Grand theories or design guidelines? Perspectives on the role of theory in computing education research

MATTI TEDRE, University of Eastern Finland, School of Computing
JOHN PAJUNEN, University of Jyväskylä, Department of Philosophy

A rich body of empirically grounded results and a solid theory base have often been viewed as signs of a mature discipline. Many disciplines have frequently debated what they should accept as legitimate kinds of theories, the proper roles of theory, and appropriate reference disciplines. Computing education research (CER) in particular has seen a growing number of calls for the development of domain-specific theories for CER, adaptation of theories from other fields, and engagement with theory-based experimental and predictive research in CER. Many of those calls share the same concerns and aims, yet they use very different vocabulary and lack a consensus over an essential concept: theory.

This article presents sticking points and trouble spots in CER's theory debates and presents a number of suggestions and ways forward. Firstly, by slightly shifting towards a model-based view of science, CER can avoid centuries of conceptual baggage related to the concept of theory. Secondly, insofar as fields like design, engineering, and social science are considered to be legitimate parts of CER, the role of theory in many CER studies needs to be judged by the criteria of the philosophy of engineering, technology, and social science, not the philosophy of (natural) science. Thirdly, instead of force-fitting elements of ill-suited research paradigms from other disciplines, the philosophy of CER should focus on building a consensus on CER's own paradigm and describing the field's relationship with theory in CER's own terms.

CCS Concepts: • **Social and professional topics → Computing education**.

Additional Key Words and Phrases: philosophy of computing education research, philosophy of CER, models, theory, design, philosophy of science

1 INTRODUCTION

Over the course of its maturation as a research discipline, the field of computing education research (CER) has seen a significant change in the field's publication profile. From its early beginnings in the 1960s until the late 1990s, the CER literature base saw more than its fair share of anecdotal experience reports, course overviews, teaching tool descriptions, proofs of concept, essays, program surveys, nifty assignment expositions, and other similar non-empirical outputs [23, 75, 83, 93]. One of the more common types was the infamous "Marco Polo"-type article: "I went there and I saw this" [93].

But at the turn of the millennium, a sea change began, as an increasing number of conferences and journals started to require researchers to present empirical evidence for their claims. Many early calls for increasing the rigor of computing education research focused on the field's research methods [75, 93]. The computing education symposia, which used to be "swap meets" for sharing one's pedagogical techniques, useful assignments, or lecture props [31], started to request the authors base their statements on empirical evidence. At the same time—and not completely unrelated—a number of nascent CER research groups and capacity-building initiatives had focused on training bona fide computing education researchers, who were able to deliver just what the

Authors' addresses: Matti Tedre, firstname.lastname@uef.fi, University of Eastern Finland, School of Computing; John Pajunen, firstname.lastname@jyu.fi, University of Jyväskylä, Department of Philosophy.

1946-6226/2022/1-ART1
https://doi.org/10.1145/3487049

symposia wanted [23, 33]. As new methodological standards for research were increasingly adopted, the field saw a gradual transformation in the 2000s [47, 75, 83].

The movement to build CER on a solid foundation primarily focused on methodological standards, but there was an important parallel in many articles: many authors identified a need to critically evaluate the field's theory base—or lack thereof. For example, Holmboe et al. [37] lamented the field's lack of reference to pedagogical theory. Goldweber et al. [31] proposed that CER needed to establish relationships between observations and theories. Report after report surveys of CER showed a weak theoretical base in the field [83], and two pioneers of the field declared in 2006 that there is no Kuhnian-type research paradigm in CER [23].

Research studies have shown that CER publishing venues differ greatly in their approaches to the role of theory. Joy et al. [39] showed that CER conferences, with their strict page limits, tend to ignore educational theory more than journals do. Clancy et al. [9] argued that CER papers in education research conferences are better positioned to involve educational theory in the context of CER.

In the 2010s, a steady stream of essays and analyses kept alive discussions about the roles and relevance of theory in CER. Nelson and Ko [63] wrote that in a design-based field, a demand for theory can detract the researcher from the main aims of the field, inhibit the search for CER-specific theory, and create a publication bias that benefits neither theory testing nor the development of better designs for learning. Szabo et al. [89] reviewed the CER body of literature for theory-related research, identified popular theories used in the field, and proposed a template for visualizing various dimensions of theory. Papamitsiou et al. [67] found that despite the marked increase in theoretically driven research in CER, theory choice was rarely featured in the author keywords of published articles.

"Theory," however, is not a well circumscribed concept, and compared to fields like mathematics education research, which has seen the rise and fall of many theoretical "-isms" and lived through eras like New Math and the science wars [85], CER has a more limited history of theory debates in the field. Like many others, Malmi et al. [51] noted that theory discussions in CER have been complicated by many different meanings of the word "theory." The variety of theories is broad: for examples of theoretically sound research, one group of conference panelists listed action research, phenomenography, socio-cultural perspectives, and cognitive theories [4]. Isomöttönen [38] theorized phenomena in computing using grounded theory. Pears et al. [70] distinguished between disciplinary, methodological, analytical, and interpretive uses of theory. On occasion, the theory discussants have felt a need to explicitly state whether they are for or against theory in CER [89, p.90]. An important aspect concerning multiple functions of theory complicates the matter: in addition to the many traditional functions of theory, theories also inform scientists about what questions are relevant and how to tackle them, and they serve as lenses for observation, or, as the philosophers of science put it: observations are theory-laden [43].

The discussions about the role of theory in CER feature a wide variety of descriptions of CER as a field [9, 23, 31]. Nelson and Ko [63] characterized CER as fundamentally a design field whose goal is to "discover designs that produce better learning of computing," and literature reviews are in consensus that designing and implementing tools specifically for computing education is a major part of the field [39, 93]. At the same time, many discussants have emphasized the field's nature as a research discipline that seeks deep, fundamental understanding of phenomena related to computing education [50, 83]. Some of them have emphasized the field's search for (possibly conditionally) generalizable theories that allow prediction [4, 22, 91]. And many combine the views, arguing that a solid foundation on theories about learning improve learning designs in computing.

The concerted push by the field's opinion leaders and gatekeepers towards a stronger theoretical basis for CER has impressed a new dimension of quality-consciousness on many CER researchers. The aims of that push are laudable, such as moving away from practice papers to contributions of theoretical relevance and empirical grounding [50], optimizing the field's search for better learning designs [63], enabling derivation of pedagogically useful insights from theory [4], and bringing CER closer to its sister fields in other disciplines, such as mathematics education research [50], physics education research [52], and other similar "trading partners" [23].

The above views on the role of theory in CER, many of which were advanced by the field's pioneers, have relied on a range of philosophical positions and theories of scientific research. As such, those views exhibit remarkable depth of disciplinary understanding, and all are well justified from their perspectives. The field has a wealth of philosophically rich contributions. For example, Isomöttönen [38] offered a discussion on theory in the context of grounded theory. Thota et al. defended paradigm pluralism [92]. In their pioneering textbook on CER, Fincher and Petre [22] presented a 76-page deep analysis of the nature and potential of theory in CER; theirs is a must-read starting point for any discussion on the use of theory in the field. Another fruitful analysis of roles of theory can be found in Pears et al. [70, e.g., Fig.1]. But similar to Almstrum et al. [2] who were, justifiably, of the opinion that "too much of the research in computing education ignores the hundreds of years of education, cognitive science, and learning sciences research that have gone before us," this article proposes that some theory discussions in computing education research ignore the latest 40 years of research on the philosophy of science, sociology of science, and science and technology studies (STS). This article is aimed at contributing some alternative perspectives to CER's theory discussions by using results and experiences from fields that have recently contributed to those debates, most of all the philosophy and sociology of science. The article assumes some familiarity with basic vocabulary in the philosophy of science but attempts to avoid jargon and technicalities.

Firstly, the article presents that the historical baggage that comes with the myriad uses of the term "theory" complicates discussions about the role of theory in CER. The article proposes for CER a modern, model-based view of scientific research. Secondly, the article discusses some challenges with borrowing theory from other fields to be used in CER, as well as some open questions about the nature of theories in general. Thirdly, the article suggests that at the heart of theory debates in CER are disagreements that are strikingly similar to the 1980s disagreements about computing's disciplinary nature and about the role of design in the discipline. The article notes that the conflicts regarding computing's disciplinary identity dissolved as a consensus was established over computing's tripartite nature and that the three branches of computing (theory, science, engineering) are irrevocably intertwined. Perhaps that is where the philosophy of CER is ultimately heading, too.

The article presents the above three topics in the context of the philosophy of science, a field with a long history in debates like CER's theory debates. In that context, the article reminds the reader that many models of scientific explanation quoted in CER's theory discussions were explicitly about natural sciences and never meant as a model of social sciences, education research, or psychology. Many philosophers of social science have nowadays largely abandoned the quest to claim the mantle of hard science for social science, and sought to understand the field in its own right and its own terms [17, pp.1–8]. Models of natural sciences are neither applicable to nor desirable for many sectors of social sciences; the concerns of the fields are fundamentally different [17]. The article also describes a rift between the early 1900s' "received view" of science that prescribed a priori rules for how successful scientists should work (see [34]), and the late 1900s "naturalistic view" of science that relies on empirical (e.g., sociological, historical, and psychological) research on how successful scientists actually do work. The article advances the view that instead of appropriating any a priori paradigm from other disciplines, and instead of relying on examples of theory use from fields like physics and medicine, CER can gradually build an understanding of its own paradigm (a disciplinary matrix of exemplars and other elements of research) by investigating how progress in CER has been made in its influential and most successful studies.

2 FROM GRAND THEORIES TO MODEL-BASED REPRESENTATIONS

Some aspects of CER's theory debates resemble mid-20th century discussions in the field of sociology, when the field was seeking a disciplinary identity and a new direction. Some sociologists yearned for a general, all-encompassing theory of sociology, while some other sociologists wished to follow the natural sciences. Both views, however, led to trouble—the former due to its overly broad scope and lack of straightforward applicability, and the latter much due to misconceptions about the natural sciences those sociologists wished to imitate, including overplaying the role of inductivist reasoning (from a lot of data to theory), ignoring the maturity phase of a discipline (thinking that sociology, a new field then, would function the same as more established fields of inquiry like physics), and speaking as if natural sciences had an all-encompassing "grand unified theory." It seemed that grand theories invariably faced the dilemma of being either trivial or wrong [60, p.70]. In this context of the field's methodological debates, the sociologist Robert K. Merton [55] presented the concept of middle-range theories.

Not unlike many CER discussions of today, Merton [55] described a scientific theory to be a set of propositions that serve as grounds for deriving "empirical uniformities." But there is much more to the picture. A theory alone is never a sufficient basis to carry out such derivations: auxiliary propositions, assumptions, and hypotheses are necessary to carry out the intended cognitive processes (e.g., [73]). Methodologically speaking, experiments never reveal anything about a single theory, but about the whole experimental setup, including also the instruments involved; one's theory about how the instruments work; a whole universe of auxiliary assumptions, hypotheses, and theories; and epistemological and ontological standpoints (e.g., [71]). Merton noted a large gap between a very general, high-level "grand" theory and what Merton called evolving working hypotheses that are used in daily research, close to data. He argued that to maintain a relatively uninterrupted chain between data and general theory, middle-range theories are needed.

Middle-range theories are separate from sweeping general theories, such as variants of constructivism in CER, as they are not logically derived from them—but they are also separate from data in the sense that they are not simple generalizations therefrom (such as statistics of eye-tracking data or keystroke patterns in CER). So, methodologically speaking, a middle-range theory is an entity that lies between theory and empirical observation (or data). It is logically (inferentially) independent from both the most abstract level of general theories as well as from data at the observation level. Middle-range theories stand in between abstract theories and empirical data also substancewise, adding domain-specific information to the picture and having, unlike general theories, a closer link to empirical observation. Different from the views of science of the early 1900s, such as logical empiricism or naïve inductivism, the image of science sketched by Merton was not purely inductive (theories are generalizations of observations) or purely deductive (observation sentences are derived logically from theory).

In CER, general theories are loose-knit, diverse, and overlapping (e.g., [55]). Take some of the many theories used in CER, for example: zone of proximal development, mental models, and Maslow's hierarchy of needs [89]. Not all the theories used in CER are straightforwardly operationalizable for the characteristic uses of theory typically mentioned in CER literature (such as prediction or interpretation of findings. For the aims of theory in CER see, e.g., [52, 63, 89]). Not all the theories used in CER offer readily applicable prescriptions for generating new learning designs, either. The concept of middle-range theories was intended to bring about a focus on distinct problems, allow for new hypotheses, direct attention in novel ways, and thereby furnish empirical work [55, 57].

The descriptions of the theoretical underpinnings of CER by Malmi et al. [52] and Szabo et al. [89] reveal challenges similar to those that vexed Merton and his successors. The desire to come up with broader and more predictive theories in sociology was aimed at extending the reach and utility of the discipline [55], the mastery of society, predictive power, control, and "human engineering" [57]. But emulating a misconceived picture of natural sciences turned out to be counterproductive, both towards those aims and towards the other aims of the discipline as well [57]. Similarly, CER debates need not be structured around simplistic, idealistic views of scientific investigation or explanation. For instance, Popper's falsificationism, the deductive-nomological model, and inductivist views of science—all referenced in CER's theory debates—advance narrow views of the roles of theory, experiment, and data [36, 71, 94].

It has been pointed out that it will be important for the nascent philosophy of computing education research to analyze what kinds of meta-theory, or "theory of theory," the various arguments for the role of theory in CER adopt (Malmi et al. [51] refer to this as the ontological nature of theory). The philosophy of science has seen upheavals since the 1980s, but those changes are not strongly visible in CER's theory discussions. There is room for richer discussion on whether the theory of theory adopted in CER theory discussions is closer to Hacking's [36] or Pickering's [71] view of theories as interventional elements that integrate and are integral for technology development, or whether CER's theory of theory is closer to the syntactic view of the 1920s and 1930s Vienna Circle (see, e.g., [7]). (For reference, instead of the classic 800-page tome on the topic by Suppe [87], Winther [96] offers a more accessible and more up-to-date overview of the structure of scientific theories).

From a modern perspective, Merton's view of science is, like its predecessors, relatively simplistic and more concerned with ideals of the time than what really happens in successful science projects (e.g., [36, 44, 71]). His idea is narrowly defined as sets of propositions, but he does broaden the view in the end to include "general orientations toward data, suggesting types of variables which theories must somehow take into account, rather than clearly formulated, verifiable statements of relationships between specified variables" [55, p.52 in original; p.458 in reprint], so Merton clearly acknowledges a variety of cognitive aspects of scientific work. Yet, the search for middle-range theories describes the task for a computing education researcher, like the task for a sociologist, according to Merton, but with a proviso that an eye for general concepts and propositions is kept.

Middle-range theories may be what many CER theory discussions seem to primarily mean by "theory." Isomöttönen [38] is explicit about that, but the view is implicit all around. A good example is Robins et al. [78, p.262], who present a hierarchy for describing CER, where theories at the functional level connect to a higher level of more abstract pedagogical concerns as well as to a lower level of cognition and behavior of individuals. Their view highlights an important point about focusing on the level that matters, even when higher and lower levels of abstraction set some necessary conditions and boundaries. How Merton saw the state of sociological research resembles in many ways how many people see CER today. Merton's proposal was to introduce an element that can bridge the most abstract level of theory and close-to-observation level of data. The following sections present different approaches to that.

2.1 Three Views of Theories

Simple models of science (think of logical positivism, naïve inductivism, falsificationism) can provide fruitful entry points for discussing with students some fundamental differences between historical schools of thought in the philosophy of science. The models fare less well in discussions about a field's disciplinary identity, methodology, epistemology, or theoretical content. An authentic description of how scientific discovery and explanation (among other things) work in a specific discipline needs to distinguish and analyze a large number of processes and concepts. Simplicity was a problem with many early models of science, with their building blocks limited to a small number of concepts like theory, data, hypothesis, and observation—which have all later turned out to be complex, multifaceted concepts. Models of science have become ever more complex towards the 2000s, and the 1980s especially saw many black boxes of scientific explanation opened. Even such basic concepts as experiments were largely taken for granted before the 1980s, after which analyses by sociologists and philosophers of science showed that experiments are a multidimensional and complex concept, understood in numerous different ways [35, 74, 90].

Decades of research have deepened the understanding of the roles and functions of theory in the scientific enterprise, and numerous competing and complementing views can be found today [30, 36, 71]. Some views are broad, others narrow, and many of them can be found in CER analyses of theory. A broad understanding is adopted by, for instance, Malmi et al. [52, p.29]:

> "*We define "theory" to mean a broad class of concepts that aim to provide a structure for conceptual explanations or established practice, and use such terms as "theories," "models," and "frameworks" to describe particular manifestations of the general concept of theory.*"

Similarly, for the purpose of their review of the use of theory in CER, Szabo et al. [89] provided a working definition of theory, in an inclusive form, leaving a more precise definition for future work. Szabo et al. [89] cited the definition of Malmi et al. [52], and their view is inclusive in the sense that it takes theory to also include models and frameworks, and its relationships between theory, model, and data are rather unspecified, aside from their direction of fit. Szabo et al.'s [89] view is that educational theory abstracts, explains, describes, or predicts learning, and theory is also related to (validates, feeds into/is derived from) models, which are further related to the phenomenon of learning via data. A number of descriptions of theory in CER (e.g., [52, 63, 89]) provide excellent entry points for further analysis of the nature of theory and theoretical explanation [63]. The same inclusivity applies to other keywords used in CER's theory discussions, such as explanation, description, abstraction, and experiment—but research in the philosophy of science has shown that neither the concepts nor their relationships are simple and straightforward (e.g., [74]).

The positions of philosophers of science over nearly a century of debates about the nature and structure of scientific theories can be summarized in three broad categories: the syntactic, semantic, and pragmatic views of theory [96]. The syntactic view of theory, advocated by logical empiricists, sees theories as sentences consisting of logical terms and domain-specific terms—ideally couched in axiomatic form. Its theory structure includes theoretical sentences, observational sentences, and sentences to link these together [96]. Although intuitively alluring, the syntactic view had lost its currency by the 1950s with the demise of logical positivism, as the task of linking the observational terms to theoretical terms proved to be problematic (for classic exposition, see [87]).

An alternative to the syntactic view of theory, the semantic view, emphasizes that the structure of a theory consists of mathematical, set-theoretical, or model-theoretical models. The model theoretical approach looks into the set of models that make a theory true, that is, what interpretations a theory has. The models of a theory are varied, including abstract axioms that depict the most general structure of the theory, models that represent the theory at a more substantial level, and models of experiment and data [96]. Models in the semantic view include the informational or knowledge side of understanding that account on all abstraction levels for phenomena and specify the kinds of relations between data, experiment, and theory. The relations may be hierarchical, similarity, or some other kind [96].

The most modern view of theory, the pragmatic view, allows the most leeway with using the term "theory," and like the above descriptions of theory in CER, it allows theory and model to include "mathematical concepts, metaphors, analogies, ontological assumptions, values, natural kinds and classifications, distinctions, and policy views" [96]. Furthermore, it allows varied types of theories and models, such as "mechanistic, historical, and mathematical models." The pragmatic view also allows theories to include non-formal aspects or make use of analogies, metaphors, and other tools not immediately available in the logical or mathematical theory orientations.

In CER, the variation of theories in terms of their types and constituents is rich and not restricted to universal law-like sentences. Theories invoke values, perceptions, practical skills and a range of concepts that are hard to formalize, let alone axiomatize: think of identity, mastery, and discovery, for example [89]. Also ontologically, the most recent wave of theory of theory—the pragmatic view of theory—is the most compatible with CER, as CER includes a range of ontologically distinct entity-types, such as learner, artifact, and mindset [89]. Different theories or models of CER would ascribe to them different attributes (e.g., emotive or cognitive characteristics, in the case of the learner) and different relations (instructor as the source of information vs. the learner as the creator of knowledge).

2.2 Some Functions of Models and Data

Malmi et al. [52, see quote above] suggest that theories, models, and frameworks play roughly similar roles in CER literature, but Malmi et al. leave the relationships of those elements unclear. Many other fields share the same concern with terminological ambiguity. To clear up some of that ambiguity, for some decades, philosophers of science have increasingly emphasized the concept of model and analyzed the nature of theory and the nature of models separately from each other, which allows for a more fine-grained analysis of science [28]. Most importantly, that separation allows one to move away from the massive historical and conceptual baggage that comes from grand theories and from debates between syntactic, semantic, or pragmatic views of theory [see 96]. Separating out the concept of model allows one to focus on representation and on the important roles that models play in scientific edifice, either as complements of theories or, if theories are viewed more like Kuhnian paradigms, as parts of theory constructions. The past decades have seen extensive discussions about what models are and how they relate to data and theory, along with some other fundamental questions (for a representative bibliography, see [26]).

Models come in many shapes and forms (e.g., [19]). Some are physical (scale models), some formal or mathematical, and some linguistic or graphical, just to mention a few. Some are human-constructed, some computer-generated. Some are related to human action, like didactic models, models of decision making, or models of care of addiction [26]. Like theories, models are intended to serve a range of functions, reflecting the different functions of scientific knowledge: to describe and explain the world or to predict and understand phenomena, for instance. But due to the long history of theory debates in sciences, conceptual baggage burdens theories more than models—take expectations like testability, falsifiability, explanatory power, and verisimilitude, for instance (although not all theories of theory require all of those). Models are allowed more flexibility, less rigor, and more pragmatic functions (even when one could grant theories the same role).

A model-oriented view of science enables one to focus on how researchers use models at different levels to represent aspects of the world for specific purposes [29]. Depending on their function, some models are abstract and cursory, while others are explicit and detailed; some focus more on processes, others more on objects. Giere [30] provides several example figures of hierarchically related models that range from abstract, principled models to models of data. Figure 1 shows an example of how different kinds of models slice the world. At the top of Fig. 1, principled models are highly abstract theories that characterize a broad perspective on a phenomenon [30]. At the bottom, a dotted line separates ontologically different spheres governed by different mechanisms. Below the line are physical and social worlds, each with their own constructive mechanisms [81]. Above the line are abstractions, representations, and information about the world. It is crucial to note that in Fig. 1, models at different levels are not in a pure inferential relation to each other: hierarchically lower-level models cannot be deduced from higher-level ones, and higher-level ones cannot be inductively generalized from lower-level ones [30]. Statements on the level below do not follow from principled models; the models at different levels are of logically different types. Instead of establishing a deductive or inductive relation, a researcher uses higher-level models to guide the formal structure of lower lever models, with an intention to find a good fit between possibly different types of models at different levels of abstraction [28].

To give an example of Figure 1, take Papert's constructionism [68]. Constructionism, as described in *Mindstorms*, can be interpreted as a *principled model*, which characterizes learning as construction of knowledge using manipulatable materials, but which alone is not well suited for making empirical claims, and not straightforwardly testable [69]. By specifying a range of conditions, specifications, actors, and constraints, it is possible to generate a broad range of representational models of real phenomena about learning. One example *representational model* for Fig. 1 is to represent learning activity in terms of specific activities of solving authentic problems in information-rich environments, where learners need to construct solutions on their own. *Specific hypotheses* in Fig. 1 can be, for instance, those tested in Kirschner et al.'s [41] oft-cited critique of unguided learning: "the hypothesis [is] that people learn best in an unguided or minimally guided environment where they are primarily asked to mimic the problem-solving activities of experts and/or learn and discover collaboratively with others." For *models of experiments and data*, take, for example, studies on the use of Logo in the classroom, which commonly employ, for instance, simple posttest-only design with non-equivalent control groups [76], with interventions such as an hour of Logo programming each week.

The shift to a model-based view of science has resulted in a sizable body of literature on the functions and types of models. That literature recognizes, for instance, computational models, phenomenological models, formal

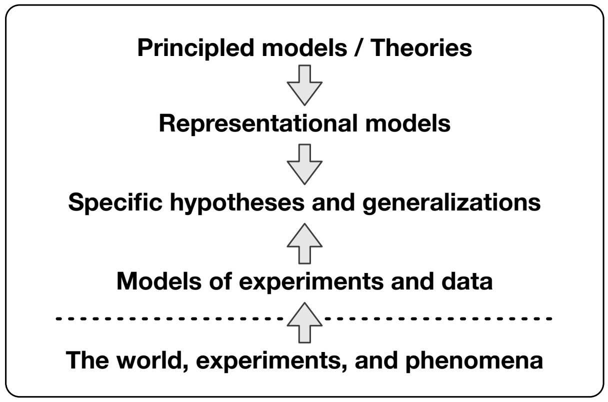

Fig. 1. One example of hierarchical relationships between models and the world (adapted from one of the several variations and examples in Giere [30])

models, and heuristic models—among dozens of others [26]. The model-based view has also breathed new life into analysis of the cognitive functions that models carry [26]. For instance, much of the time, researchers tinker with models of various kinds and not directly with reality, and, by doing that, engage in model-based reasoning [48]. Models can be used to explain even when they are not "true" [8] (as Box [5, p.424] quipped, all models are wrong, but some are useful). In Figure 1, loose representational models can play the explanatory role, and fit generalizations into a high-level theory (cf. [8]).

In CER a shift towards a model-based view of science would not be aimed at rendering the concept of theory obsolete. The shift would be aimed at broadening the idea of what kinds of entities researchers really work with [26]. Models are aimed not just at prediction, description, and explanation, but they can also help to understand phenomena, visualize, interpret, mediate, and probe. They are tools for concept formation, heuristics, and pedagogy. They can be minimal or rich, instrumental or idealized, mathematical or schematic. Models can be very useful even when they are much less developed as theories, and being much more flexible, they can relieve the CER discussions of the undue historical and conceptual baggage of theories. A focus on models encourages fitting together things on different levels of abstraction—fitting together aims, values, facts, practical actions, data collection, and theoretical entities. Grand theories always remain, for a good reason, but a shift toward models may bring into focus the kinds of entities researchers work on most of the time.

3 APPLYING EXTERNAL THEORY TO CER

In the absence of a field's homegrown theory, it is wise to investigate what others have done in similar situations [62]. For instance, mathematics education research has extensively discussed the roles of theory in the field [86], including theories "borrowed" from other fields, adapted, or developed into new "home-grown" theories particular to the discipline [85]. But in discipline-based education research (DBER) fields, the roles of learning theories and principles vary widely by field, and all DBER fields are tightly connected with their parent disciplines, complicating direct borrowing of theories and results [62, pp.52–53,188]. Still, where CER lacks theory, a neighboring field or a related field is a natural place to look for. The analogies needed to do that may be very near or far, they may prove to be useful and effective, or they may be ill-fitting and fail. For example, theories from cognitive psychology can help designers think about how to direct learners' attention to critical components of the learning material [88]. Theories about developmental stages can help understand what reasoning skills can be expected at different grades in school [88]. But the adoption or adaptation of theories from other fields requires understanding of context: the challenges of learning differ between learning multiplication tables, methods of proof, chemical reactions, playing the piano, musical composition, languages, or programming [62]. This section presents a number of concepts for analysis of analogies, theory borrowing, and strength of theory.

3.1 Straightforward Applicability of Theories Big and Small

Many theories used in CER are sweeping, high-level grand theories. Szabo et al.'s [89, pp.95–96] table of the most popular theories in CER contains, for instance, the zone of proximal development, Maslow's hierarchy of needs, and Papert's constructionism. Malmi et al.'s [52] review of 308 articles in CER from 2005 to 2011 found that the most common theory, model, or framework in their sample was constructivism or some of its variants (such as constructionism), followed by curricular frameworks, the pair programming model, and Bloom's taxonomy as the fourth. The situation is similar to broader discipline-based education research, where constructivism and sociocultural learning perspectives are common [62].

Oft-cited famous works like Papert's constructionism are good tools for thinking about how to organize learning, but at the same time they suffer from the typical problems associated with grand theories [55, 57]. As high-level umbrella theories, they are separate from the specific, typical empirical concerns of education researchers and developers (cf. [57, pp.25–49]). They have overly broad scope and lack straightforward applicability (cf. [55]). For development of learning tools, they have only a limited ability to solve specific design problems or help with interaction design.

Szabo et al. [89] found Papert's constructionism cited in 350 CER studies. Constructionism emphasizes learners' engagement with artifacts, messing about with science, and metacognitive development, and Papert [68] presented it in a tour de force that touched on countless more specific theories. The theory's grand theory nature was made clear by Papert himself: Papert held that trying to experimentally control for the "Logo variable" is fruitless [69, 77]. From the start, constructionism was the theory of choice for freewheeling builders of artifacts: "They just gave Logo to a bunch of kids, and after a while they asked them how they liked it. 'Gee, it's terrific!"' (cited in [3]). Yet, the claims about constructionism's positive effects on learning relied on mostly anecdotal evidence [3], and even the gradually accumulating evidence to the contrary [66] did not initially slow down the movement (e.g., [69]). Later empirical research has criticized some cornerstones of constructionism—such as minimally guided instruction and discovery learning [41]—but constructionism continues to be popular in CER, yet it perhaps is used more as a high-level design philosophy or principled model than a specific operationalizable theory or representational model.

For another example, take Bloom's taxonomy, another very common framework in CER [52]. The most widespread visualization of Bloom's taxonomy is the famous pyramid with lower-level educational objectives (e.g., remembering, understanding) at the bottom and higher-level ones (e.g., evaluating, creating) at the top. While that taxonomy can be useful for thinking about learning objectives, it does not serve the role(s) many advocates of a stronger theory base for CER prefer: it was never aimed at explaining or predicting phenomena, and its layered form suggests a hierarchy that has no basis in cognitive psychology [14, pp.28–32]. But while the pyramid version of Bloom's model is not useful for classifying reality, it serves well its original purpose—to provide a classification for representing educational objectives for the purposes of communicating learning goals and discussing curriculum design.

In addition to grand theories—which can be very useful for inspiration and high-level design of learning interventions—CER uses many theories from fields such as psychology, cognitive science, and learning sciences [89]. But the popularity of a theory does not mean it can be unproblematically applied to CER. Take, for instance, three theories popularly used in CER: the learning styles theory (1390 CER articles), grit theory (54 articles), and the mindset theory (633 articles). Each of them has its proponents, and some have solid evidence behind them, but each has also been criticized by meta-analyses: learning styles, since four decades ago (e.g., [10]); grit, after it became famous (e.g., [12]); and mindset, once government institutions started to notice it (e.g., [84]). Due to how they have been used and misused, and not necessarily the theories themselves, each of the theories has been listed in De Bruyckere, Kirschner, and Hulshof's two-volume popular science collection of urban myths about learning [13, 14].

The problem with applicability of theories from outside the field is known in all applied fields. A theory that was developed in field S to represent W for the purposes P may do well in its original context [see 29], but changing any of the S, W, or P may partly or completely change the construct's epistemological, ontological, technical, sociocultural, and all other frames of reference. Phenomena, processes, and interactions that were missing in the original setting may completely change the outcomes in the new setting. For example, a higher-order composite concept that works for the purpose of predicting success in completing military training may not be straightforwardly applied to predicting course marks in a programming course (as shown by [82]). But in addition to the practical problem of applicability, the philosophy of science has pointed out a more fundamental problem with theory: underdetermination of theory by data, discussed next.

3.2 Underdetermination and the Attribution Problem

Learning is at once one of the most ordinary things, one which everyone has personally experienced, and one of the most complex phenomena to study. One can study learning processes at any number of levels of abstraction, from synapses and neurotransmitters to the sociocultural practices of communities, and often those levels overlap [78]. One can study how learning shapes one's brain, how people grow as professionals, or how affective support influences learning (cf. [1]). Theories and methods applicable to each level are fundamentally different from each other. What is more, many learning-related processes are emergent: processes at one level are not reducible to, or deducible from, processes on the levels below: Most famously, consciousness is typically considered to be an emergent property of the brain, not an attribute of neurons [80]. Learning is not a single process, but a complex of complicated processes from biological to sociocultural, not explainable by theories at a single level of abstraction.

In the mid-1900s the logician W.v.O. Quine [73] articulated a major weakness in the contemporaneous dominant accounts of natural science—a weakness that has turned out to be especially pronounced in research on learning. This weakness is identified by the Duhem-Quine thesis, which states that even with simple phenomena in natural sciences—not to mention complex processes like learning—an experiment can never conclusively reject a theory, because it is always possible that some other part of the complex test situation than the theory is responsible for the unexpected results. The underdetermination problem is not just speculation; underdetermination of theory by data is a common problem in scientific practice [94]. In a field that studies learning, such as CER, multiple levels of analysis, confounding variables, materiality, and a thicket of sociocultural, cognitive, and pedagogical influences make it often impossible to judge which combination of theories would properly explain the results.[1]

In CER, variants of the Duhem-Quine thesis manifest as attribution problems, such as: To what extent can observed changes in learning be attributed to theory-derived elements of a learning intervention or to the theory itself? How reliably can observed changes in learning be attributed to one particular theory and not another, or to the myriad other possible factors at play? How feasible is it to link interface elements, learning environment designs, or other design choices with high- or low-level theories of learning? How reliably can one establish a causal link between two sets of observations at different levels? The more abstraction levels there are between two building blocks of a study—take eye movement data and examination scores, for example—the more acute the attribution problem grows. In CER, the sheer number of biological, psychological, social, and cultural processes at play in any learning situation greatly limits the predictive and descriptive power of isolated theories, making it

[1]In the philosophy of education research, it has been suggested that underdetermination is not well suited to education research because the field is fundamentally different from natural science: "What currently [in 1991] pass as theories in the human sciences are actually ideological frameworks" [56], and ideological frameworks are not even the kind of a thing that can be underdetermined by evidence. (But that citation is more related to grand theories than specific theories of learning.)

difficult to attribute effects downstream to changes upstream. That is not to say that learning-related phenomena in CER would use a different cognitive architecture, or otherwise be unreachable by learning-related theory from other fields. There are many examples of the benefits of theory on learning designs in computing. The attribution problems just emphasize that experiments on new designs for learning can never be isolated to just one theory; any change in a learning situation affects the complex of processes.

In the philosophy of science, the issue of underdetermination has been extensively discussed over the past century, and standpoints vary as to how strong a position one should hold. Underdetermination, the Duhem-Quine thesis, and the attribution problems do not necessarily state that if there is no way to conclusively tell if theory was right or wrong, then theories are useless, or that borrowing theory from other fields is hopeless. They do not need to lead to strong relativism (i.e., under some circumstances, all theories are as good as all others) or epistemological anarchism ("anything goes" [21]).

A cautious position towards underdetermination that relies on Quine [73] does not reject the importance of theory and experiment, but it does undermine the idea of "crucial experiments" that can confirm or condemn a theory in one fell swoop. Experiments do not test a single hypothesis but a whole system of theories and hypotheses and the world—and that has very real implications. For example, when the results of a study in programming education contradict the results predicted by a theory of learning, much more than a hypothesis, theory, and observations is at play. *Something* did not work out as planned: the world resisted [71]. It is the researcher's job to find a working fit among theories, instruments, learning situations, material and content, and all other pieces of the puzzle [71]. Similarly, a research study in CS1 that fails to find an effect expected by field F's theory T does not invalidate T. By the Duhem-Quine thesis, one cannot tell which element(s) of the whole test situation caused the unexpected results.

CER researchers and learning designers need not give up in the face of the daunting complexity of phenomena that accompany every learning situation. Instead, rather than relying on an oversimplified idea of the theory–observation relation, superfluousness of variables, or foundational epistemology, a researcher can embrace a less stringent idea of how science proceeds. Proofs in the logical sense are not the right building block for empirical sciences, and ampliative reasoning does not commit one to irrationalism or universal relativism. Observational sentences are not deductive consequences of the sentences of abstract theories (or the other way around), yet aiming to fit theories with observations is still important, as long as one keeps in mind that the fitting depends on multiple levels of complex processes. Most importantly, demands for strengthening the theoretical foundations of CER should account for the epistemological challenges posed by underdetermination, the Duhem-Quine thesis, and attribution problem(s).

4 CER AS A DESIGN FIELD

Many disagreements over the role of theory in CER seem to stem from lack of consensus over the nature of the field. Those who have characterized CER as an experiment-based empirical science have evoked concepts, nomenclature, and relationships between concepts from the standard literature in the philosophy of (natural) science. Studies on the topic cite Popper's falsificationism [22, 63], Kuhn's paradigm theory [22, 23], Feyerabend's anarchistic theory of science [37], positivism [92], or Hempel's deductive-nomological model [22]. Those accounts of science involve four classical aims of science: to explore (develop an initial understanding of an uncharted phenomenon), to describe (systematically record and model a phenomenon and its relationships to other phenomena), to predict (use existing knowledge to predict phenomena that have not yet come to pass), and to explain (clarify the causes, relationships, and consequences of a phenomenon).

But there are also those who see the field as primarily a design and engineering field. Many hold that the aim of CER is not primarily that of building descriptive or predictive theories but something else—such as improved designs for learning [63], innovation [54], or development of tools for learning [49]. Of Fincher and Petre's [22] ten areas that motivate research in CER, many are about developing and evaluating educational interventions for computing education, be they technical or pedagogical. A survey of ACM SIGCSE papers between 1983 and 2003 found that 22% of papers presented tools for learning, and the sample identified more “tools” papers than any other type in 1994–2003 [93]. A series of studies by Simon (reported in [83]) found similar results in various other CER publishing venues.

There are crucial paradigmatic differences between those two different views of the nature of CER. The status of theory differs between sciences, design/engineering fields, social sciences, and humanities. Research papers that focus on development of artifacts—software, hardware, or learning environments, for instance—are not expected to conform to ideals of theory-driven research presented by philosophers of (natural) science mentioned above. The most suitable reference for design and engineering in CER is not the philosophy of science but the philosophy of engineering and the philosophy of technology. Philosophers of engineering and technology have shown that the essence of design and engineering lies in their aims: to develop tools or artifacts that accomplish classes of tasks more efficiently (e.g., [25, 58, 95]). Carl Mitcham, a prominent philosopher of technology wrote that while engineers do apply theories from other fields, “artifact design is what constitutes the essence of engineering, because it is design that establishes and orders the unique engineering framework that integrates other elements” [58, pp.146–147]. In engineering, theory is not the goal; theory is subservient to the engineering aims.

Instead of theories, much of engineers’ knowledge is expressed as technical maxims, state-of-the-art solutions, and guiding principles—and similar to how Tenenberg and McCartney [91] characterized CER, that knowledge is tentative, conditional, and contextual [58, 95]. Designers and engineers achieve their aims by following constructive and descriptive methods that aim at achieving change in the affairs of the world [42]—such as developing and testing explainable AI, new ways to facilitate lifelong learning in computing, tools that enable teacher trainees to understand machine learning concepts, or tools that substantively change how collaborative learning happens in programming. As contrasted with natural sciences, the engineering method is described as “the use of heuristics to cause the best change in a poorly understood situation within the available resources” [42]. Engineers and designers often have to get things done relying on information that scientists would consider inadequate for research purposes (cf. [95]).

Yet tinkering, toying, or just building things is neither engineering nor design. The engineering-specific knowledge base that technology and engineering researchers apply in their work comprises state-of-the-art concepts, heuristic prescriptions, best practices, and procedural knowledge of what works, embodied in 1) procedural knowledge (“know-how”), 2) technical maxims, rules of thumb, heuristic strategies, or “recipes,” 3) descriptive experience-based “If A then B”-type laws, and 4) technological theories on how to apply scientific theory to practice [58].

The field of computing has a wealth of examples on how to communicate design knowledge—take human-computer interaction or software engineering, for example. Design patterns in software engineering [27] are a way to communicate heuristic strategies to work out common situations object-oriented programmers encounter. Design hints in operating systems [46] are a way to communicate procedural knowledge and technical maxims at the level of complete systems. And design principles in information protection [79] are a way to communicate rules of thumb about how to reduce the risk of compromising sensitive information. It has been suggested that in design, theories best inform design as counterfactuals, through thought experiments of the kind “If design was <like this>, then interaction would be <like that>” [65]. In those counterfactuals, theories can direct design choices, help designers choose between design choices, and expose new design spaces [65].

One central reason for the separation of the philosophy of engineering, technology, and social sciences from the philosophy of (natural) science has to do with the different ontological views of the fields they study. The subjects of physical sciences—such as properties of particles in fields of force—are mind-independent [81]. The subjects of social sciences—such as preferences, behaviors, and mental states—are mind-dependent but are separate from the researcher [81]. The subjects of engineering are human-made artifacts, which have to cater to the laws of nature, to value and worth in the human experience [11], and to a range of values from budgets and schedules to fashion statements. Insofar as one of the most common aims of CER is to develop artifacts to help the learning process in computing [49], those studies need to be judged in the framework of philosophy of engineering and technology and not by the standards developed in the philosophy of (natural) science. And insofar as CER is concerned with socially constructed phenomena—take collaborative learning, sociocultural practices of learners, or inequality, for example—then the philosophy of social sciences [16, 17, 61] offers other fruitful starting points for analyzing the role of theory in CER.

4.1 Experiences from Computing as a Discipline

The CER discussions related to the role of design and engineering in the field closely resemble computing's disciplinary debates when the field was a nascent, still-emerging discipline [90]. In the 1970s, software engineers were seen as builders of tools, relying on a weak scientific base. In the 1980s "experimental computer science" debates [18], a number of prominent people in the field wanted to see computing develop much more in an "experimental" direction, yet "experimental" meant different things to the discussants [90]. Some discussants were quick to point out that Turing Awards—the field's highest distinction—are frequently given to technology pioneers for their contributions to technology. Technology matters. The schism between software engineering, theoretical computer science, and experimental computer science characterized the field's soul-seeking efforts for decades, until a more sober view, presented in the famous report *Computing as a Discipline* [15], won public opinion. It described the field as an alliance between the field's theoretical, engineering, and scientific aspects, irrevocably intertwined. Over decades of disagreement, there grew a consensus that it would not do justice to the field to ignore the engineering branch that is intellectually challenging, societally important, and central to the field's most important achievements [90].

It has been argued that some CER arenas have started to systematically dismiss pure technological and engineering contributions in favor of studies that are hypothesis-driven or theory-driven ([83, p.17], [24]). Simon [83] showed that the reception of engineering contributions differs by the CER forum and gatekeeper—for instance, he pointed out that soon after its rebranding from JERIC to TOCE, ACM's prime journal on computing education research announced a clear shift away from engineering/technology articles [83, pp.64,80]. Some pioneers of CER have called for a clear demarcation of CER proper, in order to exclude topics such as the development of pedagogical environments, descriptive classroom studies, and building of learning tools devoid of a theory element [31, p.229]. This is a common development, familiar from other fields, too: as fields grow and mature, specializations develop within them. Similar patterns can be seen in many fields of computing, too, with some publishing venues focusing on theoretical contributions and others on technological ones [90].

In the CER theory debates, technological and engineering papers play a similar role to that which systems, software engineering, and some algorithmic innovation papers played in computing's disciplinary debates. Just as Fredrick P. Brooks Jr. stated in his ACM Allen Newell Award talk that the computer scientist is a toolsmith who studies in order to build [6], so did Nelson and Ko [63] write that computing education researchers draw upon theoretical work to develop designs to improve learning. In their 1975 Turing Award speech, Allen Newell and Herbert A. Simon went much further than the CER debates ever went when they wrote, "Each new machine that is built is an experiment. ...Each new program that is built is an experiment. It poses a question to nature" [64]. For Newell and Simon, building computers and programs were ways to discover new phenomena, and their gains sometimes paid off in the permanent acquisition of new technology. Perhaps the same can be said about artifact development in CER.

Improving the quality of empirical and theory-driven research in CER does not need to come at the cost of diminishing CER's other important traditions: CER can be seen as an empirical, experiment-based science, as an engineering and design field, as a social science, and as a branch of educational research, among other views. It is not unreasonable to expect that the development of learning tools benefits from a solid theoretical understanding of learning. But to ignore the pure technological and engineering side of CER would sever the field from its engineering tradition, it would separate the CER field from an important driver of its results, and it would not do justice to the technologically rich history of the CER field (for that history, see [33, 53]).

5 DISCUSSION

This essay has proposed that many equivocations that bedevil discussions on the role of theories in CER arise largely from the adoption of a "received view" of what theory is [34] and from a simplified view of theory's role in scientific explanation. "The models are broken, the models are broken!" would Allen Newell exclaim.

The "classical" models of scientific explanation—such as Popper's [72], Kuhn's [43], or Lakatos's [45] accounts of science—were designed with natural sciences, not arts, humanities, social sciences, or education research, in mind. Even if a procedure for conducting science works in natural sciences, it does not follow that it works in CER, a field fundamentally different in kind. (It does not follow that it would *not* work, either.) What is more, each model has been later shown to be inadequate even in natural sciences (e.g., [20]). Disciplinary discussions that cite the classical models of science should also take into account their shortcomings, their original context, and changes in scientific practice in the past 60–100 years. Logical positivism reigned a hundred years ago, Popper presented his influential view in the 1930s, and Kuhn's game-changing book was published in the 1960s.

Many models of scientific research mentioned in CER's theory discussions are a priori models, formulated in specific intellectual and historical contexts as ideals of how natural science should ideally be done. Many of those models were developed without getting much involved in the messiness of the real world when formulating methodological prescriptions for science—take Popper's, Hempel's, and Carnap's accounts of science, for instance. Their views, pioneering as they were, have been shown to be based on preconceived, idealistic ideas of science that insufficiently capture how (especially 2020s) science really works in practice [36, 43, 71] or how it should work [21].

As opposed to the earlier a priori models, towards the end of the 1900s, the naturalistic turn of the philosophy of science started to shift the field from the shibboleths of analytically oriented philosophers towards an image of science deeply rooted in the reality of scientific practice. Historians of science looked deeper into the archives and found controversy, obstinacy, and irrationalism [21, 43]. Psychologically oriented studies found ways in which many seasoned scientists neither adhere to the idealized norms and ethos of science nor believe in them (cf. [59]). Sociologically oriented philosophers opened the lab door and painted fresh, new images of how successful scientists actually work [36, 71, 94]. Naturalistic epistemology is an extension of sociology and psychology, and it treats construction of scientific knowledge as an empirical matter instead of a normative and a priori matter. Like other fields of research [20, 21, 59], CER may not be immune to many problems a priori models ignore. Discredited theories may go on for decades [32, pp.46–50], intuitively appealing folk theories may prevail despite any amount of evidence to the contrary [3], and organized dogmatism, bias, and self-interest may sometimes surface [3]. This essay suggests that it would be fruitful for the philosophy of computing education research to anchor the image of CER to naturalistic epistemology and to the naturalistic turn of the philosophy of science.

This essay presents that by shifting more towards a model-based view of science, CER can avoid some problems that arise from centuries of conceptual baggage related to the concept of theory. Compared to theory, model is a much broader and more accommodating concept, and it does not give rise to as many expectations and demands as theory does (even if those arise from misunderstanding the role of theory). For example, no one expects a didactic model be verified or proven correct [sic]. No one thinks it has to be universally true. No one expects it to be tested across all types of learning situations and circumstances [40]. Models, by virtue of their very nature, are not assumed to be "true." The demand to anchor findings in a theoretical framework somehow sounds much more intimidating than the demand to relate the findings to models related to learning—yet the actual objects can be the same.

This essay also suggests that a different perspective on the role of theory is necessary for those parts of CER that are oriented towards improved or new (artifact) designs for learning. Compared to the philosophy of science, the philosophy of engineering and the philosophy of technology provide better-suited analyses of the role of theory in the development of software or hardware artifacts. Artifact design has a rich and important history in CER, and it should be judged in terms of its own intellectual tradition instead of that of hypothesis-driven science. And compared to the philosophy of science, the philosophy of social science frames theories in ways more amenable for the logic of certain types of research on the behavior of individuals in various communities.

This essay is aimed at providing alternative perspectives to theory discussions in CER, and as such, it is limited in several important ways. It specifies only a very limited number of models, objects of interest, and their relationships. It covers only a few select perspectives of models from the many viewpoints that the philosophers of science have presented. It uses many fewer examples from CER than one would hope: the authors work in narrow sectors of education. In its advocacy of a model-based view of science, it gives much less attention than it could to the enormous body of research on the structure of scientific theories. It does not discuss the nature of products of research—and whether they are of logical type, descriptive statements, recommendations, or something else. It does not use results from the philosophy of computing education research or from the broader philosophy of computer science education. And because the strengths of theory-driven computing education research and technology development are already well covered by numerous other authors in CER and broader discipline-based education research, the essay probably fails to give enough credit to the fact that theory (of *some* kind) is often at the heart of excellence in research.

Despite its many shortcomings, we hope the essay gives rise to discussions about whether even a moderately naturalistic, model-based approach to the "theory of theory" would pay off in terms of better fit between CER theory discussions and how practicing researchers in CER see their own work. The essay suggests, not in a Quinean strand of naturalism, but perhaps in a Kuhnian strand of naturalism, that we take seriously what happens in CER in action, even when researchers do not adhere to a physics textbook depiction of theory fit in research.

Just as Kuhn's major contribution was to reveal the importance of social structure of science communities, scientists, and their relative agreement, it would be an important contribution to CER to study the generally accepted landmark studies of the CER field and analyze how those studies *actually* use theory and models (instead of how philosophers of natural science prescribe they should be used). For instance, one can study the field's most cited articles, or the most impactul ones, or perhaps the winners of best paper awards in the field's major conferences. That would help build a consensus of CER's paradigm and describe the field's take on theory in CER's own terms. One aspect to keep in mind in that undertaking is not to let an unwarranted preconception of science be the criterion of good research.

6 ACKNOWLEDGMENTS

Matti Tedre thanks January Collective for their meticulous but constructive critique on this manuscript.

REFERENCES

[1] Patricia A. Alexander, DIane L. Schallert, and Ralph E. Reynolds. 2009. What Is Learning Anyway? A Topographical Perspective Considered. *Educational Psychologist* 44, 3 (2009), 176–192. https://doi.org/10.1080/00461520903029006

[2] Vicki L. Almstrum, Orit Hazzan, Mark Guzdial, and Marian Petre. 2005. Challenges to Computer Science Education Research. *SIGCSE Bulletin* 37, 1 (Feb. 2005), 191–192. https://doi.org/10.1145/1047124.1047415

[3] Morgan G. Ames. 2018. Hackers, Computers, and Cooperation: A Critical History of Logo and Constructionist Learning. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 18:1–18:19.

[4] Mordechai Ben-Ari, Anders Berglund, Shirley Booth, and Christian Holmboe. 2004. What Do We Mean by Theoretically Sound Research in Computer Science Education? *SIGCSE Bulletin* 36, 3 (June 2004), 230–231. https://doi.org/10.1145/1026487.1008059

[5] George E. P. Box and Norman Richard Draper. 1987. *Empirical Model-Building and Response Surfaces.* John Wiley & Sons, Inc., New York, NY, USA.
[6] Frederick P. Brooks, Jr. 1996. The Computer Scientist as Toolsmith II. *Commun. ACM* 39, 3 (1996), 61–68.
[7] Rudolf Carnap. 1937. *The Logical Syntax of Language.* Kegan Paul, Trench, & Trübner, London, UK.
[8] Nancy Cartwright. 1983. *How the Laws of Physics Lie.* Oxford University Press, Oxford, UK.
[9] Mike Clancy, John Stasko, Mark Guzdial, Sally Fincher, and Nell Dale. 2001. Models and Areas for CS Education Research. *Computer Science Education* 11, 4 (2001), 323–341. https://doi.org/10.1076/csed.11.4.323.3827
[10] Richard E. Clark. 1982. Antagonism Between Achievement and Enjoyment in ATI Studies. *Educational Psychologist* 17, 2 (1982), 92–101. https://doi.org/10.1080/00461528209529247
[11] Gilbert Cockton. 2006. Designing Worth is Worth Designing. In *Proceedings of the 4th Nordic Conference on Human-Computer Interaction: Changing Roles (NordiCHI '06).* Oslo, Norway, 165–174.
[12] Marcus Credé, Michael C. Tynan, and Peter D. Harms. 2017. Much ado about grit: A meta-analytic synthesis of the grit literature. *Journal of Personality and Social Psychology* 113, 3 (2017), 492–511.
[13] Pedro De Bruyckere, Paul A. Kirschner, and Casper Hulshof. 2015. *Urban Myths about Learning and Education.* Academic Press, London, UK.
[14] Pedro De Bruyckere, Paul A. Kirschner, and Casper Hulshof. 2020. *More Urban Myths about Learning and Education: Challenging Eduquacks, Extraordinary Claims, and Alternative Facts.* Routledge, New York, NY, USA.
[15] Peter J. Denning, D. E. Comer, David Gries, Michael C. Mulder, Allen Tucker, A. Joe Turner, and Paul R. Young. 1989. Computing as a Discipline. *Commun. ACM* 32, 1 (1989), 9–23.
[16] Norman K. Denzin and Yvonna S. Lincoln (Eds.). 2018. *The SAGE Handbook of Qualitative Research* (5th ed.). SAGE, Thousand Oaks, CA, USA.
[17] Brian Fay. 1996. *Contemporary Philosophy of Social Science.* Blackwell Publishing, Malden, MA, USA.
[18] Jerome A. Feldman and William R. Sutherland. 1979. Rejuvenating Experimental Computer Science: A Report to the National Science Foundation and Others. *Commun. ACM* 22, 9 (1979), 497–502.
[19] James H. Fetzer. 1999. The Role of Models in Computer Science. *Monist* 82, 1 (1999), 20–36.
[20] Paul Karl Feyerabend. 1970. Consolations for the Specialist. In *Criticism and the Growth of Knowledge*, Imre Lakatos and Alan Musgrave (Eds.). Cambridge University Press, London, UK, 197–230.
[21] Paul Karl Feyerabend. 1975. *Against Method* (1st ed.). Verso, London, UK.
[22] Sally Fincher and Marian Peter (Eds.). 2004. *Computer Science Education Research.* Taylor & Francis, Lisse, The Netherlands.
[23] Sally Fincher and Josh Tenenberg. 2006. Using Theory to Inform Capacity-Building: Bootstrapping Communities of Practice in Computer Science Education Research. *Journal of Engineering Education* 95, 4 (2006), 265–277. https://doi.org/10.1002/j.2168-9830.2006.tb00902.x
[24] Sally A. Fincher, Josh Tenenberg, Brian Dorn, Christopher Hundhausen, Robert McCartney, and Laurie Murphy. 2019. Computing Education Research Today. In *The Cambridge Handbook of Computing Education Research*, Sally A. Fincher and Anthony V. Robins (Eds.). Cambridge University Press, Cambridge, UK, 40–55.
[25] Samuel C. Florman. 1994. *The Existential Pleasures of Engineering* (2nd ed.). St. Martin's Press, New York, NY, USA.
[26] Roman Frigg and Stephan Hartmann. 2020. Models in Science. In *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (Ed.). Stanford University. http://plato.stanford.edu/archives/spr2003/entries/moral-epistemology/
[27] Erich Gamma, Richard Helm, Ralph Johnson, and John Vlissides. 1994. *Design Patterns: Elements of Reusable Object-Oriented Software.* Addison-Wesley, Boston, MA, USA.
[28] Ronald N. Giere. 1999. *Science Without Laws.* The University of Chicago Press, Chicago, IL, USA.
[29] Ronald N. Giere. 2004. How Models Are Used to Represent Reality. *Philosophy of Science* 71, 5 (2004), 742–752. https://doi.org/10.1086/425063
[30] Ronald N. Giere. 2010. An agent-based conception of models and scientific representation. *Synthese* 172, 2 (2010), 269–281. https://doi.org/10.1007/s11229-009-9506-z
[31] Michael Goldweber, Martyn Clark, Sally Fincher, and Arnold Pears. 2004. The Relationship between CS Education Research and the SIGCSE Community. *SIGCSE Bulletin* 36, 3 (2004), 228–229.
[32] Mark Guzdial. 2015. *Learner-Centered Design of Computing Education: Research on Computing for Everyone.* Morgan & Claypool, San Rafael, CA, USA.
[33] Mark Guzdial and Benedict du Boulay. 2019. The History of Computing Education Research. In *The Cambridge Handbook of Computing Education Research*, Sally A. Fincher and Anthony V. Robins (Eds.). Cambridge University Press, Cambridge, UK, 11–39.
[34] Ian Hacking (Ed.). 1981. *Scientific Revolutions.* Oxford University Press, Oxford, UK.
[35] Ian Hacking. 1983. Experimentation and Scientific Realism. *Philosophical Topics* 13, 1 (1983), 71–87.
[36] Ian Hacking. 1983. *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science.* Cambridge University Press, New York, NY, USA.

[37] Christian Holmboe, Linda McIver, and Carlisle George. 2001. Research agenda for computer science education. In *Proceedings of the 13th Workshop of the Psychology of Programming Interest Group*, G. Kadoda (Ed.). PPIG, 207–223.

[38] Ville Isomöttönen. 2011. *Theorizing a One-Semester Real Customer Student Software Project Course*. Ph.D. Dissertation. Faculty of Information Technology of the University of Jyväskylä, Jyväskylä, Finland.

[39] Mike Joy, Jane Sinclair, Shanghua Sun, Jirarat Sitthiworachart, and Javier López-González. 2009. Categorising Computer Science Education Research. *Education and Information Technologies* 14, 2 (2009), 105–126.

[40] Jeremy Kilpatrick. 2010. Preface to Part I. In *Theories of Mathematics Education: Seeking New Frontiers*, Bharath Sriraman and Lyn English (Eds.). Springer, Berlin / Heidelberg, Germany, 3–5.

[41] Paul A. Kirschner, John Sweller, and Richard E. Clark. 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 (2006), 75–86.

[42] Billy Vaughn Koen. 2003. *Discussion of the Method: Conducting the Engineer's Approach to Problem Solving.* Oxford University Press, Oxford, UK.

[43] Thomas S. Kuhn. 1962. *The Structure of Scientific Revolutions* (1st ed.). University of Chicago Press, Chicago, USA.

[44] Thomas S. Kuhn. 1970. Reflections on My Critics. In *Criticism and the Growth of Knowledge*, Imre Lakatos and Alan Musgrave (Eds.). Cambridge University Press, London, UK, 231–278.

[45] Imre Lakatos and Alan Musgrave (Eds.). 1970. *Criticism and the Growth of Knowledge.* Cambridge University Press, London, UK.

[46] Butler W. Lampson. 1983. Hints for Computer System Design. *SIGOPS Operating Systems Review* 17, 5 (1983), 33–48.

[47] Alex Lishinski, Jon Good, Phil Sands, and Aman Yadav. 2016. Methodological Rigor and Theoretical Foundations of CS Education Research. In *Proceedings of the 2016 ACM Conference on International Computing Education Research* (Melbourne, VIC, Australia) *(ICER '16)*. ACM, New York, NY, USA, 161–169. https://doi.org/10.1145/2960310.2960328

[48] Lorenzo Magnani, Nancy J. Nersessian, and Paul Thagard (Eds.). 1999. *Model-Based Reasoning in Scientific Discovery.* Springer, New York, NY, USA.

[49] Lauri Malmi. 2014. Tools Research — What Is It? *ACM Inroads* 5, 3 (2014), 34–35.

[50] Lauri Malmi, Judy Sheard, Päivi Kinnunen, Simon, and Jane Sinclair. 2019. Computing Education Theories: What Are They and How Are They Used?. In *Proceedings of the 2019 ACM Conference on International Computing Education Research* (Toronto ON, Canada) *(ICER '19)*. ACM, New York, NY, USA, 187–197. https://doi.org/10.1145/3291279.3339409

[51] Lauri Malmi, Judy Sheard, Päivi Kinnunen, Simon, and Jane Sinclair. 2020. Theories and Models of Emotions, Attitudes, and Self-Efficacy in the Context of Programming Education. In *Proceedings of the 2020 ACM Conference on International Computing Education Research* (Virtual Event, New Zealand) *(ICER '20)*. ACM, New York, NY, USA, 36–47. https://doi.org/10.1145/3372782.3406279

[52] Lauri Malmi, Judy Sheard, Simon, Roman Bednarik, Juha Helminen, Päivi Kinnunen, Ari Korhonen, Niko Myller, Juha Sorva, and Ahmad Taherkhani. 2014. Theoretical Underpinnings of Computing Education Research: What is the Evidence?. In *Proceedings of the Tenth Annual Conference on International Computing Education Research* (Glasgow, Scotland, United Kingdom) *(ICER '14)*. ACM, New York, NY, USA, 27–34. https://doi.org/10.1145/2632320.2632358

[53] Lauri Malmi, Ian Utting, and A. J. Ko. 2019. Tools and Environments. In *The Cambridge Handbook of Computing Education Research*, Sally A. Fincher and Anthony V. Robins (Eds.). Cambridge University Press, Cambridge, UK, 639–662.

[54] Lauren E. Margulieux, Brian Dorn, and Kristin A. Searle. 2019. Learning Sciences for Computing Education. In *The Cambridge Handbook of Computing Education Research*, Sally A. Fincher and Anthony V. Robins (Eds.). Cambridge University Press, 231–275. https://doi.org/10.1017/9781108654555.010

[55] Robert K. Merton. 1949. *Social Theory and Social Structure.* Free Press, New York, NY, USA.

[56] Steven I. Miller and Marcel Fredericks. 1991. Postpositivistic Assumptions and Educational Research: Another View. *Educational Researcher* 20, 4 (1991), 2–8.

[57] C. Wright Mills. 1959. *The Sociological Imagination.* Oxford University Press, New York, USA.

[58] Carl Mitcham. 1994. *Thinking Through Technology: The Path Between Engineering and Philosophy.* The University of Chicago Press, Chicago, USA.

[59] Ian I. Mitroff. 1974. Norms and Counter-Norms in a Select Group of the Apollo Moon Scientists: A Case Study of the Ambivalence of Scientists. *American Sociological Review* 39, 4 (1974), 579–595.

[60] Nicos Mouzelis. 1995. *Sociological Theory: What Went Wrong?* Routledge, London, UK.

[61] Michiru Nagatsu and Attilia Ruzzene (Eds.). 2019. *Contemporary Philosophy and Social Science: An Interdisciplinary Dialogue.* Bloomsbury Academic, London, UK.

[62] National Research Council. 2012. *Discipline-Based Education Research: Understanding and Improving Learning in Undergraduate Science and Engineering.* The National Academies Press, Washington, DC, USA. https://doi.org/10.17226/13362

[63] Greg L. Nelson and A. J. Ko. 2018. On Use of Theory in Computing Education Research. In *Proceedings of the 2018 ACM Conference on International Computing Education Research* (Espoo, Finland) *(ICER '18)*. ACM, New York, NY, USA, 31–39. https://doi.org/10.1145/3230977.3230992

[64] Allen Newell and Herbert A. Simon. 1976. Computer Science as Empirical Inquiry: Symbols and Search. *Commun. ACM* 19, 3 (1976), 113–126.
[65] Antti Oulasvirta and Kasper Hornbæk. 2021. Counterfactual Thinking: What Theories Do in Design. *International Journal of Human–Computer Interaction* (2021), 1–15. https://doi.org/10.1080/10447318.2021.1925436
[66] David B. Palumbo. 1990. Programming Language/Problem-Solving Research: A Review of Relevant Issues. *Review of Educational Research* 60, 1 (1990), 65–89.
[67] Zacharoula Papamitsiou, Michail Giannakos, Simon, and Andrew Luxton-Reilly. 2020. Computing Education Research Landscape through an Analysis of Keywords. In *Proceedings of the 2020 ACM Conference on International Computing Education Research* (Virtual Event, New Zealand) *(ICER '20)*. ACM, New York, NY, USA, 102–112. https://doi.org/10.1145/3372782.3406276
[68] Seymour Papert. 1980. *Mindstorms: Children, Computers, and Powerful Ideas.* Basic Books, New York, NY, USA.
[69] Seymour Papert. 1987. Computer Criticism vs. Technocentric Thinking. *Educational Researcher* 16, 1 (1987), 22–30.
[70] Arnold Pears, Neena Thota, Päivi Kinnunen, and Anders Berglund. 2012. Harnessing Theory in the Service of Engineering Education Research. In *2012 Frontiers in Education Conference Proceedings.* 1–5. https://doi.org/10.1109/FIE.2012.6462292
[71] Andrew Pickering. 1995. *The Mangle of Practice: Time, Agency, and Science.* The University of Chicago Press, Chicago, USA.
[72] Karl Popper. 1934. *Logik der Forschung.* Mohr Siebeck GmbH & Co., Tübingen, Germany.
[73] Willard Van Orman Quine. 1951. Two Dogmas of Empiricism. *The Philosophical Review* 60, 1 (1951), 20–43.
[74] Hans Radder. 2003. Toward a More Developed Philosophy of Scientific Experimentation. In *The Philosophy of Scientific Experimentation*, Hans Radder (Ed.). University of Pittsburgh Press, Pittsburgh, PA, USA, 1–18.
[75] Justus J. Randolph. 2007. *Computer Science Education Research at the Crossroads: A Methodological Review of the Computer Science Education Research: 2000–2005.* Ph.D. Dissertation. Utah State University, Logan, UT, USA.
[76] Justus J. Randolph. 2008. *Multidisciplinary Methods in Educational Technology Research and Development.* HAMK University of Applied Sciences, Hämeenlinna, Finland.
[77] Lloyd P. Rieber. 1987. LOGO and Its Promise: A Research Report. *Educational Technology* 27, 2 (1987), 12–16.
[78] Anthony V. Robins, Lauren E. Margulieux, and Briana B. Morrison. 2019. Cognitive Sciences for Computing Education. In *The Cambridge Handbook of Computing Education Research*, Sally A. Fincher and Anthony V. Robins (Eds.). Cambridge University Press, 231–275. https://doi.org/10.1017/9781108654555.010
[79] Jerome H. Saltzer and Michael D. Schroeder. 1975. The Protection of Information in Computer Systems. *Proc. IEEE* 63, 9 (1975), 1278–1308.
[80] John R. Searle. 1992. *The Rediscovery of the Mind.* The MIT Press, Cambridge, MA, USA.
[81] John R. Searle. 1996. *The Construction of Social Reality.* Penguin Press, England.
[82] Nikki Sigurdson and Andrew Petersen. 2018. An Exploration of Grit in a CS1 Context. In *Proceedings of the 18th Koli Calling International Conference on Computing Education Research* (Koli, Finland) *(Koli Calling '18)*. ACM, New York, NY, USA. https://doi.org/10.1145/3279720.3279743
[83] Simon. 2015. *Emergence of Computing Education as a Research Discipline.* Ph.D. Dissertation. Aalto University, Finland.
[84] Victoria F. Sisk, Alexander P. Burgoyne, Jingze Sun, Jennifer L. Butler, and Brooke N. Macnamara. 2018. To What Extent and Under Which Circumstances Are Growth Mind-Sets Important to Academic Achievement? Two Meta-Analyses. *Psychological Science* 29, 4 (2018), 549–571. https://doi.org/10.1177/0956797617739704
[85] Bharath Sriraman and Lyn English. 2010. Surveying Theories and Philosophies of Mathematics Education. In *Theories of Mathematics Education: Seeking New Frontiers*, Bharath Sriraman and Lyn English (Eds.). Springer, Berlin / Heidelberg, Germany, 7–32.
[86] Bharath Sriraman and Lyn English (Eds.). 2010. *Theories of Mathematics Education: Seeking New Frontiers.* Springer, Berlin / Heidelberg, Germany.
[87] Frederick Suppe. 1977. *The Structure of Scientific Theories.* University of Illinois Press, Urbana, IL, USA.
[88] Marilla D. Svinicki. 2010. *A Guidebook On Conceptual Frameworks For Research In Engineering Education.* University of Texas, Austin, TX, USA.
[89] Claudia Szabo, Nickolas Falkner, Andrew Petersen, Heather Bort, Kathryn Cunningham, Peter Donaldson, Arto Hellas, James Robinson, and Judy Sheard. 2019. Review and Use of Learning Theories within Computer Science Education Research: Primer for Researchers and Practitioners. In *Proceedings of the Working Group Reports on Innovation and Technology in Computer Science Education* (Aberdeen, Scotland Uk) *(ITiCSE-WGR '19)*. ACM, New York, NY, USA, 89–109. https://doi.org/10.1145/3344429.3372504
[90] Matti Tedre. 2014. *The Science of Computing: Shaping a Discipline.* CRC Press / Taylor & Francis, New York, NY, USA.
[91] Josh Tenenberg and Robert McCartney. 2011. Entry Points for Computing Education Research. *ACM Transactions on Computing Education* 11, 1, Article 1 (Feb. 2011), 5 pages. https://doi.org/10.1145/1921607.1921608
[92] Neena Thota, Anders Berglund, and Tony Clear. 2012. Illustration of Paradigm Pluralism in Computing Education Research. In *Proceedings of the Fourteenth Australasian Computing Education Conference - Volume 123* (Melbourne, Australia) *(ACE '12)*. Australian Computer Society, Inc., 103–112.

[93] David W. Valentine. 2004. CS Educational Research: A Meta-Analysis of SIGCSE Technical Symposium Proceedings. *SIGCSE Bulletin* 36, 1 (2004), 255–259. https://doi.org/10.1145/1028174.971391

[94] Bas C. van Fraassen. 1980. *The Scientific Image.* Oxford University Press, Oxford, UK.

[95] Walter G. Vincenti. 1990. *What Engineers Know and How They Know It: Analytical Studies from Aeronautical History.* The Johns Hopkins University Press, Baltimore / London.

[96] Rasmus Grønfeldt Winther. 2020. The Structure of Scientific Theories. In *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (Ed.). Stanford University. https://plato.stanford.edu/archives/spr2021/entries/structure-scientific-theories/