Computing Education: An Overview of Research in the Field

Tom Crick

April 2017

Background

In May 2016 the Royal Society announced it was commissioning a study\(^1\), funded by Google and Microsoft, to understand the challenges faced by educators delivering computing and computer science and share best practice which can be adopted more widely. The research will establish the progress that has been made since the introduction of the new English computing curriculum (Department for Education, 2013) in September 2014 – as well as recognising related curriculum and qualifications reforms across the rest of the UK – identify areas that still need to be addressed, and will be used as the basis of a wider action plan to transform computing in schools. This new project builds upon the Royal Society’s impactful work in this space, *Shut Down or Restart?* (Royal Society, 2012), a review of ICT and computing education in schools in the United Kingdom (UK). The introduction of a new computing curriculum for 5-16 year-olds in England has, for the first time anywhere, established computer science and computational thinking as foundational subjects alongside English, mathematics and the sciences. England is in effect pioneering a brand new school subject from the age of five\(^2\), with the rest of the world watching to see the impact. This is alongside a number of national and international initiatives to reinforce computer science’s position as a mainstream Science, Technology, Engineering and Mathematics (STEM) discipline (Guzdial and Morrison, 2016).

At the same time, the government is standing back from guiding everyone about how to implement this seismic curriculum change in England, and inviting, employers, universities, professional societies, and educational establishments themselves to play the leading role on how to deliver this curriculum. Academia, industry, parents and educators have responded enthusiastically with a range of initiatives such as code clubs and informal networks for professional development. But there are limits to informal activity; a subject in its infancy needs high-quality teacher training and development, best practice in the classroom, and reliable materials for students; it also needs solid, evidence-based research about what works, proper co-ordination of activity, and effective dissemination of best practice.

Thus, the key objectives which the Royal Society is seeking to understand by commissioning this research project are:

\(^1\)https://royalsociety.org/topics-policy/projects/computing-education/

\(^2\)However, it is important to note that the new national curriculum in England is only mandatory for state-maintained primary and secondary schools; free schools and academies have more freedom and flexibility to decide the curriculum they follow; see here for more information on the types of schools in England: https://www.gov.uk/types-of-school/overview
- What current literature says about computing education, including effective pedagogy and assessment methods;

- Current practice in educational establishments with reference to computing education, including pedagogy; learning goals and content and assessment;

- The attitudes of school leaders and teaching staff towards computing education, including timetabling, teacher confidence, qualifications and motivation as well as engagement of the senior leadership team.

This is a two-stage programme by the Royal Society designed to improve the quality and scale of computing education in English educational establishments, with transferability across the UK and internationally. Stage one will identify and prioritise goals for stage two, which will achieve those goals through a series of separate projects. Stage two projects may include areas such as:

- Producing classroom resources, teacher guidance, and continuing professional development (CPD) programmes;

- Developing effective assessment tools that educators can use to understand and guide progress;

- Publishing guidance about how to address gender imbalance in the uptake of computing;

- Identifying opportunities for project work in schools, perhaps with corporate partners.

This literature review covers the first part of Work Package 1 of the project (*What do we know about pedagogy and assessment in computing?*), providing a comprehensive literature review on effective computing pedagogy.

**Methodology**

There is an emerging corpus of academic and pedagogic literature in computing education, including assessment, attainment and baselining of core digital and technology skills, contextualised by substantial policy reports that transcends education, skills and wider digital economy agendas.

Much of the recent UK-focused literature has been driven by the curricula reforms (and emerging professional practice) in England and Scotland, alongside more recent developments in Wales. There exists active international networks of computer science education researchers (for example, in Germany, US, Israel and New Zealand), providing quantitative and qualitative educational research insight into effective pedagogies, assessment, models for cascading best practices for in-service training and professional development, as well as curriculum changes and wider education policy reform.

The proposed methodology for this literature review was initially based on a systematic review (*Gough et al., 2012*) of the existing corpus of research and policy contributions in this space to identify, appraise, select and synthesise all high-quality research evidence and arguments relevant to the project aims and objectives. Furthermore, it aimed to evaluate
the existing “grey literature” and semi-formal evaluations of practice and experience produced in the UK and internationally, which has required careful categorisation and evaluation of the methodologies and cohort sizes presented. However, during the preliminary phases of the review process, it became apparent that due to the significant quantities of references to pedagogy in computer science in the academic literature, conducting a rigorous systematic review in the relevant time frame would have been problematic. Therefore, multiple methods have been employed to search the literature to ensure a thorough coverage and to minimise the possibilities of omitting any promising research from the review. A number of key academic databases have been targeted, including Google Scholar, ACM Digital Library and IEEE eXplore, followed by searching individual leading international journals and conferences in computer science/ICT/digital/technology education. This traditional literature search of recent English-language literature in refereed educational research journals, using key words and then selecting for relevance to the detailed research questions, produced a selection of some hundreds of papers: these were organised into themes, and in each of these we have attempted to identify a recent meta-analysis or review that we could have confidence in. This survey then suggested some important gaps which we endeavoured to fill through wider searches. Finally, certain areas have been covered to add context to the wider overview of research in this field, but are not comprehensively covered; for example, the section on gender and diversity.

Effective Pedagogies in Computing

Introduction

There is significant international focus on recent and ongoing computing curriculum reform in the UK. A number of audits and studies of national-level curricula models have been conducted over the past five years (CAS, 2011; Hazzan et al., 2008; Hubwieser et al., 2011; Snyder, 2012; Sturman and Sizmur, 2011), with numerous nations and states engaged in efforts to revamp their compulsory-level computer science curricula (Hubwieser et al., 2015b; Webb et al., 2017). Relevant examples include the USA (both nationally (ACM et al., 2016) and at the state level (Ericson et al., 2016; Guzdial et al., 2014)), France (Baron et al., 2014), Italy (Bellettini et al., 2014), India (Raman et al., 2015), Israel (Armonia and Gal-Ezer, 2014; Gal-Ezer and Stephenson, 2014), New Zealand (Bell, 2014; Bell et al., 2012), Russia (Khenner and Semakin, 2014) and Sweden (Rolandsson and Skogh, 2014), with each country having different issues to address in implementing a concepts-rich computer science curriculum (along with ensuring they have the expertise in the teaching profession to deliver it effectively). It is worth noting that recommendations for academic computer science curricula have a long pedigree (Atchison et al., 1968). Despite the increasing number of success reports from several countries – with establishing computer science as a worthwhile and high-value subject a frequently-named educational goal – it is one that represents significant challenge in terms of research, assessment and teacher training (Vahrenhold, 2012), especially for early years education (Beauchamp, 2016; Bird et al., 2014; Manches and Plowman, 2017) and in the wider context of re-thinking effective pedagogies for the digital age (Beetham and Sharpe, 2013). Due to the substantial differences of preconditions, circumstances and influence factors, it is often difficult to compare or transfer research results in the field of computer science education in schools from one country to another; a framework has been developed – the “Darmstadt model” (Hubwieser, 2013) – to reflect all factors that might be relevant for computer
science education in schools (Hubwieser et al., 2014, 2015a; Passey, 2017). Furthermore, throughout all of these reforms, there is a clear imperative to provide high-quality professional service to upskill existing and new educators to be able to effectively deliver the curriculum (Ericson et al., 2014; Hazzan et al., 2015; Sentance et al., 2012, 2013, 2014), reflecting a variety of challenges (Ni, 2009; Sentance and Csizmadia, 2017).

Policy and practice in the higher education sector is also of importance, particularly from a curriculum standards and guidelines perspective; for example from the UK’s Quality Assurance Agency and its Subject Benchmark Statement for Computing (QAA, 2016) and the US Association of Computing Machinery (ACM) model curricula for undergraduate degree programmes (ACM, 2010, 2013, 2014, 2016), especially from a pedagogic perspective (Dziallas and Fincher, 2015). For universities across the UK offering computer science degrees, the ongoing school curriculum reform has had uncertain (and emerging) impact on the delivery (and thus pedagogy) of their undergraduate programmes, with the diversity of the educational background of applicants likely to increase before it narrows: it is certainly possible now for prospective students to have anywhere from zero to four or five years experience (and potentially two school qualifications) in computer science with some knowledge of programming.

Theories of Learning

There are long-established groundings in the literature to key educational research themes and how they support effective pedagogies and underpin high-quality learning and teaching in ICT and computing (Cox et al., 2004).

Constructivism is a philosophical viewpoint about the nature of knowledge, representing an epistemological stance. Constructivists interpret learning as the development of personalised knowledge frameworks that are continually refined, with learners actively constructing mental models to understand the world around them, rather than have understanding passively “dumped into their brains” (Esper et al., 2012). According to this theory, to learn, a student must actively construct knowledge, rather than simply absorbing it from textbooks and lectures (Davis et al., 1990; Meyer et al., 2010). There is an important distinction to note between social constructivism – which primarily focuses on the development of an individual’s understanding – and cognitive constructivism – which focuses on knowledge constructed through interaction with others. The influential idea of cognitive apprenticeship (Brown et al., 1989), which holds that masters of a skill often fail to take into account the implicit processes involved in carrying out complex skills when they are teaching novices. To combat these tendencies, cognitive apprenticeships are designed, among other things, to bring these tacit processes into the open, where students can observe, enact, and practice them with help from the teacher, which honors the situated nature of knowledge. Further work in social constructivism includes communities of practice (Lave, 1991), with a primary focus on learning as social participation – that is, an individual as an active participant in the practices of social communities, and in the construction of his or her identity through these communities (Wenger, 2000).

The work of Piaget (1950) and Vygotsky (1978) is prominent in this space. Piaget described constructivism in education as being the process whereby students constructed their own unique systems of knowing (Piaget, 1950), in consequence of which the teacher should focus on this individual process of internal construction rather than standing at the front and dictating their own models. Students develop their own self-constructed rules, or “alternative frameworks” (Ben-Ari, 2001). This appears to be a common theme in programming, in which these alternative frameworks “naturally occur as part of the
transfer and linking process” (Clancy, 2004); they represent the prior knowledge essential to the construction of new knowledge. When learning, the student modifies or expands his or her framework in order to incorporate new knowledge. Piaget’s theory of constructivist learning has had wide-ranging impact on learning theories and teaching methods in education, and has been an underlying theme of many education reform movements across the world; it has also been proposed as a suitable pedagogy for information and computing sciences in the past (Boyle, 2000), particularly for computer programming (Lister, 2016). This constructivist approach informs the US National Research Council’s recommendations for the adoption of active learning pedagogies in the classroom (Eberlein et al., 2008), particularly focused around the perceived utility of the “classroom lecture” and exploratory homework, a mechanism to support active learning for teaching programming languages by seeking to develop a model for students of how to explore and understand key constructs and concepts (Esper et al., 2012).

Constructivism does not refer to a specific pedagogy, although it is often confused with constructionism, an educational theory developed by Seymour Papert, inspired by constructivist and experiential learning ideas of Piaget. There are thus strong pedagogical links to the work of Papert (1993) and social constructionism, where learners work together to construct solutions (Kafai and Resnick, 1996), potentially through distributed means (Resnick, 1996), as well as how constructionist learning in science and mathematics can be applied to computing and the digital world (Kafai, 2006; Kafai and Resnick, 1996). Pedagogical approaches that emphasise constructive and collaborative learning in CS1 classrooms, based on empirical results, have been conducted (Van Gorp and Grissom, 2001), with potential application to schools. This leads into the pedagogic approaches around designing and making artefacts, particularly on the idea of “computational participation” (Kafai et al., 2014), as well as using robotics (Barreto and Benitti, 2012), although there are tensions with “technocentrism” and taking a constructionist approach to learning (Brennan, 2015).

While there have been criticism of Bloom’s taxonomy (Bloom, 1965) – particularly as the classification was not a properly constructed taxonomy, as it lacked a systemic rationale of construction – work has been done on whether the taxonomy is appropriate for computer science, especially with the structure of observed learning outcomes (SOLO) taxonomy (a general educational taxonomy that describes levels of increasing complexity in student’s understanding of subjects). Bloom’s taxonomy of the cognitive domain and the SOLO taxonomy are being increasingly widely used in the design and assessment of courses, but there are some drawbacks to their use in computer science (Anderson et al., 2013); researchers have advocated its use for specifying learning outcomes in computer science prior to assessment (Starr et al., 2008), for computational thinking and teaching programming (Selby, 2015), as well as attempting to ‘benchmark’ the content of a computing degree (Fuller et al., 2007; Johnson and Fuller, 2006).

There is also extensive work in threshold concepts (Land et al., 2008; Meyer et al., 2010) – as a subset of the core concepts in a discipline – for computing, and specifically in programming (Khalife, 2006); while not necessarily pedagogy, identifying and understanding which concepts are transformative provides valuable insight into key parts of the curriculum, as these are the building blocks that must be understood. Threshold concepts must be:

- **transformative**: they change the way a student looks at things in the discipline;
- **integrative**: they tie together concepts in ways that were previously unknown to the
Students who have mastered these threshold concepts have, at least in part, “crossed over from being outsiders to belonging to the field they are studying” (Eckerdal et al., 2006), although there is some dispute on how they apply to computer science (Boustedt et al., 2007). Furthermore, perceptions (Lonati et al., 2011) and issues with domain terminology is still a barrier to many – both practitioners and policymakers – although there is ongoing work in the UK to address this (Simon et al., 2015; UKForCE, 2016).

There are a number of other pedagogical approaches that have been used to identify key introductory computing topics; Delphi processes – a structured multi-step process that uses a group of experts to achieve a consensus opinion (Clayton, 1997) – have been used to identify topics that are important and perceived to be difficult in each of three introductory computing subjects: discrete math, programming fundamentals, and logic design. These topic rankings can then be used to guide both the coverage of standardised tests of student learning (i.e. concept inventories) and can be used by educators to identify which topics merit emphasis (Goldman et al., 2008).

This naturally leads into the use of concept inventories, specialised assessment instruments that enable educational researchers to investigate student (mis)understandings of concepts in a particular domain i.e. how a student’s conceptual framework matches the accepted conceptual framework of a discipline. While students experience a concept inventory as a set of multiple-choice items taken as a test, this belies its purpose, its careful development, and its validation (Almstrum et al., 2006). A concept inventory is not intended to be a comprehensive instrument, but rather a tool that probes a student’s comprehension of a carefully-selected subset of concepts that give rise to the most common and pervasive mis-modellings; concept inventories have been developed and used in a number of STEM fields, with application to computing, for example digital logic (Herman et al., 2010), algorithms and data structures (Danielsiek et al., 2012).

Peer instruction (PI) is an effective active learning method that supports student-centric classrooms, where students construct their own understanding through a structured approach featuring questions with peer discussions. PI has been shown to increase learning in STEM disciplines such as physics and biology (Crouch and Mazur, 2001), as well as being an indicator of student success. The potential implications of widespread adoption of PI in computing has also been discussed (Porter et al., 2011, 2013a), with the general focus of PI research been on the in-class portion of PI: multiple choice questions and group discussion (Zingaro et al., 2013). It is generally assumed that early success in CS1 is crucial for success on the examination and course as a whole. Particularities of students, densely-connected CS1 content, and recurring core topics each suggest that it is difficult to rebound from early misunderstandings. PI data, in addition to examination data, has been used to explore relationships between in-class assessments and performance on the end of term assessment; early course performance very quickly and strongly predicts performance on the final examination and that subsequent weeks provide no major increase in that predictive power (Porter and Zingaro, 2014). Peer instruction has also
been used to increase engagement in lectures, for example with students answering a multiple choice question typically using hand-held remote devices, discuss the question with their peers, and then answer the question again; studies have reflected on this approach, understanding the value for the teacher, and reporting the attitudes and opinions of the students (Simon et al., 2010). Furthermore work on the benefits of lab-centric instruction, a collection of pedagogical techniques enabled by converting class time in lecture to time in a supervised closed lab, with the intermediate goal of supporting students’ in developing skills of self-assessment (Lewis et al., 2011).

Finally, there is a body of pedagogic approaches on the use games-based learning to support the acquisition of knowledge in computing education (Egenfeldt-Nielsen, 2007; Giannakos, 2013; Schmitz et al., 2011; Theodoropoulos et al., 2017), with a variety of factors affecting learner perception, engagement and performance Bourgonjon et al. (2010); Farrell and Moffat (2014); Jiau et al. (2009); Lee and Hammer (2011). For example, enhancing self-motivation to learn programming using game-based simulation and metrics, primarily to avoid tedious trial-and-error refinement processes by providing helpful clues on how the student might reprogram the strategy to improve the result (Jiau et al., 2009).

Computer Science Fundamentals

There are a range of pedagogic approaches to developing knowledge and understanding of key fundamental topics in computer science, for example understanding automata (Isayama et al., 2017), finite state machines and Turing machines (Korte et al., 2007), as well as key issues such as concurrency (Kolikant, 2004). A wide range of work has been conducted in algorithm design strategies (Levitin, 1999, 2016) and in particular, the use of analogies (Cao et al., 2016; Chec, 1993; Repenning and Perrone, 2000) and metaphors (Forišek and Steinová, 2012; Hidalgo-Céspedes et al., 2014). This is linked to students’ misconceptions relating to algorithms and data structures, building in work on concept inventories (Coffey, 2013; Danielsiek et al., 2012). Tracing – to understand the notional machine provides insight into the process a computer goes through when executing a section of code is manually stepped through by the learner in order to understand how it will work and what the expected output should be (Sorva, 2013). Just showing a visualisation of a piece of code or an algorithm does not appear to improve learners’ depth of understanding; annotation on, and next to, the code provides a visual guide of what actually happens when it executes. This makes the abstract concepts and hidden mechanisms visible and therefore easier to think and reason about (Sorva, 2013). However, just viewing program or algorithm visualisations does not necessarily lead to improved understanding: how students use the technology has a greater impact on effectiveness than what the technology shows them (Hundhausen et al., 2002). Instead, active engagement with the visualisation is the key factor either by predicting the next step, choosing suitable input data to achieve a particular output or learners constructing the visualisation themselves (Grissom et al., 2003; Petre and Blackwell, 1999).

Various studies have been conducted to determine factors that promote success in introductory college computer science courses, identifying possible predictive factors including: math background, attribution for success/failure (luck, effort, difficulty of task, and ability), domain specific self-efficacy, encouragement, comfort level in the course, work style preference, previous programming experience, previous non-programming computer experience, and gender (Cantwell Wilson and Shrock, 2001). Studies have revealed three predictive factors: comfort level, math, and attribution to luck for success/failure; comfort level and math background were found to have a positive influence on success, whereas at-
tribution to luck had a negative influence. Furthermore, by considering the different types of previous computer experiences (including formal programming class, self-initiated programming, internet use, game playing, and productivity software use) that both a formal class in programming and game playing were predictive of success, as well as outreach programs (Franklin et al., 2013). Formal training had a positive influence and games a negative influence on class grade (Cantwell Wilson and Shrock, 2001). Much work has focused on planning and assessment of CS1 level topics – especially programming – and how this impacted on attainment of students (Matthasdóttir and Arnalds, 2015). Many university computing academics report bimodal grade distributions in their CS1 classes, believing that such a distribution is due to there being an innate talent for programming, a so-called “geek gene” (Ahadi and Lister, 2013). Robins (2010) introduced the concept of learning edge momentum, which offers an alternative explanation for the purported bimodal grade distribution, with studies analysing empirical data from introductory programming class, looking for evidence of geek genes, learning edge momentum and other possible factors (Ahadi and Lister, 2013). Further work in this area is discussed later on in this chapter, particularly with how developing conceptual understanding of theoretical concepts apply to programming and computational thinking.

Programming and Coding

Since the launch of the curriculum in England, there has been significant media and public attention on this new “coding curriculum”3, with programming dominating the focus on what is a much broader discipline; even the terminology used can be divisive: “coding” or “programming”? However, as part of the initial search of the wider computing education literature, it is clear there is a substantial corpus of work on developing and understanding effective pedagogies on how to teach programming.

For many years – and increasingly at all levels of compulsory and post-compulsory education – the choice of programming language to introduce the “art” (Knuth, 2011), “science” (Gries, 1981) and “discipline” (Dijkstra, 1976) of computer programming via key principles, constructs, syntax and semantics has been regularly revisited. So what is a good first programming language? The issues surrounding choosing a first language (Gupta, 2004; Kaplan, 2010) – and a Google Scholar search identified a number of papers of the form “X as a first programming language”, going as far back as the 1970s (Gries, 1974) – appear to be legion, especially with discussions of what precisely we aim to achieve from teaching programming (Fincher, 1999; Schulte and Bennedsen, 2006), to psychological approaches (Winslow, 1996), gender gaps (Angel Rubio et al., 2015), adults learning programming (Guo, 2017), attitudes (Fesakis and Serafeim, 2009) and a new focus on developing transferable computational thinking and problem solving skills (Tedre and Denning, 2016; Wing, 2008). While in the past, research on the teaching of introductory programming had limited effect on classroom practice (Pears et al., 2007), increasingly relevant research exists across several disciplines including education and cognitive science, disciplinary differences have made this material inaccessible to many computing educators. Furthermore, computer science educators have not had access to comprehensive surveys of research in this area (McCracken et al., 2001; Pears et al., 2007).

The topics addressed in the literature span a wide range of problems and solutions as-

---

3e.g. https://www.theguardian.com/technology/2014/sep/04/coding-school-computing-children-programming
associated with the teaching of programming such as introductory programming courses, exposition of the programming process, apprentice-based learning, functional programming first, problem-based learning, the use of on-line tutorials, object-oriented programming and Java, environments to introduce programming, model-driven programming as opposed to the prevailing language-driven approach, teaching software engineering, testing, extreme programming, frameworks, feedback and assessment, active learning, technology-based individual feedback, and mini project programming examinations (Bennedsen et al., 2008).

Mastery of basic syntactic and logical constructs is an essential part of learning to program. Unfortunately, practice exercises for programming basics can often be tedious (or perceived to be tedious), making it difficult to motivate students (Dasgupta and Resnick, 2014; Dasgupta et al., 2015). The idea of using “problets” – problem solving software assistants for learning, reinforcement and assessment of programming concepts – was introduced by Kumar (2005). They are designed to help students learn programming concepts through small-scale problem-solving e.g. code-tracing problems (Kumar, 2015), and as a supplement to large-scale programming traditionally used in introductory programming courses. Other simple “programming puzzles” was introduced by Parsons and Haden (2006): automated, interactive tools that provides practice with basic programming principles in an entertaining puzzle-like format. Careful design of the items in the puzzles allows the tutor to highlight particular topics and common programming errors. Since each puzzle solution is a complete sample of well-written code, use of the tool thus exposes students to good programming practices, particularly when coupled with using sub-goals (Morrison et al., 2016).

This links with work focusing on issues with students who are unable to “trace” code, implying they are unable to explain its syntax and function. Students who tend to perform reasonably well at code writing tasks have also usually acquired the ability to both trace code and explain code (Lister et al., 2009) – the performance of students on code tracing tasks correlates with their performance on code writing tasks. A correlation was also found between performance on “explain in plain English” tasks and code writing. Further work in this space suggests the possibility of a hierarchy of programming-related tasks: knowledge of programming constructs forms the bottom of the hierarchy, with “explain in English”, Parson’s puzzles, and the tracing of iterative code forming one or more intermediate levels in the hierarchy (Lopez et al., 2008). The structure of observed learning outcomes (SOLO) taxonomy – a general educational taxonomy that describes levels of increasing complexity in student’s understanding of subjects – has also been used to describe differences in the way students (particularly novice programmers) and educators solve small code reading exercises (Lister et al., 2006).

Programming is a hard craft to master and its teaching is challenging. An apprentice model, where students learn their craft from a master is an approach that can lead to improved student engagement (Astrachan and Reed, 1994; Vihavainen et al., 2011). Although traditionally applied to physical and vocational skills, the apprenticeship model can also be applied to the acquisition of cognitive skills such as those required for programming, as discussed in the Theories of Learning section. In this context, the master is required to focus on the programming process and demonstrates it through writing, debugging and running ‘live’ programs. This takes place whilst being observed by the student cohort; scaffolding is provided through the provision of regular practical exercises with good quality formative feedback (Crick et al., 2015).

There is a belief that programming – as opposed to, say analysis of algorithms, a
closely related theoretical skill – is fundamentally a craft that needs immersion and practice (Fincher, 1999; Milne and Rowe, 2016). This is related to both general and discipline-specific pedagogy, with some positions implying that minimal guidance during instruction is less effective and efficient than guidance specifically designed to support the cognitive processing necessary for learning (Kirschner et al., 2006). Connected to this general point is a curious paradox in the teaching of programming: students are generally taught to write programs, and not particularly to read them (essentially overlooking the importance of code literacy), whereas in natural languages, be it the mother tongue or foreign language instruction, students are taught to read before they are taught how to write (Crick et al., 2015). Approaches such as discussion classes (Hagan and Sheard, 1998) and live-coding – defined as “the process of designing and implementing a [coding] project in front of class during lecture period” – can be effective in teaching introductory programming, with experimental data indicating that teaching via live-coding is as good as if not better than using static code examples (Rubin, 2013).

Pair programming (Williams and Kessler, 2000) is another common theme; prior research on pair programming has found that compared to students who work alone, students who pair have shown increased confidence in their work, greater success in CS1, and greater retention in computer-related majors (Hanks et al., 2004). A systematic review by Salleh et al. (2011) showed that students’ skill level was the factor that affected pair programming’s effectiveness the most. The most common measure used to gauge pair programming’s effectiveness was time spent on programming; in addition, students’ satisfaction when using pair programming was overall higher than when working solo and was effective in improving students’ grades on assignments. Furthermore, pairing students were more likely to turn in working programs, and these programs correctly implemented more required features. An unexpected but significant finding was that pairing students were more likely to submit solutions to their programming assignments (Hanks et al., 2004). From a qualitative, student-focused approach, how do students define, experience, and value the pair programming experience; especially how do they experience and value it compared to solo programming? Students get stuck less and explore more ideas while pairing, and believe that pair programming helped them in CS1; however, students reported that when solo programming they were more confident and better understood their programs. Many students also said that they started work on their assignments earlier when soloing. Students also continue to use other students as resources even when working solo (Simon and Hanks, 2008). Finally, pair programming produces more proficient, confident programmers – and may help increase female representation in the field (McDowell et al., 2006).

As expected, there is a broad corpus on work on innovative and effective assessment mechanisms for programming, asking fundamental questions about what are we aiming to do when we teach programming (Fincher, 1999), through to misconceptions, attitudes and perceptions (Clancy, 2004). There have been multi-national, multi-institutional studies of assessment of programming skills of first-year computer science students in higher education institution (Lister et al., 2004; McCartney et al., 2013; McCracken et al., 2001), looking at effective pedagogies (Pears et al., 2007; Schulte and Bennedsen, 2006), the potential impact of programming expertise on learning motivation and academic achievement (Kori et al., 2016) – particularly failure rates (Bennedsen and Caspersen, 2007) – as well as the choice of introductory programming language and the potential impact on students’ grades and attainment (Bergin and Reilly, 2006; Ivanović et al., 2015; Porter et al., 2013b; Simon et al., 2006). Research into effective study habits has provided in-
sight into preparation for assessment, as well as resilience (Nam Liao, 2016; Willman et al., 2015); in many instances, high-quality exemplification is a key feature of successful initiatives (Scott, 2013). Recent international surveys of coding practices in K-8 have provided insight into demographics, experience, teaching practices and teacher confidence.

From a higher education perspective, there is a substantial body of work around effective pedagogies to support high-quality learning, teaching and assessment of computing. In the UK, much work has been supported by the Higher Education Academy – an independent non-profit organisation committed to world-class teaching in higher education; they work in partnership with institutions and individuals in higher education supporting student success through collaboration and share teaching strategies and practice. This has been primarily done through the HEA’s STEM disciplinary theme (HEA, 2017). As the national body for promoting high-quality learning and teaching in higher education institutions in the UK, they run a programme of discipline-specific events, including STEM teaching and learning conferences, funding pedagogic research on innovative pedagogies, learning & teaching resources and case studies to inform and improve practice (for example, looking at retention and attainment in computer science (Gordon, 2016), a key sector issue). In the context of what are perceived to be the most difficult introductory topics in computer science degrees, numerous key themes frequently appear (Dale, 2006), with innovative approaches to addressing them (Hazzan et al., 2011).

Furthermore, the teaching of introductory programming in many higher education institutions has started to move away from focusing primarily on syntax to developing a deeper understanding of principles of programming, transferable language semantics, underlying constructs and structures, as well as developing useful and usable software artefacts: in summary, building upon the craft of programming via software carpentry (Wilson et al., 2014), codemanship (Crick et al., 2015) and teamwork (Martinez et al., 2014). Although controversial (Guzdial, 2009), this practice has a dual focus: firstly, it develops a high-level appreciation for why programming is being taught – essentially to solve real-world problems, using the most appropriate languages, tools and environments; secondly, it allows embedding the use of tools, methodologies and techniques so as to start to develop best practice for real-world software development. While the aim of this approach is not to create industry-level programmers at the end of the degree programme, by fostering and supporting the development of a particular culture around creating useful and usable software artefacts, underpinned by rigorous knowledge and theory, ensures that students understand how software is designed, developed and maintained in industry, and have the internal framework for developing knowledge and understanding in new languages, tools and environments and methodologies when required to do so (Davenport et al., 2016). However, it is important to note that university educators have the freedom and flexibility to select whichever programming languages, tools and environments they wish; educators in schools and colleges may be constrained in their choice of programming languages by informal requirements of the UK awarding bodies (especially in regards to their specific GCSE Computer Science qualification).

In this evolving national and international environment of emerging policy and curricula, as well as the demands of developing innovative pedagogies and high-quality learning and teaching for computer science degree programmes, national surveys of introductory programming languages provide valuable insight into pedagogy and practice. Longitudinal studies has been conducted in Australia and New Zealand over the past 15 years; in
2001 and 2003, and censuses were conducted on Australian and New Zealand universities to examine trends in the programming languages and environments used in introductory programming courses (Raadt et al., 2004). In 2010 and 2013 (Mason and Cooper, 2014; Mason et al., 2012) similar surveys were conducted online, providing further data on the types of programming languages used (the prevalence of Python), along with some insight into pedagogies and environments. A related national-scale survey conducted in the USA in 2011 also provided some insight into the state of computing education in that country (Davies et al., 2011), along with more recent surveying (Guo, 2014) (again, an increasingly preference for Python). An inaugural survey of UK universities was conducted in 2016 (Murphy et al., 2017), mirroring the Australasian questions and structure; across the UK there is a diversity of practice, with Java appearing to have most traction – perhaps from perceived industry usefulness – but with Python also on the rise, most likely due to school curriculum reforms. Further to this large corpus of work on computer science – and in particular, programming – pedagogies, there is an opportunity to identify best practice that can be adapted for use in schools and colleges (Tangney et al., 2010).

Computational Thinking

Computational thinking (Wing, 2008) is a key theme of the new computing curriculum in England – appearing in the first sentence of the purpose of study (Department for Education, 2013):

“A high-quality computing education equips pupils to use computational thinking and creativity to understand and change the world.”

It refers to a collection of computational ideas and habits of mind that people in computing disciplines acquire through their work in designing programs, software, simulations, and computations performed by machinery. Known in the 1950s and 1960s as “algorithmic thinking”, it meant a mental orientation to formulating problems as conversions of some input to an output and looking for algorithms to perform the conversion (Denning, 2009). An early definition of computational thinking as presented in Wing (2010):

“Computational thinking is the thought processes involved in formulating problems and their solutions so that the solutions are represented in a form that can be effectively carried out by an information-processing agent.”

While increasingly being discussed in the context of being a “21st century skill” (Bocconi et al., 2016b,c), computational thinking is perhaps perceived to be one of the more challenging concepts to teach and assess, partly due to lack of clarity about its precise definition and core components. Over the past decade, computational thinking and related concepts have received increasing attention across the educational field, which has given rise to an increasing amount of academic and grey literature (Lockwood and Mooney, 2017), but also numerous collaborative initiatives (Meerbaum-Salant et al., 2015). Despite this widespread interest and policy focus, successful computational thinking integration in compulsory education faces unresolved issues and challenges (Bocconi et al., 2016a). A 2014 working group report on computational thinking in K-9 education (Mannila et al., 2014) aiming to revealing to what extent different aspects of computational thinking are

---

5See: https://idleclicks.wordpress.com/computational-thinking/
already part of teachers’ classroom practice. Well-established initiatives such as CSUnplugged\(^6\), Barefoot Computing\(^7\), as well as aspects of CAS Tenderfoot\(^8\) in the UK, attempt to build teacher confidence in teaching computer science through algorithmic and computational thinking principles (Csizmadia et al., 2015).

Considerable evidence indicates that domain specific knowledge in the form of schemas is the primary factor distinguishing experts from novices in problem-solving skill. Evidence that conventional problem-solving activity is not effective in schema acquisition is also accumulating. It is suggested that a major reason for the ineffectiveness of problem solving as a learning device, is that the cognitive processes required by the two activities overlap insufficiently, and that conventional problem solving in the form of means-ends analysis requires a relatively large amount of cognitive processing capacity which is consequently unavailable for schema acquisition (Sweller, 1988). This links with the “use-modify-create” framework (Lee et al., 2011), representing three phases of students’ cognitive and practical activity in computational thinking, as well as the use of sub-goals (Morrison et al., 2015).

Linking back to the previous section on effective pedagogies for programming, a study by an ITiCSE working group (McCracken et al., 2001) established that many students do not know how to program at the conclusion of their introductory courses; although this study from 2001 may be dated by recent reforms, it is clear more research needs to be done in this space. A popular explanation for this incapacity is that the students lack the ability to problem-solve; that is, they lack the ability to take a problem description, decompose it into sub-problems and implement them, then assemble the pieces into a complete solution (Lister et al., 2004). While this could also be attributed to the fact that many students have a fragile grasp of both basic programming principles and the ability to systematically carry out routine programming tasks (such as tracing through code), there appears to be a strong perceived link between programming and computational thinking (Davies, 2008; Hu, 2011).

As the term became more accepted and the ‘for everyone’ manifesto generated interest, the focus shifted to curricula (ACM et al., 2016; Barr and Stephenson, 2011; Bell et al., 2010; Brinda et al., 2009; CAS, 2012; Gal-Ezer et al., 1995; Hubwieser et al., 2015b; Iyer et al., 2010), classroom experiences (Ater-Kranov et al., 2010; Cooper et al., 2010; Demšar and Demšar, 2016; Lee et al., 2011; Rodriguez et al., 2017; Yadav et al., 2014) and cross-curricular activities (Basawapatna et al., 2011; Curzon et al., 2009; Eisenberg, 2010; Goldberg et al., 2013; Kubica, 2012; Kules, 2016; Perković et al., 2010); at this time, other fields were also exploring their connections to computational thinking (Eisenberg, 2010; Gal-Ezer and Zur, 2002; L’Heureux et al., 2012; Lu and Fletcher, 2009).

However, modern computational thinking initiatives should be well aware of the broad and deep history of computational thinking, or risk repeating already refuted claims, past mistakes, and already solved problems, or losing some of the richest and most ambitious ideas in computational thinking (Denning, 2007). Since Wing (2006), there have been attempts to provide a formal and comprehensive definition of computational thinking by authoritative individuals and groups (Aho, 2012; Barr and Stephenson, 2011; Denning, 2007; Google, 2013; Grover and Pea, 2013; Guzdial, 2008, 2012; National Research Council, 2010, 2011; Wing, 2008). Recent work by Tedre and Denning (2016) has examined a number of threats to computational thinking initiatives: lack of ambition, dogmatism,
knowing versus doing, exaggerated claims, narrow views of computing, overemphasis on formulation, and lost sight of computational models. Two interesting questions that have been raised in the literature: i) computational thinking a unique and distinctive characterisation of computer science? and ii) is computational thinking an adequate characterisation of computer science? (Denning, 2009).

Policy Context

Overview

With significant international focus on recent computing education policy reform (and emerging practice) in the UK, it is useful to briefly frame the wider policy context, particularly disaggregating the educational, societal and economic drivers. Since the development of ICT as a new curriculum subject in the late 1990s (McKinsey & Co., 1997; Stevenson, 1997), there has been scrutiny of its breadth, depth and impact on students’ learning and attainment (Ofsted, 2001, 2002, 2004). The publication of the Nesta Next Gen. report in 2011 (Livingstone and Hope, 2011) and the Royal Society’s Shut Down or Restart? report (Royal Society, 2012), coupled with the disapplication of the ICT programme of study in England and scrutiny of the quality of a range of vocational qualifications (Royal Academy of Engineering, 2012; Wolf, 2011), left many educational establishments uncertain about the direction of this subject area, as well as poor wider perceptions of the discipline (Quinlan, 2015). However, a new computing curriculum (and associated qualifications) was starting to emerge (Brown et al., 2013, 2014), being published in 2013 for first delivery in September 2014 (Department for Education, 2013) from aged five (Manches and Plowman, 2017) to qualifications at 16 and 18. Alongside reforms in Scotland as part of their Curriculum for Excellence (Scottish Government, 2008), with a clear strand of computing science, as well as emerging education and skills reform in Wales (Arthur et al., 2013; Crick and Moller, 2015; Donaldson, 2015) and Northern Ireland (Perry, 2015). This has been supported by a strong European theme (Informatics Europe and ACM Europe Working Group, 2013), particularly as part of the wider Digital Agenda for Europe9 initiative to promote digital literacy, skills and inclusion. Throughout all of these reforms, there is a clear imperative on providing high-quality professional service to upskill existing and new educators to be able to effectively deliver the curriculum (Cutts et al., 2017; Sentance and Csizmadia, 2017; Sentance et al., 2012, 2013, 2014).

From a societal and economic policy context, the “digital skills gap” has been highlighted by a number of high-profile policy reports, including by the UK Digital Skills Taskforce in 2014 (UK Digital Skills Taskforce, 2014), the House of Lords Select Committee on Digital Skills in 2015 (House of Lords Select Committee on Digital Skills, 2015), as well as the House of Commons Select Committee on Science & Technology in 2016 (House of Commons Select Committee on Science & Technology, 2016). From a wider societal context, a 2017 report from the Children’s Commissioner for England’s Growing Up Digital Taskforce stated that: “The current Computing curriculum sets out in detail the technical skills and some of the legal knowledge a child should have at different ages. The Childrens Commissioner however believes this is too narrow, and often too late; your data protection rights, for instance, are not taught until GCSE level, and GCSE Computing is

Policy and practice in the higher education sector is also of relevance to this project, particular from a curriculum standards and guidelines perspective, for example from the QAA in the UK (QAA, 2016) and the ACM in the USA (ACM, 2010, 2013, 2014, 2016). For universities across the UK offering computer science degrees, the school curriculum reform has had uncertain (and emerging) impact on the delivery of their undergraduate programmes, with the diversity of the educational background of applicants likely to increase before it narrows: it is certainly possible now for prospective students to have anywhere from zero to four or five years experience (and potentially two school qualifications) in computer science with some knowledge of procedural, object-oriented and/or functional programming. Furthermore, over the past three years, there has been increasing scrutiny of the quality of teaching in UK universities, partly linked to the current levels – and potential future increases – of tuition fees (generally paid by the student through government-supported loans), as well as relative levels of graduate employability and the perceived value of professional body accreditation by industry. In February 2015, the Department of Business, Innovation & Skills initiated independent reviews of STEM degree accreditation and graduate employability\textsuperscript{10}, with a specific focus – the Shadbolt review – on computer science degree accreditation and graduate employability, reporting back in May 2016 (Shadbolt, 2016). A number of recommendations were made to address the relatively high unemployment rates of computer sciences graduates, particular on the quality of data, course types, gender and demographics.

Qualifications

The uptake of computing/computer science qualifications at GCSE and A-level in the UK can obviously be made by looking at the educational establishments that offer the qualification and the students sitting it. Not all schools and colleges offer computing qualifications at GCSE or A-level and not all students sit qualifications in computing. This is further complicated by the increasing divergence of the four education systems in the UK, two of whom are still undergoing curriculum reform. However, all of the major examination boards – AQA, OCR, Edexcel, CCEA and WJEC/Qualifications Wales) – offer GCSEs and/or A-Levels in computing/computer science. Even where a qualification is taught by a school, subject requirements might limit the type of student who is able to take the course. Whilst at A-level, computing is a well-established subject, it is only offered by a minority of centres, with some areas having no provision. Until recently, the number of students taking A-level computing has been in decline, but Joint Council for Qualifications figures show that since 2014 numbers have been increasing year-on-year (Joint Council for Qualifications, 2014, 2015a, 2016a). A new computing GCSE was introduced by OCR in 2011\textsuperscript{11} with the first cohort of students sitting examinations in 2013. Understandably, not all schools and colleges adopted this qualification immediately, and whilst the number of centres and students have been increasing, the numbers have not yet matched those of ICT (Joint Council for Qualifications, 2016b). A similar picture has been observed at A-level with numbers of computing students rising 50% in five years but still well below ICT (Joint Council for Qualifications, 2011, 2015b). Additionally, with recent school funding changes at A-level, from a per subject to per student system

\textsuperscript{10}https://www.gov.uk/government/collections/graduate-employment-and-accreditation-in-stem-independent-reviews
\textsuperscript{11}http://www.ocr.org.uk/qualifications/gcse-computing-j275-from-2012/
in England, the computing cohort size of A-level providers now becomes a greater concern for the ongoing financial viability of the subject, with smaller subject cohorts potentially making the course too expensive for smaller providers (Kemp et al., 2016).

Gender and Diversity

While a comprehensive analysis of the history and challenges of gender and diversity in computer science is out of scope for this review, it is worth highlighting some of the key issues. A number of different strategies have been proposed for involving females in computing – particularly in higher education – but there is a clear and recognised issue with engaging and retaining females in computing from the 1990s onwards. Despite the current growing popularity of the computer science major in the USA, women remain sorely underrepresented in the field, continuing to earn only 18% of bachelors degrees; understanding women’s low rates of participation in computer science is important given that the demand for individuals with computer science expertise has grown sharply in recent years (Lehman et al., 2017). Numerous studies (Adam et al., 2004; Bunderson and Christensen, 1995; Clarke, 1992; Wilson, 2003) have explored some of the reasons that may underlie the gender segregation and declining levels of female participation within the field of computing, including gender differences and differences in computer science majors vs. non-majors in ability in quantitative areas, educational goals and interests, experience with computers, stereotypes, as well as domain knowledge, confidence, personality, support and encouragement, stress and financial issues, gender discrimination, and attitudes toward the academic environment in computer science. Three themes often appear: communicative processes, social networks and legitimising claims to knowledge, overlaid by gendered-power relations (Robertson et al., 2001). One of the challenges is to conceptualise women’s computer skills as real computing and to instead ask what is wrong with computing rather than what is wrong with women (Clegg and Trayhurn, 2000). Finally, it is worth noting that there is growing evidence that instructors in computer science tend to believe in the importance of innate ability (Lewis, 2007) and thus the resulting negative effects this has on diversity (Murphy and Thomas, 2008) and thus performance (Patitsas et al., 2016). Addressing these gender stereotypes about intellectual ability from an early age (Bian et al., 2017), as well as improving these fixed mindsets in computing courses can have a positive effect on learning outcomes (Cutts et al., 2010).

Men had more confidence in using computers than did women even when statistically-controlling quantitative ability; in fact, female CS majors had less computer confidence than did male non-majors (Beyer et al., 2003). Studies have been conducted to determine factors that promote success in introductory college computer science courses – particularly programming (Angel Rubio et al., 2015) – and what differences appear between genders on those factors (Clarke and Chambers, 1989), including gender-equitable pedagogical practices (Vekiri, 2013); three predictive factors have appeared: comfort level (with a positive influence), mathematics background (with a positive influence), and attribution to luck (with a negative influence) (Cantwell Wilson, 2002).

While a number of bodies have been active in promoting diversity in computing, especially addressing the poor gender balance, e.g. BCSWomen12, CAS #include13 and ACM-W14, more recent work has been conducted in England to better understand the

12http://bcswomen.bcs.org
13http://casinclude.org.uk
14https://women.acm.org/
landscape – particularly, the A-level and GCSE computing cohorts – beyond the widely publicised disparity in gender (Joint Council for Qualifications, 2015b), looking at provider type, provider location, subject mix, the ethnicity and socio-economic status of students (Kemp et al., 2016).

Conclusions & Recommendations

There are a number of recognised issues with the existing corpus of work associated with effective pedagogies and assessment mechanisms for computing. In the first instance, much of the existing research has been conducted in post-compulsory or higher education, addressing some of the specific issues associated with degree-level study. While there is certainly transferability from higher education to compulsory education, we have to make sure that the approaches and pedagogies are relevant and appropriate to the age group and topics being covered. Furthermore, we have to consider the quality and transferability of some of the studies presented, especially with regards to cohort size, sampling, statistical significant and claimed impacts, particularly when studies are based on single-institution, non-sampled interventions.

Building upon the previous section, in which we discussed some of the limitations of the existing research base, in this section we will summarise some of the key points from the literature with an aim of capturing future research priorities and requirements to best support the ongoing implementation of the computing curriculum in England.

• Identifying a UK Research “Grand Challenge” Theme: there is a significant opportunity to identify and develop a long-term strategic research programme (and associated funding models) for the UK to “develop effective pedagogies for computing education in schools”. However, it is important to recognise that this is a significant task that may take a number of years; it will not be possible to do everything. This could consist of a number of sub-strands around some of the key themes identified in this literature review, particularly focused on concepts, skills and pedagogies. For example, it could be framed around “big ideas in computing”, in a similar way to work done on the Big Ideas in Science (Association for Science Education, 2017). It should aim to establish cross-cutting understanding of a set of “big ideas” in computing which include key strategic themes and concepts, as well as and appreciation of computing’s role and impact on society.

• Identifying research with immediate applicability to schools: a clear short-term priority exists in identifying “low-hanging fruit” that could be quickly transferred and translated for rapid adoption in schools and colleges by educators, especially from the higher education sector. For example, there are key areas of pedagogy that have a strong empirical evidence base and could be applied to schools, as well as calling for more school-specific research; for example, tracing, peer instruction and pair programming, as well as work on effective pedagogies of programming and the apprenticeship model from higher education.

• Improving the strength of the UK research base: it is a known issue that there is a lack of UK capacity for computer science education research, with small clusters at a few institutions across the UK. It is clear that the researchers and groups at these few institutions – namely, King’s College London, University of
Kent, Queen Mary University of London, University of Glasgow, as well as individuals at a number of other institutions across the UK – do not currently provide the critical mass necessary to drive forward computer science education research, from both a pedagogical and practice perspective. This is contrasted by international strengths in STEM education, but is also an artefact of clear (if any) funding mechanisms to support computer science education research in the UK. Therefore, a priority recommendation is to identify funding models and schemes to support existing researchers, as well as develop researchers and groups at other institutions across the UK. Funding models could support doctoral and/or post-doctoral research, as well as identifying strategic partners (e.g. Nuffield, UK Research Councils, Leverhulme, etc) for co-funded schemes. It is imperative that the level of research and number of researchers increases in the UK to support the curriculum and engage with and contribute to the international research domain. Precisely what these models of funding look like are somewhat dependent on the partner organisations – particularly if this includes UK Research Councils, charities and/or industry support. This literature review has provided the start of a gap-analysis to enable a prospective research programme to be developed.

- **Adopting interdisciplinary approaches**: addressing some of the cognitive issues with learning some of the key topics, how can we best expand our research collaborations to ensure that we have the educational/cognitive/psychological expertise to achieve what we need to achieve? For example, research themes in programming, computational thinking, creativity, diversity, perceptions, etc.

- **Addressing the long-standing terminological diversions**: it is clear that domain terminology is still a barrier to many, both practitioners and policymakers; while there is existing work to address this (UKForCE, 2016), as a community we will need to agree on consistent terminology and nomenclature going forward so as to be able define and deliver a coherent research programme.

- **Relationship with digital literacy/competency**: further to the recommendations in Royal Society (2012) – and the resulting curriculum reform in England – digital literacy has been largely sidelined during the expansion of computer science. While there are active national and international research communities in this space, there is a need to better understand how digital literacy/competency fits in the wider “computing” landscape (especially across the four nations of the UK).

- **Defining and evidencing the value and impact of computational thinking**: for example, how do you teach computational thinking effectively? The idea is frequently espoused that it is something you get from learning computing (or programming. Teaching it specifically, do you teach the separate CT elements separately, or do you teach them holistically as an overall “CT package”? Also, teaching skills vs. teaching the concepts – how do you teach them to students and how do you teach them to educators? Finally, how does “unplugged” work and how best to teach it in the context of CT?

- **Effective programming pedagogies**: in particular, tools for teaching programming; software tools, hardware gadgets and pedagogies for teaching effective programming. We have been asking similar research questions for nearly 30 years,
but we do not yet have the rigorous evidence base; some of this work clearly requires more empirical evidence, contextualised to the specific school or university environment.

- **Improving the perceptions of young people towards computer science**: students’ perceptions, wider public perception (potentially linking with the existing Wellcome science education work), as well as looking at gender/class/ethnicity/inclusion.

- **Local, regional and national mapping**: what is actually being taught in schools and colleges e.g. published schemes of work/curricula on school websites, aggregating at the local, regional and national level to better understand the landscape, as well as being able to contribute to international initiatives.

In summary, we need to identify and sustainably fund long-term research focused on supporting students’ learning, so that they are able to obtain secure, resilient and transferable knowledge of the subject of computing in the UK.
References


