empirical research on discovery learning (Kersh 1958, 1962; Davis 1960) was already reviewed in a nuanced debate (Bittinger 1968; Shulman and Keislar 1966). For example, Bittinger wondered if some conditions had had a fair chance.

Discovery learning as a means to other ends, and individual differences

Even when using a focus on discovery as a means for learning, multiple different purposes are mentioned in the literature: improving memory, motivation (Kersh 1962), intrinsic reward, or interest (Bruner 1960b). One could even consider a discovery attitude (Cronbach 1966, cited in Bittinger 1968) or discovery skills as learning goals worth striving for. Not surprisingly, discovery methods—with sufficient support—seem an appropriate way to foster discovery skills (De Jong and Van Joolingen 1998).

Cronbach (1966) further noted the relevance of individual differences: "I am tempted by the notion that pupils who are negativistic may blossom under discovery training, whereas pupils who are anxiously dependent may be paralyzed by demands for self-reliance" (cited in Bittinger 1968, p. 145).

Human values and the quality of interventions

There are many more aspects to the idea of discovery, two of which I discuss in relation to a mathematician and mathematics educator, cited in several contributions to this special issue: Hans Freudenthal (La Bastide-van Gemert 2015). Whenever he came across a theorem, he tried to prove it himself, which for him was reinvention rather than discovery. Similarly he wanted students to engage actively in doing mathematics, much like how people learn to swim (Freudenthal 1971; see Trninic 2018). Freudenthal (1971) pointed with pathos to the *value* of treating students as humans when he wrote "Telling a kid a secret he can find out himself is not only bad teaching, it is a crime" (p. 424). This is not to deny that there is, and always be, a tension between transmitting disciplinary knowledge developed over many centuries on the one hand, and discovery by students on the other (Freudenthal 1983).

Where much of the debate on discovery learning was on effectiveness, Freudenthal had an eye for the *quality* of instruction. For example, he criticized common approaches to discovery learning for a very different reason than its effectiveness. In his characterization, discovery learning too often came down to "uncovering what was covered by somebody

else—*hidden Easter eggs*" (1991, p. 46). He was a proponent of guided reinvention instead—active mathematical thinking supported by high quality tasks and teaching.

Apart from quality of the methods evaluated in experiments (cf. Bittinger 1968), one could criticize the *length* of most experiments on discovery versus instruction. For example, Dean and Kuhn (2007) showed that findings from experiments on discovery learning may be different when taking a longer-term perspective (maintenance).

Opportunities of technology

Another milestone in the history of discovery learning and related approaches was the introduction and development of technology, which offered new opportunities and challenges (Papert 1980); see also De Jong and Van Joolingen (1998) for a nuanced story in relation to computer simulations. An interesting distinction is between exploratory models that are designed by experts so that students can discover things by experimenting within a microworld (Doerr 1997), versus expressive models (model building) that allows what might be called reinvention. Further technology-related issues arise with reference to some of the contributions to this special issue.

The elephant in the room

Here finally comes the elephant in the room, the reason why I bothered to dive into the history of the special issue's topic: The more recent debate has been dominated by attacks on a version of discovery learning that no sensible educator would endorse, and that all contributors to this special issue quickly dismiss: minimal guidance (Kirschner et al. 2006; Mayer 2004). These papers on minimal guidance have received considerable criticism (Hmelo-Silver et al. 2007; Schmidt et al. 2007) and yet have high citation numbers—a phenomenon worth being studied by sociologists, philosophers, and historians of science. I am not sure if there is a causal relationship, but the term discovery then seems to have lost popularity. In a recent meta-analysis of innovative approaches to science and mathematics education, Savelsbergh et al. (2016) found many experiments on inquiry-based and context-based approaches, but hardly any called discovery methods (see also Furtak et al. 2012; Lee and Anderson 2013).

Was the topic of discovery burnt down? How is it possible that compared to the debates over 60 years ago, the current debate seems so flat? Even longer ago, Goethe wrote the following:

> *All wise thoughts have been thought before; we must only try to think them again.*
> Alles Gescheite ist schon gedacht worden, man muß nur versuchen, es noch einmal zu denken. (von Goethe, 1828, Wilhelm Meisters Wanderjahre, Zweites Buch, Section 41)

If he is right, it is worth doing more historical research on the topic of discovery learning and related approaches, and try to revive wise ideas from great minds. The latter is one of my aims in this commentary.

Zombie, phoenix, or elephant?

Turning gradually to the current special issue, let me introduce three figures. The critics of minimally guided discovery learning have come up with very strong images. With reference to several review studies, Mayer (2004) asked if there should be "a three-strikes rule against pure discovery learning" (p. 14). Because the idea of discovery learning kept popping up over the time span of decades, he characterized discovery learning as "some zombie that keeps returning from its grave" (p. 17).

When I read about this special issue coming up, my initial response was admiration for the Guest Editors' courage to revive the debate about discovery learning. This admiration grew when I received the contributions to the special issue with nuanced discussions of the topic and the careful thinking about the role of technology. The Guest Editors' title is a nice pun on Freudenthal's preference for reinvention: Re-inventing discovery learning (Abrahamson and Kapur 2018). The reader may thus wonder: Is this special issue like the mythical phoenix arising from its own ashes?

When reading some of the well-known publications on discovery learning from many decades ago as well as the contributions to this special issue, another image came to my mind: The ancient Indian parable of several blind men touching an elephant (Fig. 1). These men were asked to describe what they felt, but they all touched different parts of the elephant. Hence the man feeling the trunk thinks it is a snake, the one feeling the ear infers it is a rug, and the third man touching the tusk judges it as a spear. The others feel a tree, a wall, and a rope.

I hasten to write that I do not want to suggest that the six articles in this special issue equate the six men in Fig. 1. What I tried to highlight in my short history of discovery learning is that different authors emphasize different aspects of discovery: the purpose of education is to be able to make discoveries; the means of the discovery method for multiple purposes (effective learning, intrinsic motivation, interest, discovery skills etc.); the quality and length of discovery approaches; but also the new opportunities technology offers students to explore relationships and discover patterns, and get immediate feedback. And the discussion is much richer. For example, there is a strong link with constructivism: certain things cannot be transmitted; they have to be constructed or experienced. For those things that cannot be told it makes sense to think through how educators create opportunities for construction or new experiences. Of course, some things can and should be told (Baxter and Williams 2010; Schwartz and Bransford 1998).

The aforementioned aspects are main issues in *didactics* (in the continental European sense of domain-specific pedagogy): What to cover up as a teacher or designer and what to let students discover? Teachers and designers always keep many things out of the problem space in which they want students to discover or become aware of something. I find this point worth spelling out, because it shows that there is no teaching for discovery learning without covering up (Hovinga 2007), or in terms of several contributions to this special issue: putting constraints on the problem space in which students do particular tasks (e.g., Levy et al. 2018). It is the designer or teachers who makes the judgments of what constraints will be productive in learning some target knowledge or skill. The paradox is that the more autonomy we want to give students, the better we have to design the tasks—a lesson I learned in statistics education (e.g., Bakker and Gravemeijer 2003). I mention these more general aspects to point to the enormous size of the discovery topic to frame my discussion of the contributions to this special issue.
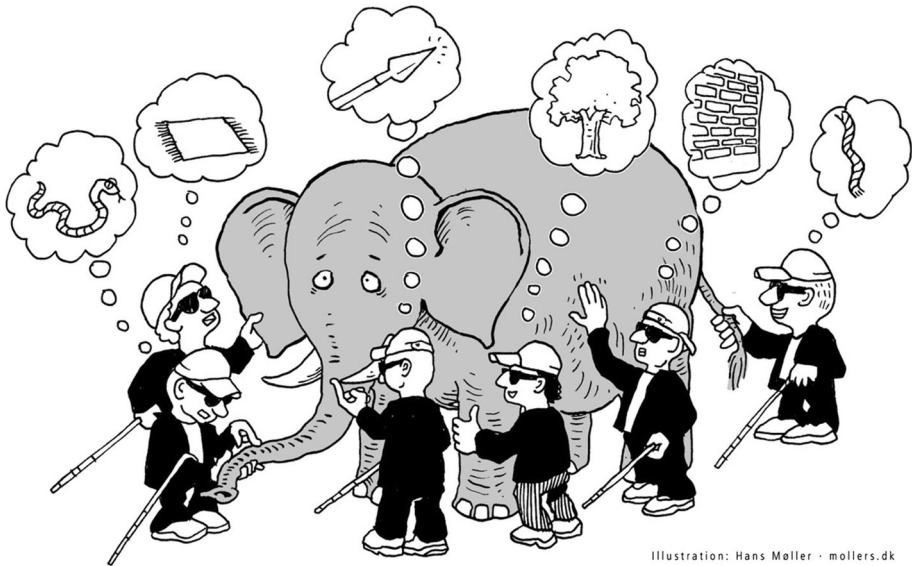
Fig. 1 Parable of the blind men touching an elephant. Reprinted with permission from Hans Møller

The contributions to the special issue

The various contributions to the special issue point to further aspects of the elephant of discovery learning. For example, Wilkerson et al. (2018) convincingly argue that there should be *alignment* between on the one hand the epistemic games (modeling strategies) that designers and teachers invite students to engage in and on the other hand the epistemic forms (model types) that the modeling environment is designed to support. They show in their design-based research how they supported students by means of curricular activities with computer tools in playing the intended epistemic game. Moreover, like others (Roll et al. 2018; Levy et al. 2018), these authors found differential effects across student groups.

Levy et al. (2018) also designed curricular materials with the intention of fostering particular discoveries. The power of their article lies in presenting both a general design framework and a concrete example. More explicitly than other authors, they write about the importance of social interaction for learning, a point I return to in relation to scaffolding. An interesting point is the design of constraints into the system to limit possible actions so as to guide and focus students on particular aspects of the system. This is another manifestation of the covering up and uncovering that I wrote about when teaching or designing. Their findings about discovery learning are differential: effects were not uniform—neither across students nor across learning goals. What I learn from this is that it is hardly possible to claim anything about discovery-related approaches in general.

Roll et al. (2018) combine various aspects of the elephant of discovery learning: They take into account differences between students but also measure several dependent variables (transfer, attitudes, behaviors), and show awareness of the distinction between short-term and long-term effects (cf. Dean and Kuhn 2007). As such, their research clearly goes beyond Cronbach's (1966) assumption about individual differences and Bruner's (1960b, 1961) assumptions on possible effects of discovery learning. Roll et al.'s study indicates the complex interaction between guidance and student characteristics. I admire

the work done to relate the aforementioned aspects, but I also became worried that larger and larger studies would be necessary to investigate the interactions between all relevant factors. To me it seems there is a limit to such quantitative approaches. In my view, additional, more detailed, and smaller-scale studies are necessary to understand the mechanisms behind the complex interaction as indicated by Roll and colleagues.

Kapur (2018) raises the interesting question of the benefits of problem-posing per se or problem-posing with problem-solving, for making discoveries. With reference to Kapur's study, I make a plea for research teams with diverse expertise to be aware of various sides of the elephant. In particular, I missed the involvement of experts in statistics education. Before I underpin this point, I ask the reader to solve the task Kapur used with 14 to 15-year-olds who had just first learned about the standard deviation:

An equal number of students competed in the 100 m sprint and 100 m swim finals. The timings (in s) of the champions of the 100 m sprint and 100 m swim are shown below, as are the average timings and the SDs of the finalists in the two competitions

	100 m sprint (s)	100 m swim (s)
Champion	11	40
Average of the finalists, M	12	45
SD of the finalists	1	10

Assuming all else being equal, between the two champions, who is the better performer?

A. The sprint champion

B. The swim champion

C. Both are equally good

D. Not enough information to decide (p. 9–10 in the online-first version)

My first problem with this item is that it cannot be expected that students having just learned about the SD can answer this question. Even university students and statistics teachers struggle with the concept of SD in relation to distribution and mean (delMas et al. 2007; delMas and Liu 2005; Groth and Meletiou-Mavrotheris 2018; Peters 2009). My second problem is that I found it very hard to imagine a distribution with the following characteristics: minimum of 40 s, arithmetic mean of 45, and an SD of 10. For me the validity problem of using such a task implies that the findings on transfer in this paper have to be bracketed, and cannot be more than working hypotheses.

Based on my communication with several statistics educators, I predict that experts in statistics education would not have chosen this item to measure students' understanding of the SD, hence my plea to work in bigger and diverse teams on such important research. I want to emphasize that my point of criticism is not meant to diminish the value of Kapur's study on problem-posing and problem-solving, only to highlight a validity problem that I—being raised in a tradition of didactical research in mathematics and statistics education—have encountered quite often. For example, mathematics educators also frowned about a study by Berends and Van Lieshout (2009) because the items used were inappropriate in the eyes of didactical experts (van den Heuvel-Panhuizen and Peltenburg 2011).

A strength of the work by Chase and Abrahamson (2018) is that they carried out both an experiment and conducted posthoc qualitative analysis of what happened in the two conditions (Discovery vs. No-Discovery). The detailed examples shed light on the

proposed mechanisms that may have led to the differences in outcomes. Their work matches well with Sandoval's (2014) plea to check—in addition to measuring learning outcomes—whether a design really embodies the intended theoretical characteristics and to study which mediating processes take place that explain the learning outcomes. The theoretically rich analysis provides useful insights, even though the conditions are dichotomized. The topic of the next section is how to go, with the help of the last contribution (Trninic 2018), beyond such dichotomies.

Sublation

The contribution by Trninic (2018) convinced me that there is a way out of the dichotomies in the debate about discovery learning. The common dichotomies dissolve in his case study about martial arts, skillfully compared with mathematics education. It is by repetition and continued guidance that the student discovers what the martial arts teacher wants him to discover. The story resonates well with my own experiences in sports, music, and mathematics, that teachers can prepare and guide some lessons, but that learners can only discover these lessons in their full embodied sense after much deliberate practice.

The dissolving dichotomies in Trninic's study resonates with Hegel's idea of sublation (*Aufhebung* in German). The German term has three different meanings, which Hegel often combined (Pinkard 1996; Inwood 1992). In the literal sense, *aufheben* is to lift up to a higher level, like a glass when saying "cheers" or getting up from a chair (the first sense: to raise, to lift up). When Hegel uses the term, what is lifted up is often elements of both sides of a tension (or contradiction) to a new integration preserving something of what was integrated (*aufheben* in the second sense: to keep, preserve, save). One can think of drafting a paper that is not fully consistent; even when completely rewriting it, traces of the old draft will be preserved in an improved version of the paper. Inwood (1992, p. 282) gives the example of feeling and thinking, which are often considered opposites but can be combined in, for instance, an experience of beauty (or interest in something). The term *aufheben* also has the connotation of ending the tension as in the meaning of lifting a ban (*aufheben* in the third sense: to abolish, annul, cancel, suspend).

How can one concept combine seemingly contradictory meanings? Inwood (1992, p. 283) gives the example of the concept of development. In development, the different senses of sublation play a role: Early stages of developmental process are sublated into later and higher stages. In development, certain elements are abolished while others are preserved, though perhaps at a higher level of organization. Another example that may make the idea of sublation more concrete is that of the mathematical concept of proportion. It illustrates how also mathematical concepts can integrate seemingly contradictory meanings: How could one thing at the same time be also two? How can something be both continuous and discrete? Proportion, as an equivalence class of ratios, shows that these opposites can be integrated. For example, 1:2, 2:4, and 3:6 all refer to the same proportion even though their constituent elements are different. One could say that proportion is both one and two at the same time: It is one relationship of two elements. One could also argue that proportion is both discrete—think of all discrete instances (4:8, 5:10, etc.)—and continuous: Put your hands on the table and move them up in such a way that the right hand is moving twice as fast as the left one; your hands then move proportionally in a continuous way (Abrahamson et al. 2014).

I propose to use this concept of *Aufhebung* to lift the tension between guidance and discovery, and lift up the discussion about discovery learning to a higher level. After all, Ausubel (1963) already concluded that "learning by discovery is not necessarily anti-thetical to programed instruction, despite the howls of anguish which teaching machines frequently elicit from discovery enthusiasts" (p. 162, "programed" in the original spelling). The question, however, is how? I suggest three ways that in my view can assist in this endeavor, two theoretical and one methodological.

Three possible ways to sublate

Scaffolding

One source of problems in the discovery learning discussion is that teachers and students are often conceptualized as separate systems with a causal relationship between teaching and learning, instead of a dialectic subject|object ensemble: Students are not only objects of a pedagogical process but also subjects of it, as Roth and Lee (2006) write. Teachers further make discoveries about their students and might in some sense be guided by their students. In Western educational psychology, a turn to the student and more knowledge-able other (teacher, parent) as a dynamical system was made in the scaffolding literature. Wood et al. (1976) wrote about scaffolding as "an interactive system of exchange" (p. 99), in their case of mother and child.

As soon as educators see teacher and students as such a system, questions become more nuanced than whether the discovery method or direct instruction should be applied. The question becomes what is in the region of sensitivity to instruction (Wood and Middleton 1975), in the zone of proximal development (Vygotsky 1978), or in the zone of next development (Smagorinsky in press). What needs to be told, what can only be discovered, what needs to be repeated and practiced? The answer to such questions requires profes-sional judgment, there and then. As long as researchers can stick to this original sense of scaffolding with a dynamical systems lens, they should be able to stay out of a one-directional view in which a scaffold is seen as a support system working on students as objects. Within such a systems view, however, guidance and discovery have to form into a coordinated whole. To study the mechanisms in such coordination processes, I think inferentialism and design research can be of help.

Inferentialism

Inferentialism is a recent philosophical semantic theory (Brandom 2000) that bears potential for education (Bakker and Hußmann 2017; Derry 2013; Guile 2006). It has already shown to be able to overcome dichotomies such as theoretical versus practical knowledge (Heusdens et al. 2016) and the acquisition versus participation metaphors for learning (Taylor et al. 2017). Understanding this theory takes some investment, but I think the return is worth the investment, also for the current topic of discovery learning.

Inferentialism is a holistic, pragmatist, expressivist, and rationalist theory of meaning. More particularly, it is about the use and content of concepts (Brandom 2000). Rather than taking representation as the initial basis for explaining meaning, Brandom takes social reasoning practices as the starting point. His philosophy is in line with a long tradition of anti-representationalist philosophy, which has also influenced educational theory (e.g.,

Cobb et al. 1992). Put simply, a concept for Brandom is not a representation of, say, a class of referents, but the norm governing the use of the concept. So one could see a concept as entailing the inferences that can be made with it. This explains why his theory of concepts (and therefore knowledge) is holist:

> inferentialist semantics is resolutely *holist*. On an inferentialist account of conceptual content, one cannot have *any* concept unless one has *many* concepts. For the content of each concept is articulated by its inferential relations to *other* concepts. Concepts, then must come in packages (though it does not yet follow that they must come in just one great big one). (Brandom 2000, pp. 15–16; emphases original)

This idea implies that one cannot understand the concept standard deviation (SD) as one entity. The meaning of SD is intricately connected with all inferences that can be made with this concept in relation to, for example, mean, distribution, data, the square root, summation, and the inflection points in a normal distribution.

> Brandom (2000) further takes a pragmatist stance on understanding concepts:
> To grasp or understand (…) a concept is to have practical mastery over the inferences it is involved in—to know, in the practical sense of being able to distinguish, what follows from the applicability of a concept, and what it follows from. (p. 48)

As Derry (2013) noted, this idea is in line with how Vygotsky thought about concepts:

> We must seek the psychological equivalent of the concept not in general representations, … not even in concrete verbal images that replace the general representations—we must seek it in a system of judgments in which the concept is disclosed. (Vygotsky 1998, p. 55)

This implies that understanding a concept such as SD (in relation to other concepts) is always a matter of degree: Users of a concept may have practical mastery of many inferences it is involved in but not all.

From an inferentialist perspective it is impossible, as Trninic (2018) also observes, that direct instructional guidance would be able to provide "information that fully explains the concepts and procedures that students are required to learn" (Kirschner et al. 2006, p. 75). The transfer item on SD in Kapur's article may illustrate this. It is known in the statistics education literature that students who know the computational definitions of arithmetic mean and of SD often do not see the mean as a measure of center or the SD as a measure of variation. For example, when comparing two data sets they often do not use the mean and/ or SD as relevant characteristics of the distributions that could tell which of two situations is favorable (delMas and Liu 2005; Konold and Pollatsek 2002). It is therefore to be expected that the 14 to 15-year-old students in Kapur's study learned a few things about SD but most likely had not learned to navigate the rich inferential terrain of the concepts required to answer the item.

Let me make a confession to hammer this point home. When I saw Kapur's (2018) item on the sprinters and swimmers, I suspected students were supposed to look at the SD relative to the mean. My first stumbling block was actually contextual: 100 finalists? I have an image of 6 or 8 finalists in such competitions. More importantly, I find it important that students see mean and SD as characteristics of a distribution (Bakker and Gravemeijer 2004), so I tried to envision what the distribution of the swimmers could look like yet initially was not able to. My approach, instigated by Kapur's comment on normalization, was as follows: I sketched a normal distribution with a mean of 45; then indicated an SD of 10 at the lower inflection point, and realized that I had to cut off the distribution at 40,

because this was the champion's time. Then I worried that I had to cut off the higher tail of the distribution too, otherwise the mean would drift to the right. And then I wondered: Is there a distribution with these characteristics at all? I inserted fifty times 40 and fifty times 50 into Excel and calculated the SD: too low (about 7). And here comes my confession: It took me a few hours and a hint from a colleague to make the distribution skewed. I discovered something new, which was yet a direct consequence of what I knew already. I had just explored the inferential terrain of the concepts further, with the help of a tool and a colleague. The correct answer, by the way, was (A): the sprint champion, because this time was a full SD ahead of the average time whereas the swim champion was only half an SD ahead.

To hint at how inferentialism may assist in sublating the discussion on instruction versus discovery, I need two more metaphors from Brandom's inferentialism. The first is the *game of giving and asking for reasons* (GoGAR); the second is *scorekeeping*. In Brandom's view, people involved in reasoning practices may be seen as playing a kind of rational game. When they say things, they commit themselves to their statements and are held responsible for their claims. Whether they are entitled to say these things is up to others in the GoGAR. The mechanism at stake here is scorekeeping: People keep track of what they and others say and can confirm, nuance, or correct these statements. These reasoning practices are governed by historically and culturally developed norms of people in touch with the world and expert in particular disciplines.

One could envision a teacher with her students as being engaged in a GoGAR, with all doing some scorekeeping. The teacher keeps track of what students seem to know and need, and the students keep track of the norms of what counts as correct (cf. Yackel and Cobb 1996). Such scorekeeping can of course be informed by digital technology (learning analytics, feedback, etc.). From an inferentialist view, it would be silly to think of students discovering or constructing their own knowledge (Noorloos et al. 2017), because it is the teacher (with designed materials) who represents the norms of the discipline to be learned. Likewise it would be silly to think of a teacher fully explaining what is to be learned, because—as mentioned earlier—the inferences that can be made with the help of concepts are boundless. I think it is worth studying teacher–student interaction from an inferentialist perspective to understand better the development of disciplinary normativity.

Design research

Why do researchers dichotomize discovery versus direct instruction? I conjecture that randomized controlled trials as the so-called gold standard force researchers to dichotomize, and thus compare extremes. Moreover, thinking in terms of independent and dependent variables elicits a one-directional thinking about cause and effect. When the learning and educational sciences are indeed a design science (Glaser 1976; Simon 1967; Wittmann 1995), the key challenge becomes to design so as to improve educational practice. This requires more complex models of causality and change.

Design research (design-based research, design experiments) has been proposed as a methodological orientation to do so (Collins et al. 2004; Cobb et al. 2003). When doing design research, such as some authors in this special issue (Chase and Abrahamson 2018; Levy et al. 2018; Wilkerson et al. 2018), one loses methodological control but typically gains ecological validity (Brown 1992). When really trying to improve practice, and figuring out how the intricate interplay between guidance and discovery works, one is likely to face many aspects of the elephant I identified, notably the quality of the intervention, the detailed considerations of what to cover up and what to highlight in the design (with or

without technology), the unique characteristics of students you work with, et cetera. Generalization then typically takes another form than statistical generalization from sample to population. Rather, theoretical generalization and transportation of useful design ideas become the productive ways of generalizing.

Conclusion

In this commentary I asked if discovery learning should be characterized as a zombie returning from its grave (Mayer 2004), or a phoenix arising from the ashes to which the topic of discovery seemed to have been burned. I argued it was more like an elephant—a huge topic in the educational and learning sciences with many aspects. Are the six contributions to this special issue like the six blind men feeling part of the elephant of discovery learning? I am more optimistic than that. My conclusion is that they have managed to lift it up together. What they show, some more explicitly than others, is that there can indeed be an integration of repeated guidance by a teacher and discovery by a student (Trninic 2018); that careful technology design and teaching can cover up what students should not (yet) see so that they can dis-cover particular insights that are not easily "told" (Chase and Abrahamson 2018; Levy et al. 2018; Wilkerson et al. 2018).

I hope that this special issue contributes to lifting the spell that has been on discovery learning for some years (one sense of sublation). With the simplistic extremes exorcized, educational and learning scientists can return to the hard work of doing design and research on how to support learners to reinvent or discover what is worth learning but cannot be told, and thus raise the discussions to a higher level (another sense of sublation). I happily embark on the research program sketched by the Guest Editors and raise my glass to all contributors.

References

Abrahamson, D., & Kapur, M. (2018). Reinventing discovery learning: A field-wide research program. *Instructional Science*. https://doi.org/10.1007/s11251-017-9444-y.

Abrahamson, D., Lee, R. G., Negrete, A. G., & Gutiérrez, J. F. (2014). Coordinating visualizations of polysemous action: Values added for grounding proportion. *ZDM Mathematics Education, 46*(1), 79–93. https://doi.org/10.1007/s11858-013-0521-7.

Ausubel, D. P. (1963). *The psychology of meaningful verbal learning*. New York, NY: Grune & Stratton.

Bakhurst, D. (2011). *The formation of reason*. Oxford: Wiley.

Bakker, A., & Derry, J. (2011). Lessons from inferentialism for statistics education. *Mathematical Thinking and Learning, 13,* 5–26.

Bakker, A., & Gravemeijer, K. P. E. (2003). Planning for problem solving in statistics. In R. Charles & H. R. Schoen (Eds.), *Teaching mathematics through problem solving: Grades 6–12* (pp. 105–117). Reston, VA: National Council of Teachers of Mathematics.

Bakker, A., & Gravemeijer, K. P. E. (2004). Learning to reason about distribution. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 147–168). Dordrecht: Kluwer Academic Publishers.

Bakker, A., & Hußmann, S. (2017). Inferentialism in mathematics education: Introduction to a special issue. *Mathematics Education Research Journal, 29,* 395–401.

Baxter, J. A., & Williams, S. (2010). Social and analytic scaffolding in middle school mathematics: Managing the dilemma of telling. *Journal of Mathematics Teacher Education, 13*(1), 7–26.

Beberman, M. A. X., & Meserve, B. E. (1956a). An exploratory approach to solving equations. *Mathematics Teacher, 49*(1), 15–18.

Beberman, M. A. X., & Meserve, B. E. (1956b). Graphing in elementary algebra. *Mathematics Teacher, 49*(4), 260–266.

Berends, I. E., & Van Lieshout, E. C. (2009). The effect of illustrations in arithmetic problem-solving: Effects of increased cognitive load. *Learning and Instruction, 19*(4), 345–353.

Biesta, G. (2015). What is education for? On good education, teacher judgement, and educational professionalism. *European Journal of Education, 50*(1), 75–87.

Bittinger, M. L. (1968). A review of discovery. *The Mathematics Teacher, 61*(2), 140–146.

Brandom, R. B. (2000). *Articulating reasons: An introduction to inferentialism*. Cambridge, MA: Harvard University Press.

Brown, A. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences, 2*(2), 141–178.

Bruner, J. S. (1960a). On learning mathematics. *The Mathematics Teacher, 53*(8), 610–619.

Bruner, J. S. (1960b). *The process of education: A searching discussion of school education opening new paths to learning and teaching*. New York, NY: Vintage.

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

Bruner, J. S. (2006). *The selected works of Jerome Bruner, 1957-1978, Volume I: In search of pedagogy*. New York, NY: Routledge.

Bryson, B. (2003). *A short history of nearly everything*. New York: Broadway Books.

Chase, K., & Abrahamson, D. (2018). Searching for buried treasure: uncovering discovery in discovery-based learning. *Instructional Science*. https://doi.org/10.1007/s11251-017-9433-1.

Cobb, P., Confrey, J., DiSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher, 32*(1), 9–13.

Cobb, P., Yackel, E., & Wood, T. (1992). A constructivist alternative to the representational view of mind in mathematics education. *Journal for Research in Mathematics Education, 23,* 2–33.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *Journal of the Learning Sciences, 13*(1), 15–42.

Cronbach, L. J. (1966). The logic of experiments on discovery. In L. S. Shulman & E. H. Keislar (Eds.), *Learning by discovery: A critical appraisal* (pp. 76–92). Chicago, IL: Rand McMally.

Davis, R. B. (1960). The "Madison Project" of Syracuse University. *The Mathematics Teacher, 53*(7), 571–575.

De Jong, T., & Van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research, 68*(2), 179–201.

Dean, D., & Kuhn, D. (2007). Direct instruction vs. discovery: The long view. *Science Education, 91*(3), 384–397.

delMas, R., Joan, G., Ooms, A., & Chance, B. (2007). Assessing students' conceptual understanding after a first course in statistics. *Statistics Education Research Journal, 6*(2), 28–58.

delMas, R., & Liu, Y. (2005). Exploring students' conceptions of the standard deviation. *Statistics Education Research Journal, 4*(1), 55–82.

Derry, J. (2013). *Vygotsky: Philosophy and education*. Oxford: Wiley Blackwell.

Dewey, J. (1938). *Experience and education*. New York, NY: Touchstone.

Doerr, H. M. (1997). Experiment, simulation and analysis: An integrated instructional approach to the concept of force. *International Journal of Science Education, 19*(3), 265–282.

Ellis, M. W., & Berry, R. Q., III. (2005). The paradigm shift in mathematics education: Explanations and implications of reforming conceptions of teaching and learning. *Mathematics Educator, 15*(1), 7–17.

Freudenthal, H. (1971). Geometry between the devil and the deep sea. *Educational Studies in Mathematics, 3*(3/4), 413–435.

Freudenthal, H. (1983). The implicit philosophy of mathematics history and education. In Z. Ciesielski & C. Olech (Eds.), *Proceedings of the International Congress of Mathematicians* (pp. 1695–1709). Amsterdam: Elsevier.

Freudenthal, H. (1991). *Revisiting mathematics education: China lectures*. Dordrecht: Kluwer Academic Publishers.

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.

Glaser, R. (1976). Components of a psychology of instruction: Toward a science of design. *Review of Educational Research, 46*(1), 1–24.

Groth, R., & Meletiou-Mavrotheris, M. (2018). Research on statistics teachers' cognitive and affective characteristics. In D. Ben-Zvi, K. Makar, & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 327–355). Cham: Springer.

Guile, D. (2006). Learning across contexts. *Educational Philosophy and Theory, 38*(3), 251–268.

Heusdens, W. T., Bakker, A., Baartman, L. K. J., & De Bruijn, E. (2016). Contextualising vocational knowledge: A theoretical framework and illustrations from culinary education. *Vocations and Learning, 9*(2), 151–165. https://doi.org/10.1007/s12186-015-9145-0.

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.

Hovinga, D. (2007). *Discover and cover: Learning about the manifold relationship of humans to nature in learning trajectories about sustainable development*. Utrecht: CD-β Press.

Inwood, M. J. (1992). *A Hegel dictionary*. Oxford, UK: Blackwell.

Kapur, M. (2018). Examining the preparatory effects of problem generation and solution generation on learning from instruction. *Instructional Science*. https://doi.org/10.1007/s11251-017-9435-z.

Kersh, B. Y. (1958). The adequacy of "meaning" as an explanation for the superiority of learning by independent discovery. *Journal of Educational Psychology, 49*(5), 282–292.

Kersh, B. Y. (1962). The motivating effect of learning by directed discovery. *Journal of Educational Psychology, 53*(2), 65–71.

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.

Konold, C., & Pollatsek, A. (2002). Data analysis as the search for signals in noisy processes. *Journal for Research in Mathematics Education, 33*(4), 259–289.

La Bastide-van Gemert, S. (2015). *All positive action starts with criticism. Hans Freudenthal and the Didactics of Mathematics*. Houten: Springer.

Lee, H. S., & Anderson, J. R. (2013). Student learning: What has instruction got to do with it? *Annual Review of Psychology, 64,* 445–469.

Levy, E. T., Peleg, R., Ofeck, E., Tabor, N., Dubovi, I., & Bluestein, S., et al. (2018). Designing for discovery learning of complexity principles of congestion by driving together in the TrafficJams simulation. *Instructional Science* (this issue).

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *American Psychologist, 59*(1), 14–19.

Noorloos, R., Taylor, S. D., Bakker, A., & Derry, J. (2017). Inferentialism as an alternative to socioconstructivism in mathematics education. *Mathematics Education Research Journal, 29*(4), 437–453. https://doi.org/10.1007/s13394-017-0189-3.

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. New York: Basic Books.

Peters, S. A. (2009). *Developing an understanding of variation: AP statistics teachers' perceptions and recollections of critical moments*. Unpublished doctoral dissertation, The Pennsylvania State University, State College.

Pinkard, T. (1996). *Hegel's phenomenology: The sociality of reason*. Cambridge: Cambridge University Press.

Roll, I., Butler, D., Yee, N., Welsh, A., Perez, S., Briseno, A., et al. (2018). Understanding the impact of guiding inquiry: The relationship between directive support, student attributes, and transfer of knowledge, attitudes, and behaviours in inquiry learning. *Instructional Science*. https://doi.org/10.1007/s11251-017-9437-x.

Roth, W.-M., & Lee, Y.-J. (2006). Contradictions in theorizing and implementing communities in education. *Educational Research Review, 1*(1), 27–40.

Rousseau, J.-J. (1762/1979). *Emile, or on education* (A. Bloom, Trans.). New York: Basic Books.

Sandoval, W. A. (2014). Conjecture mapping: An approach to systematic educational design research. *Journal of the Learning Sciences, 23*(1), 18–36.

Savelsbergh, E. R., Prins, G. T., Rietbergen, C., Fechner, S., Vaessen, B. E., Draijer, J. M., et al. (2016). Effects of innovative science and mathematics teaching on student attitudes and achievement: A meta-analytic study. *Educational Research Review, 19,* 158–172.

Schmidt, H. G., Loyens, S. M., Van Gog, T., & Paas, F. (2007). Problem-based learning is compatible with human cognitive architecture: Commentary on Kirschner, Sweller, and Clark (2006). *Educational Psychologist, 42*(2), 91–97.

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction, 16*(4), 475–522.

Shulman, L. S., & Keislar, E. R. (Eds.). (1966). *Learning by discovery: A critical appraisal*. Chicago, IL: Rand McNally.

Simon, H. A. (1967). *The sciences of the artificial*. Cambridge, MA: MIT Press.

Smagorinsky, P. (in press). Deconflating the ZPD and instructional scaffolding: Retranslating and reconceiving the zone of proximal development as the zone of next development. *Learning, Culture and Social Interaction*.

Stanley, J. C. (1949). The role of instruction, discovery, and revision in early learning. *Elementary School Journal, 49*(8), 455–458. https://doi.org/10.1086/459075.

Suchman, J. R. (1964). The elementary school training program in scientific inquiry. *AV Communication Review, 12,* 246–247.

Taylor, S. D., Noorloos, R., & Bakker, A. (2017). Mastering as an inferentialist alternative to the acquisition and participation metaphors for learning. *Journal of Philosophy of Education, 51*(4), 769–784. https://doi.org/10.1111/1467-9752.12264.

Trninic, D. (2018). Instruction, repetition, discovery: Restoring the historical educational role of practice. *Instructional Science*. https://doi.org/10.1007/s11251-017-9443-z.

van den Heuvel-Panhuizen, M., & Peltenburg, M. (2011). Continuous research lines in mathematics education research for low-achievers. *Orthopedagogiek, 50,* 16–20.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge: Harvard University Press.

Vygotsky, L. S. (1998). *The collected works of L.S. Vygotsky* (Vol. 5). In R. W. Reiber (Ed.), Child psychology. New York, NY: Plenum.

Wilkerson, M. H., Shareff, R., Laina, V., & Gravel, B. (2018). Epistemic gameplay and discovery in computational model-based inquiry activities. *Instructional Science*. https://doi.org/10.1007/s11251-017-9430-4.

Wittmann, E. C. (1995). Mathematics education as a 'design science'. *Educational Studies in Mathematics, 29*(4), 355–374.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, 17*(2), 89–100.

Wood, D., & Middleton, D. (1975). A study of assisted problem-solving. *British Journal of Psychology, 66*(2), 181–191.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education, 27*(4), 458–477.