C·E·P·S *Journal*

Center for Educational Policy Studies Journal
Revija centra za študij edukacijskih strategij

Vol.12 | Nº2 | Year 2022

Vol.12, Nº2, Year 2022
Focus issue editors / Uredniki številke:
Gregor Torkar, Miha Kovač and Mojca Kovač Šebart

Revija Centra za študij edukacijskih strategij
Center for Educational Policy Studies Journal
issn 2232-2647 (online edition / spletna verzija)
issn 1855-9719 (printed edition / tiskana verzija)
Publication frequency: 4 issues per year
Subject: Teacher Education, Educational Science
Published by / Založila: University of Ljubljana Press / Založba Univere v Ljubljani / *For the publisher:* Gregor Majdič, The Rector of the University of / rektor Univerze v LjubljaniLjubljana / *Issued by:* Faculty of Education, University of Ljubljana / *For the issuer:* Janez Vogrinc, The dean of Faculty of Education / dekan
Technical editor: Lea Vrečko / *English language editor:* Terry T. Troy and Neville J. Hall / *Slovene language editing:* Tomaž Petek / *Cover and layout design:* Roman Ražman / *Typeset:* Igor Cerar / *Print:* Birografika Bori

C·E·P·S *Journal*

Center for Educational Policy Studies Journal
Revija Centra za študij edukacijskih strategij

The CEPS Journal is an open-access, peer-reviewed journal devoted to publishing research papers in different fields of education, including scientific.

Aims & Scope

The CEPS Journal is an international peer-reviewed journal with an international board. It publishes original empirical and theoretical studies from a wide variety of academic disciplines related to the field of Teacher Education and Educational Sciences; in particular, it will support comparative studies in the field. Regional context is stressed but the journal remains open to researchers and contributors across all European countries and worldwide. There are four issues per year. Issues are focused on specific areas but there is also space for non-focused articles and book reviews.

About the Publisher

The University of Ljubljana is one of the largest universities in the region (see www.uni-lj.si) and its Faculty of Education (see www.pef.uni-lj.si), established in 1947, has the leading role in teacher education and education sciences in Slovenia. It is well positioned in regional and European cooperation programmes in teaching and research. A publishing unit oversees the dissemination of research results and informs the interested public about new trends in the broad area of teacher education and education sciences; to date, numerous monographs and publications have been published, not just in Slovenian but also in English.

In 2001, the Centre for Educational Policy Studies (CEPS; see http://ceps.pef.uni-lj.si) was established within the Faculty of Education to build upon experience acquired in the broad reform of the national educational system during the period of social transition in the 1990s, to upgrade expertise and to strengthen international cooperation. CEPS has established a number of fruitful contacts, both in the region – particularly with similar institutions in the countries of the Western Balkans – and with interested partners in EU member states and worldwide.

~

Revija Centra za študij edukacijskih strategij je mednarodno recenzirana revija z mednarodnim uredniškim odborom in s prostim dostopom. Namenjena je objavljanju člankov s področja izobraževanja učiteljev in edukacijskih ved.

Cilji in namen

Revija je namenjena obravnavanju naslednjih področij: poučevanje, učenje, vzgoja in izobraževanje, socialna pedagogika, specialna in rehabilitacijska pedagogika, predšolska pedagogika, edukacijske politike, supervizija, poučevanje slovenskega jezika in književnosti, poučevanje matematike, računalništva, naravoslovja in tehnike, poučevanje družboslovja in humanistike, poučevanje na področju umetnosti, visokošolsko izobraževanje in izobraževanje odraslih. Poseben poudarek bo namenjen izobraževanju učiteljev in spodbujanju njihovega profesionalnega razvoja.

V reviji so objavljeni znanstveni prispevki, in sicer teoretični prispevki in prispevki, v katerih so predstavljeni rezultati kvantitativnih in kvalitativnih empiričnih raziskav. Še posebej poudarjen je pomen komparativnih raziskav.

Revija izide štirikrat letno. Številke so tematsko opredeljene, v njih pa je prostor tudi za netematske prispevke in predstavitve ter recenzije novih publikacij.

Contents

Reviews

doi: 10.26529/cepsj.1479

Editorial

The Role of Textbooks in Teaching and Learning Processes

The theme of this issue of the CEPS Journal is the role of textbooks in teaching and learning processes. Nowadays, a range of learning resources are used to support students in their learning process, among which textbooks are the most notable. Textbooks play an important role in almost every school system in the world, representing a useful resource for both teachers as course designers and students acquiring knowledge. However, the use of commercially published textbooks and other teaching and learning resources has advantages and disadvantages, and it is this aspect that we initially wanted to explore in this focus issue. We aimed to investigate the advantages and disadvantages of using a textbook, as well as ways of adapting textbooks that provide teachers with an opportunity to personalise teaching material and enable students to be more actively involved in the learning process. In addition, we intended to focus on differences in the use of printed and digital textbooks, as well as addressing the question of the extent to which printed textbooks are being replaced by digital learning tools and other non-textbook materials and how this affects the quality of education. In this context, we invited authors to investigate the quality criteria for evaluating textbooks, such as whether they address individual learners' needs, maintain high standards of design and production, follow the aims and objectives of syllabuses or course programmes, use effective teaching and learning approaches and methods, are accompanied by various learning resources, are sensitive towards gender equality, etc. In response, we received a selection of theoretical and empirical articles that bring new insights into the role of textbooks in teaching and learning processes.

In the first article, entitled *The Changing Role of Textbooks in Primary Education in the Digital Era: What Can We Learn from Reading Research?,* the authors discuss how textbooks, a primary teaching tool for most children since the nineteenth century, have evolved from a linear to a dynamic layout, and how, in the last decades, screens have started to complement printed books as one of the main textbook substrates. There has been a great deal of research on how the content of textbooks has changed in line with changing values in different societies and over different periods of time, but little research has been done on how different types of textbook substrates and design have changed, and how these changes have affected learning and comprehension. The authors looked at a set of reading studies, studies on multimodal learning and PISA 2021 results to draw the conclusions that readers can learn about in the article. The article was produced as part of the project "Za kakovost slovenskih učbenikov" (For the Quality of Slovenian Textbooks), which is co-funded by the Republic of Slovenia and the European Union from the European Social Fund.

In the second article, entitled *Textbooks and Students' Knowledge*, the authors also focus on textbooks as a source of students' knowledge, by using data on student knowledge measured independently by national assessments and the Trends in International Mathematics and Science Study, TIMSS 2015. They explore the differences in knowledge and attitudes to learning between students who are taught by different textbooks. Although the study has considerable limitations due to missing data, the results of the analyses indicate some profound differences in knowledge and attitudes between groups of students using different textbooks. We would like to especially highlight the authors' conclusion and valuable recommendation for education policymakers "that these findings could serve as support for improving the criteria in the national system of validation of textbooks in the future. The link between the use of textbooks and student learning outcomes also highlights the need to systematically collect information on the use of textbooks among students and follow the effects on achievement in order to improve the quality of future textbooks." This article was also produced as part of the aforementioned project "Za kakovost slovenskih učbenikov".

The next article, entitled *Textbooks and Teaching Materials in Rural Schools: A Systematic Review*, presents a systematic review of research papers in the last decades aimed at analysing the concept of multigrade teaching resources and the teaching materials used by teachers in rural schools. Their use and dimensions are studied in order to promote inclusion and learning in multigrade classrooms, organised with children of different ages mixed together. The article effectively presents the challenges of multigrade didactic materials and their relation to teaching-learning processes. The need to personalise and adapt printed or digital textbooks and other teaching materials to involve students actively in the learning process and respond to the needs of rural students in multigrade classrooms is highlighted.

The following two articles explore different aspects of textbooks in mathematics. The first article, entitled *Differences in the Requirements of Digital and Printed Mathematics Textbooks: Focus on Geometry Chapters,* encompasses an analysis and comparison of the tasks in the printed and digital versions of the same mathematics textbook set, which covers Grades 1 to 4 of primary education in Croatia. The results show that both the printed and the digital textbook tasks have traditional requirements, with an emphasis on closed answer forms. The findings highlight the need for reflection on the meaningful use of the possibilities that the digital environment brings to textbook creation, and indicate that textbook authors do not yet know how to make full use of these possibilities in the creation of learning materials. In the second article on mathematics education, entitled *The Teaching of Initial Multiplication Concepts and Skills in Croatian Textbooks*, the authors examine the teaching of initial multiplication by comparing Croatian mathematics textbooks with textbooks from Singapore, Japan and England. The analysis provides evidence that practice and automation are the focus of the initial learning of multiplication in Croatia, and that students are not encouraged to use different calculation strategies in a flexible manner.

The next article, entitled *Nature of Science in Greek Secondary School Biology Textbooks,* analyses the presence of Nature of Science (NOS), an essential aspect of scientific literacy, in all Greek biology textbooks, workbooks, lab guides and teachers' books of lower secondary education. The authors find that most NOS references in the studied material were lacking the explicit references of NOS. The article also has a methodological value. The presented content analysis would be worth replicating in other school systems/countries, because NOS is essential for the learning of science and the achievement of scientific literacy (Lederman et al., 2013).

The last article of this focus issue is entitled *Theorising Textbook Adaptation in English Language Teaching.* It proposes a research-informed framework to contribute to a systematic description of textbook adaptation in foreign and second language teaching. It is worth mentioning that even though it is related to English language teaching, the article also has useful implications for research on textbook use in other disciplines.

We are very pleased to be able to publish an interview with Richard E. Mayer on the topic of multimedia materials and textbooks in this focus issue. The interview was conducted online in February 2021 (via an exchange of emails between Ljubljana, Slovenia, and Santa Barbara, California, USA). Multimedia learning theory (Mayer, 2001, 2005) is a fundamental theory for the design and use of textbooks and other types of multimedia learning tools. With this interview, we wanted to provide some important guidance for ensuring the present and future quality of textbooks and other educational materials.

A book review completes the contents of the focus issue. The book is entitled *The Palgrave Handbook of Textbook Studies* (2018) edited by Eckhardt Fuchs and Annekatrin Bock, published by Palgrave Macmillan (ISBN: 978-1-137-53141-4).

In addition to the focus issue, there are also four articles in the Varia section and a book review dealing with *Becoming Scientific: Developing Science* *across the Life-Course* (2020) written by Saima Salehjee and Mike Watts, published by Cambridge Scholars Publishing (ISBN: 1-5275-5498-8).

We wish you pleasant reading and abundant new insights.

Gregor Torkar, Miha Kovač and Mojca Kovač Šebart

References

Lederman, N. G., Lederman, J. S., & Antink, A. (2013). Nature of science and scientific inquiry as contexts for the learning of science and achievement of scientific literacy. *International Journal of Education in Mathematics, Science and Technology, 1*(3), 138–147.

Mayer, R. (2001). *Multimedia learning.* Cambridge University Press.

Mayer, R. E. (Eds.). (2005). *The Cambridge handbook of multimedia learning.* Cambridge University Press.

doi: 10.26529/cepsj.1290

The Changing Role of Textbooks in Primary Education in the Digital Era: What Can We Learn from Reading Research?

Miha Kovač[1] and Alenka Kepic Mohar*[2]

Textbooks have been the primary teaching tool since the 19th century. By their nature, they contained a comprehensive compilation of the content of a particular subject with the intention of explaining it; this knowledge, in turn, was usually filtered to conform to a particular society's expectations of elementary knowledge about the natural and social environments. There has been a great deal of research on how the content of textbooks has changed in line with changing values in different societies and over different periods. However, little research has been done on how textbook reading substrates and design have changed and how these changes have affected learning and comprehension: studies that systematically examined the effects of different reading substrates and different layouts on reading and learning comprehension did not appear until the late 20th century and early 21st century. We examine such studies and PISA 2021 results to draw five conclusions for future textbook research. These conclusions indicate that screens are worse than printed texts for some types of reading, while interactivity and dynamic design are not values per se but require coherent design to improve reading performance and higher-level thinking skills.

Keywords: textbook, educational publishing, reading, audio reading, reading substrate

1 Department of Library and Information Science and Book Studies, Faculty of Arts, University of Ljubljana, Slovenia.

2 *Corresponding Author. Mladinska knjiga Publishing House, Slovenia; Alenka.KepicMohar@mladinska-knjiga.si.

Spreminjajoča se vloga učbenikov v osnovnošolskem izobraževanju v digitalni dobi: česa se lahko naučimo s pomočjo raziskav branja

Miha Kovač in Alenka Kepic Mohar

Učbeniki so od 19. stoletja osnovno učno sredstvo. Po svoji naravi so vsebovali celovit povzetek vsebine določenega predmeta z namenom, da bi jo razložili; to znanje pa je bilo običajno filtrirano tako, da je ustrezalo pričakovanjem določene družbe glede osnovnega znanja o naravnem in družbenem okolju. Veliko je bilo raziskav o tem, kako se je vsebina učbenikov spreminjala skladno s spreminjajočimi se vrednotami v različnih družbah in v različnih časovnih obdobjih. Malo pa je bilo raziskav o tem, kako so se spreminjale bralne podlage in oblika učbenikov ter kako so te spremembe vplivale na učenje in razumevanje. Študije, ki so sistematično preučevale učinke različnih bralnih podlag in različnih oblik na branje in učno razumevanje, so se sistematično pojavile šele konec 20. stoletja in v začetku 21. stoletja. Preučili smo več takšnih študij in rezultate raziskave PISA 2021 ter oblikovali pet sklepov za prihodnje raziskave učbenikov. Ti kažejo, da so zasloni pri nekaterih vrstah branja slabši od tiskanih besedil, medtem ko interaktivnost in dinamična zasnova nista vrednoti sami po sebi, ampak za izboljšanje bralne uspešnosti in spretnosti mišljenja na višji ravni zahtevata skladno zasnovo.

Ključne besede: učbenik, izobraževalno založništvo, branje, zvočno branje, bralna podlaga

Introduction

For a long time, textbook research was primarily concerned with the content of textbooks and their embeddedness in cultural contexts (for more on this, see Fuchs & Bock, 2018). This paper examines textbooks from a different perspective: it focuses on the changing materiality and design of textbooks and asks what technological, social, didactic, and economic forces were behind these changes, whether changes in design and textbook substrates affected learning and reading processes, and whether these changes correlate with students' abilities to higher levels of reading such as deep reading. In the context of this paper, we see deep reading as an 'array of sophisticated processes that propel comprehension and that include inferential and deductive reasoning, analogical skills, critical analysis, reflection, and insight' (Wolf & Barzillai, 2009; see also Baron, 2021; Wolf, 2016, 2018). In short, we consider deep-reading abilities as something going beyond the bare extraction of information from the text: it forms one of the foundations of critical and creative thinking.

The aim of this paper is thus modest; as the correlations and causations among textbook design, textbook substrates, reading abilities, and critical thinking are rarely systematically researched, we will outline the reasons for the relevance of such research. By doing so, we want to contribute to the debate on creative thinking assessment measures that will be at the centre of PISA 2022 (OECD, 2019)

As a starting point, we will examine changes in textbook design that occurred in the previous 150 years and show that these changes correlate with the development of printing technologies. In the second step, we examine the results of three meta-studies on correlations between reading formats and reading comprehension as a counterpart to research on the reading of digital texts for children enriched with digital objects and on the effectiveness of multimodal learning tools. As we will show in the third step, there is enough circumstantial data to hypothesise that transformations of the textbook substrates and design have not been neutral, either in terms of their impact on reading comprehension or in terms of the societal and economic forces that have prompted them.

A terminological note: we take textbooks to mean a long-form text with artwork in codex format accessed in print or on a digital substrate, 'containing a comprehensive compilation of content in a branch of study with the intention of explaining it'(Wikipedia, 2021). We leave aside the workbooks with practise exercises and blank spaces for answers to be written directly into the book, nor do we take into consideration the digital-only interactive learning tools used primarily for repetitive learning and drill exercises. In the previous two decades, these learning tools have surrounded textbooks and, at least according to research in Slovenia, have begun to marginalise the role of textbooks in the learning process. The pandemic became an important catalyst for this process (Kepic Mohar & Kovač, 2021). We assume that similar processes are also taking place in other countries.

In the context of this paper, we consider textbook reading as decoding of content that is by default multimodal (i.e., textual and visual) and, in the case of digital textbooks, also auditory and augmented with mouse-over hyperlinks. Such an approach is based on the assumption that, even when using digital textbooks enhanced with audio and video objects, reading is central to learning. When learning is being done for the sake of memorisation, comprehension, inferential and deductive reasoning, reflection, and similar, we consider such reading to be long-form deep reading. When other modes of reading are considered, such as skimming and audio reading, for the clarity of argumentation, we avoid using the word 'reading'.

The Great Textbook Transformation

The history of textbooks in the late 19th and 20th centuries can be described in terms of changes in printing technology. For example, in examining how the visual representation of the same topic (i.e., history of ancient Greece in Slovene textbooks in the Habsburg Empire, Yugoslavia and Slovenia) has changed over the last 150 years, Kepic Mohar (2019) has shown how the layout has evolved from a linear text interrupted only by chapters, subchapters, and occasional boldface sentences in the 1870s to a linear text interrupted by occasional black-and-white photographs and captions in the 1930s. For the purposes of this article, we will refer to such linear text-based organisation of content as 'linear layout'.

Six decades later, in the 1990s, textbooks with similar content were printed in four colours; the text occupied only two thirds of the page and was supplemented with explanations, drawings, photos, and instructional coloured maps. In order to introduce various levels of text, different typography was used, which appeared against coloured backgrounds. In the 1990s, some textbooks were accompanied by CDs with additional instructional materials, and a CD icon marked the places where students could use supplemental digital learning materials (predominantly video or audio).

Over the next decade, the complexity of the printed textbook continued to evolve: in addition to icons directing students to the CDs, the double-page spread featured artwork with tables, keywords, and collocations in the outer margins. The main text was printed in black, and keywords were in semi-bold type. Motivational texts were printed in different colours at the beginning of chapters, and metadata was located at the top and bottom of the page, indicating the main content as explained on the page. In general, numerous elements have emerged that no longer consist of text alone so that the reader decides immediately which block to read next, the reading order being dictated in part by the typography and the various graphic blocks and not necessarily by the linearity of the main text.

On the other end of Europe, in the United Kingdom, similar transformations of textbook layout were described by Bezemer and Kress (2016), comparing English textbooks for Science, Mathematics and English in the 1930s, 1980s, and 2000s:

> Textbooks from the 1930s are A5 sized or smaller. Their pages are, typically, designed following a rigid grid, in a single column, with consistent margins, baselines, headers and footers, allowing the writing to flow continuously from one column to the next from top left to bottom right; it runs across pages. In the 2000s, the book is bigger, and we see a move away from the rather rigid, writing-driven grid which was common in the 1980s. We also see an increase in the use of the two-page spread from the 1980s, providing an entirely different 'canvas' or 'site of display'. Most textbooks now use varying numbers of columns per page, varying column widths, allowing writing to be 'wrapped around' - often irregularly shaped - images. Writing may still be running across pages but more often page breaks coincide with separations of different parts of the text, marked off by line boxes and background colours. (pp. 15–17)

This complex structure finally migrated from the printed page to a digital textbook with a reflowable layout and hyperlinks. In this way, the limitations of textbooks by the codex format were put to an end and (with a series of hyperlinks and digital materials) a book with no real beginning or end was created, allowing the learner to access a variety of online content (more Kepic Mohar, 2019). For the purposes of this article, we will refer to such complex organisation of content either in print or on-screen as a dynamic layout.

As if in a parable, with all these layout changes, in a hundred years, an ordinary spruce became a Christmas tree. All these changes correspond with developments in printing and screen technology. In the 1870s, the use of photography in printed publications was still in its infancy and far from mass use; in the 1930s, the development of printing enabled the mass use of photographs, especially in monochrome publications; and in the 1990s, four-colour printing became economical enough to be widely used in publications with relatively large print runs, such as textbooks (see more: History of printing timeline, 2020). Last but not least, in the second half of the 2010s, especially after the wide usage of the mobile phone, screen technologies became commonplace in education, and all kinds of educational content moved from print to screen media and (as we have indicated above) textbooks were no exception. Much like their counterparts publishing magazines, newspapers, mass-market non-fiction books and commercial catalogues, educational publishers went with the flow, changing and improving textbook layouts at the same pace as print and later screen technology evolved and made such changes possible.

As the results of studies done by Kepic Mohar (2019) and Bezemer and Kress (2016) indicate, at least in Europe, these changes occurred at about the same time, yet they were introduced without any systematic didactic evaluation of their impact on the learning process: with our modest research reach, we could not find any study in the 20th century that addressed the question of whether students understand and memorise information better when they learn from textbooks with linear text occasionally interrupted by photographs, chapters, and subchapters, or whether the learning process is more effective when they use textbooks with dynamic layouts. The didactic superiority of multicolour textbooks with dynamic layouts compared to their older relatives with linear layouts was simply taken for granted.

We can only hypothesise why this change has occurred without in-depth debate and research on the didactic implications of different textual and graphic design of learning materials. The simplest and most straightforward answer seems to be that, for practical reasons, it has been almost impossible to design such studies. For measuring the impact of layout on the learning process, it would require a group of students to learn from two types of textbooks with the same content, one with a linear layout and one with a dynamic layout (research with within-participant design) and then, using interviews to assess how their understanding correlates with the layout of the textbooks with which they used to learn. Alternatively, each participant could read on both paper and digital presentations and then, their comprehension when learning from each type of layout could be measured and compared.

The lesson of current reading research on comprehension differences between reading from print and screen is that such studies are not replicable in the same way as experiments in natural sciences: no two texts are exactly alike, and because of a set of uncontrollable variables, the attitudes and focus of participants in study can vary from experiment to experiment, thus leading to variations in results. Therefore, these results can only be considered significant when a set of studies using the same methodology, conducted in different countries by similarly educated participants of similar age, yield similar results. (For more on this, see Clinton, 2019; Delgado et al., 2018; Singer & Alexander; 2017).

As far as we know, throughout the 20th century, there was never a situation in which textbooks with linear and dynamic layouts coexisted, so there were no tools at hand to conduct such studies, even less in a number that would allow meta-analysis. More to the point, because of the slow pace of textbook development (as mentioned, the evolution of layout from linear to dynamic was gradual and took about a hundred years), it is likely that it never occurred to researchers that there might be a correlation or even causality between textbook layout and comprehension. The social and historical context simply did not lead to such research questions being asked and studied, thus turning the impact of long-term changes in textbook design on learning and comprehension into a blind spot of pedagogy.

With the digital transformation, all that changed. Its pace was faster than the evolution of print in the 19th and 20th centuries, and suddenly, with the advent of e-books and electronic textbooks, the same fiction and informational texts existed in print and on the screen. Since the late 1980s and early 1990s, this triggered a body of research in reading studies on the differences between print and screen reading. By the 2010s, there were a few hundred research papers on this topic (for more, see Delgado, 2018; Singer & Alexander, 2017). However, the methodology used in these papers was not consistent, making the overall comparison of results quite complicated. These discrepancies led to three different meta-studies comparing reading studies with similar research designs (Clinton, 2019; Delgado et al., 2018; Singer & Alexander, 2017). Surprisingly, all three meta-studies came to similar conclusions: When reading long informational texts, comprehension is better when the text is read in print than in the screen version. Furthermore, Delgado et al. (2018) found that the effect of screen inferiority has increased over the past 18 years, which is consistent with findings that digital technology is having a detrimental impact on students' comprehension skills and indicates that so-called digital natives perform worse in screen reading than digital migrants (see for example also Duncan et al., 2015; Pfost et al., 2013). In summary, as long as the replicability of research results remains the golden standard of science, we can conclude that print is a more suitable medium for reading longer linear informational texts than screens.

All of these studies were primarily conducted with university students and never considered primary and secondary textbooks. How can such findings about the reading of long-form linear texts be applied to textbooks, where the linear layout was replaced by a dynamic one some sixty years ago?

Not all textual digital content is created equal

One of the clues to how textbook layout affects comprehension may be hidden in a few reading studies on how the understanding of digital texts changes when enriched with various digital objects. In a research synthesis of 29 studies involving 1272 young children, Takacs et al. (2015) found evidence that multimedia stories were more conducive to story comprehension and word learning than when children were exposed to the same stories in a linear format in print without adult support; moreover, there were no differences between the benefits of multimedia elements embedded in the text (such as animated illustrations, background music, and sound effects) read without an adult and reading the linear text with adult support. However, interactive elements such as hotspots and games were not found to be beneficial for story comprehension as they require switching between the story and the interactive elements, thus interfering with story comprehension and language acquisition. Similarly to Takacs et al. (2015), Bus et al. (2014) also point out that interactions that have only a decorative function interfere with reading and learning, thus undermining comprehension. We can therefore hypothesise that the detrimental effects of overly dynamic print layout are similar to those of overly enriched digital texts; furthermore, if the visual elements are not decorations that distract attention from the main text but reinforce its content and help the reader follow the narrative in a way that visual and textual elements are coherent, a dynamic, multi-modal print layout could enhance comprehension.

Another interesting reading study was conducted by cognitive and metacognitive researchers (Sidi et al., 2017), who tested the inferiority of reading on screens compared to paper in effort regulation, test performance, and levels of overconfidence. The researchers hypothesised that the medium would provide a contextual cue that would lead to shallower on-screen processing regardless of text length, especially when task characteristics indicated that shallow processing was legitimate. The results suggest that metacognitive processes are sensitive to contextual cues that indicate expected processing depth, regardless of the associated reading burden involved (Sidi et al., 2017). Similarly, several studies show that a smartphone is a distraction even if it is just sitting on the table next to us (Thorton, 2014; Ward, 2017; for more on smartphone distractions and downfalls, see also Spitzer, 2019).

There is one more layout and substrate caveat coming from reading studies: not all screen media are created equal. Dedicated reading devices, for example, are used only for long-form reading, and they do not allow distractions, thus providing different contextual cues than smartphones: consequently, it might make a difference whether reading the text on a dedicated reader like Kindle or Kobo than on a smartphone. Again, we did not find any study examining such differences. However, there is some circumstantial evidence to suggest that such differences indeed do exist: Mangen et al.'s (2019) study of the differences between reading on a print medium and reading on a Kindle suggests that comprehension of long narrative texts does not differ when reading on these two media; the differences only occur in terms of temporal and spatial orientation in the text, with reading on a print medium yielding better results than reading on a Kindle. The authors concluded that this is likely because the reading device does not provide the same sensorimotor cues as the printed book, where the reader can tangibly see how much they have read and how much is left to read in the book (see also Baron, 2021).

These findings were confirmed, again circumstantially, by Salmeron et al. (2018), who conducted a study on the differences between reading photocopies and reading authentic documents such as books, magazines, and printed newspapers. The study found that students remember better when they read from authentic documents. Again, the researchers assumed that this was due to the visual and tactile characteristics of the documents used (see also Baron, 2021). Another factor that contributes to decreased spatial and temporal orientation when reading from screens might be scrolling, as it does not provide markers for beginnings and ends, while the borders of a particular page give the reader a sense of location (for more on this, see Baron 2021, pp. 87–91).

To summarise, if there are no significant differences in reading comprehension when reading print and from dedicated reading devices, but there are differences in reading comprehension when reading print, and from phones/tablets, we can assume that there are also differences in reading comprehension when reading from phones and from dedicated reading devices. This allows us to hypothesise that what also matters in digital reading is what kind of device is used for reading. Less distracting devices allow for better comprehension, but when temporal and spatial information is important, print is still better than digital.

If we apply these findings to textbooks, we can hypothesise that the adverse effects of the contextual cues of screen medium and of the (too) dynamic layout on comprehension might be replaced by a) dedicated reading/learning devices and b) artwork (in print) and multimedia elements (on-screen medium) that support rather than interfere with the main text; if the latter is the case, the detrimental effects of screens and/or too dynamic layout are exacerbated.

These findings correspond with the cognitive load theory (Clark et al., 2006) and the cognitive theory of multimedia learning theory (Mayer, 2020) which postulates that optimal learning occurs when visual and verbal learning materials are presented simultaneously (Torkar, 2021). Similar to Takacs et al. and Bus et al., on the basis of a set of studies on this issue, Richard Mayer states (in Torkar, 2021):

> If I had to choose one principle for revising textbooks, I would start by choosing the coherence principle and seek to remove irrelevant and distracting elements so students can focus on learning the essential material in the lesson. Next, I would add the spatial contiguity principle, which calls for removing the captions on figures and moving the essential text (in segments) next to the corresponding part of the graphic. When a textbook has graphics with long captions or legends, that is an indication of poor design. (p. 3)

Such findings should be seen as an additional warning that dynamic layouts and highly interactive and multimodal textbooks are not by default a value per se, yet when adequately designed, they could represent a better match. The PISA 2021 Report on 21st Century Readers (OECD, 2021) confirmed such conclusions and underscored the importance of linear fiction texts for reading comprehension as it found that:

> […] a higher frequency of reading fiction texts, texts that include tables and graphs, and texts that include diagrams more frequently is significantly associated with reading performance after accounting for students and schools' socio-economic profile on average across OECD countries' (p. 121)

In contrast, digital texts with links 'show a negative association with reading performance after accounting for students and schools' socio-economic profile' (OECD, 2021, p. 121), while countries in which students have to read 'longer pieces of text for school (101 pages or more) achieved 31 points more in reading than those who reported reading smaller pieces of text (10 pages or less) after accounting for students' and schools' socio-economic profiles and students' genre' (OECD, 2021, p. 120).

In the language of this paper, using only digital learning tools with an overly dynamic layout leads to lower reading performance.

New kid on the block: audio

Audio brings additional complexities to the assessment of learning tools. Audiobooks have a long history, back to the days of vinyl records and later audio cassettes. However, until the early 21st century, audiobooks remained a niche; other than being used by the visually impaired and commuters, they never gained a significant share of the book market (for more on the history of audiobooks, see Rubery, 2016). That began to change with the advent of smartphones. Data from the U.S. market for 2020, for example, showed that one in six books was sold in digital audio format, consumed predominantly through online audiobook subscription services accessed by smartphones. Adult non-fiction accounted for the highest share, and the most popular categories were business, self-help, and humour (Audio Publishers Association, 2021). On the other side of the Atlantic, in Sweden, digital books accounted for about half of the total fiction market in 2020, and about 90% of these books were audiobooks, consumed mostly on subscription platforms, similar to the US (more on the Swedish book market: Bokförsäljningsstatistiken, 2020 and The Swedish Book Market, 2020).

This considerable growth of the audiobook market has stimulated interest in exploring the comprehension and memorisation differences between reading and listening to textual content. Two views of listening to fiction and non-fiction trade audiobooks have emerged: the dual-process view assumes that listening and reading share some elements but are essentially two separate cognitive processes; the unitary process view, in contrast, assumes that the same comprehension mechanisms underlie both processes (for more on this, see Baron 2021, pp. 165–170). At the time of writing this text, the latter view seemed to prevail and even gave rise to the rather controversial notion of audio reading that was used not only in audiobook publishing but also among some researchers.

The reason we find this notion questionable is that it ultimately leads to a paradox: since audiobooks can be listened to by illiterate people, the notion of audio reading hints that illiterate people can also read or can learn to read by listening. However, reading and listening involve two different sensory systems, the auditory and the visual: when we read visually, we connect signs with sounds and assemble sounds into words and words into sentences, thus making meaning in our minds (for more on the reading process, see Dehaene, 2009; Willingham, 2017; Wolf, 2008;). None of this happens when an illiterate person listens only to audio-only content, clearly suggesting that audio reading cannot make an illiterate person literate.

From this point of view, it is not surprising that (as found by Diakidoy et al., 2005), children in early primary education have better comprehension when listening than reading, which is reversed when children are older and learn to read and write fluently. In psychological terms, younger children use all of their working memory for decoding when reading, whereas older children, who have already automated their decoding system, have more processing space left in working memory for comprehension. These findings were confirmed by Daniel and Woody (2010), who conducted a study with two groups of older students, one of whom learned by listening to podcasts and the other by reading the text. The podcast group performed worse than the students who read the text, leading the researchers to conclude that while podcasts can be a useful tool for supplementing course-related material, they are not as effective as texts in teaching primary content.

As Baron suggests, the better comprehension in reading compared to audio is due to the absence of several aids and signalling devices that facilitate comprehension when reading printed text or on a screen: in listening, unlike reading written texts, we have no control over the pace, re-listening is much more tedious than re-reading, it is virtually impossible to skim pages, and there are no markers such as paragraphs and subheadings that are present in texts to help readers orient themselves (Baron 2021, p. 166). Even more, some studies have found that students digress more when listening than when reading (Barao Sousa in Baron, 2021).

Nevertheless, as the field of digital audio learning tools is still in its early stages, so is the research about it: more studies need to be conducted to draw definitive conclusions about the effectiveness of audio learning tools in primary and secondary education. Learning Ally (n.d.), a US non-profit volunteer organisation (previously named Recording for the Blind and Dyslexic) for example, developed a set of audiobooks that allow parallel listening and reading and, according to them, such texts with parallel audio input significantly help struggling readers and readers with dyslexia. In other words, we know from the pre-digital era that audio can be a helpful learning tool in language and music courses, and there is some evidence that the combination of text and audio helps struggling readers and students; nevertheless, it remains to be seen how audio can adequately complement text materials as one of the primary tools in education.

Conclusion: What does reading research tell us about textbooks?

If we summarise all that we have found about the impact of textbook reading substrates and design on comprehension and learning, five critical conclusions regarding the design of printed and digital textbooks stand out:

1. In the pre-digital era, changes in the design of learning tools were not systematically evaluated from the perspective of reading and learning

comprehension and were taken for granted as a natural result of technological development.

2. When reading long-form linear informational texts, print seems to be a better medium than screens. As this finding was confirmed in three meta-studies covering a few hundred studies, we can consider this difference between print and digital reading to be an established fact.
3. When reading from screens, circumstantial evidence indicates it matters what kind of reading device is being used. Less disruptive devices, such as dedicated reading devices, afford better comprehension than smartphones and tablets. However, when temporal and spatial information matters, circumstantial evidence indicates that print remains superior in comparison to all screen devices.
4. As shown by a meta-study, at least in reading materials for children, the inferiority of screens could be compensated by visual and audio objects that support the main narrative of the text. However, if these objects disrupt the main narrative (such as hotspots and games), the inferiority of the screen increases. Studies on multimedia learning tools produced similar conclusions. PISA 2021 results indicate that besides properly designed digital learning tools, long-form linear reading significantly contributes to reading performance and, consequently, to critical thinking (OECD, 2021). From this point of view, combining coherently designed print and digital learning tools could be the optimal solution.
5. Audio-only seems to be an inferior format to textual learning tools. However, there is evidence that in combination with textual media it may be helpful for struggling readers. This indicates that a combination of experimentation and evaluations will be needed to find the proper place for audio among learning tools.

When looking for answers to these dilemmas, two issues stand out: a technical one and a cultural one.

The cultural question refers to the general understanding of the role of long-form reading in contemporary civilisation. Authors from Birkerts (1993) to Baron (2021) have warned that with digital media, the human mindset is changing in such a way that our ability to read long-form texts is declining or not developing (see also Carr, 2011; Firth et al., 2019; Kovač & van der Weel, 2018; Wolf, 2018). The reason for these warnings is that one of the positive externalities of reading books/long-form linear texts is the broadening and deepening of readers' vocabulary and the training of focus, which is necessary for the acquisition of critical cognitive skills, such as logical and abstract thinking. These positive externalities diminish with the use of digital media: As we have shown, screens are inferior to print when reading informational long-form texts. In addition, screens are primarily viewed as entertainment tools with which immersion is driven by visual stimuli, whereas the concentration required to read print texts is just the opposite: a concentration that is immune to distractions outside the text. As we have shown, coherently designed multimodal textbooks can surpass these flaws of screens in comparison to print, yet at least, for now, there are no studies showing that digital media can entirely replace printed books in performing these tasks.

Regarding long-form linear reading, print seems to remain the optimal medium. Therefore, if analytical thinking and the ability to describe and discuss complex social and natural phenomena remain desirable societal values and learning outcomes, schools should train students to use printed textbooks and read printed books in combination with using properly designed digital media, thus keeping the printed textbook as one of the core learning tools, at least until we have solid evidence that screen media can do this job better without exceptions than their printed predecessors. As shown in a study conducted in Slovenia, textbooks might be losing this position in learning processes (Kepic Mohar & Kovač, 2021). If we are correct, this might lower reading literacy as measured by PISA in the coming years.

The technical question relates to the evaluation of changes in the design of learning tools. Both the pace of technological development of screen and print media on the one hand and available research equipment on the other enable studying how pupils and students use the learning tools: by using eye trackers, for example, we can determine whether the design of the learning tool is distracting, or that allows students to use it in a way that makes it easy for them to follow the narrative and understand the content. PISA provides an enormous amount of data on correlations between reading performance and reading substrates and formats. As we have shown, the PISA data and results of studies on reading comprehension and on multimodal learning can produce robust results on correlations between the medium and reading performance, while raising a new set of research questions on why and when using which learning tools.

In short, for the first time in history, educational publishers will find themselves in a privileged position to develop learning tools by tinkering with them while evaluating their instructional impact. It would be a pity to squander this opportunity by taking the benefits of digital media for granted.

Acknowledgement

The article was produced as part of the project "Za kakovost slovenskih učbenikov" (For the Quality of Slovenian Textbooks, https://kauc.splet.arnes.si), which is co-funded by the Republic of Slovenia and the European Union from the European Social Fund.

References

American Printing History Association. (2020, October 10). *History of printing timeline. American Printing History Association.* https://printinghistory.org/timeline/

Audio Publishers Association. (2021). A first look at digital audiobook tracking. https://www.audiopub.org/blog/a-first-look-at-digital-audiobook-tracking-from-the-npd-group

Baron, N. (2021). *How we read now*. Oxford University Press.

Benson, J. (2021). *A first look at digital audiobook tracking from the NPD Group*. A first look at digital audiobook tracking from The NPD Group - APA (en-US). https://www.audiopub.org/blog/a-first-look-at-digital-audiobook-tracking-from-the-npd-group

Bezemer, J., &Kress, G. (2016). The textbook in a changing multimodal landscape. In N. M. Klug & H. Stöckl (Eds.), *Handbuch Sprache im multimodalen Kontext* [Handbook of language in multimodal contexts]. De Gruyter.

Birkerts, S. (1994). *The Gutenberg elegies. The fate of reading in an electronic age*. Faber and Faber.

Bokförsäljningsstatistiken. (2020). https://forlaggare.se/bokforsaljningsstatistik/

Bus, A. & Takacs, Z., & Kegel, C. (2014). Affordances and limitations of electronic storybooks for young children's emergent literacy. *Developmental Review, 19*. 10.1016/j.dr.2014.12.004

Carr, N. (2011). *The shallows*. Atlantic Books.

Clark, R., & Mayer, R. (2016). *E-learning and the science of instruction*. John Wiley & Sons.

Clark, R., Nguyen, F., & Sweller, J. (2006). *Efficiency in learning*. John Wiley & Sons.

Clinton, V. (2019). Reading from paper compared to screens: A systematic review and meta-analysis. *Journal of Research in Reading*, *42*(2), 288–325. https://doi.org/10.1111/1467-9817.12269

Daniel, D. B., & Woody, W. D. (2010). They hear, but do not listen: Retention for podcasted material in a classroom context. *Teaching of Psychology, 37*(3), 199–203. https://doi.org/10.1080/00986283.2010.488542

Dehaene, S. (2009). *Reading in the Brain*. Penguin Books.

Delgado, P., Vargas, C., Ackerman, R., & Salmerón, L. (2018). Don't Throw away your printed books: A meta-analysis on the effects of reading media on reading comprehension. *Educational Research Review*, *25*, 23–38. https://doi.org/10.1016/j.edurev.2018.09.003

Diakidoy, I.-A. N., Stylianou, P., Karefillidou, C., & Papageorgiou, P. (2005). The relationship between listening and reading comprehension of different types of text at increasing grade levels. *Reading Psychology*, *26*(1), 55–80. https://doi.org/10.1080/02702710590910584

Duncan, L. G., McGeown, S. P., Griffiths, Y. M., Stothard, S. E., & Dobai, A. (2015). Adolescent reading skill and engagement with digital and traditional literacies as predictors of reading comprehension. *British Journal of Psychology*, *107*(2), 209–238. https://doi.org/10.1111/bjop.12134

Firth, J., Torous, J., Stubbs, B., Firth, J. A., Steiner, G. Z., Smith, L., Alvarez-Jimenez, M., Gleeson, J., Vancampfort, D., Armitage, C. J., & Sarris, J. (2019). The "Online Brain": How the internet may be changing our cognition. *World Psychiatry*, *18*(2), 119–129. https://doi.org/10.1002/wps.20617

Fuchs, E., &Brock, A. (Eds.) (2018). *Introduction to the Palgrave handbook on textbook studies*. Palgrave Macmillan.

Kepic Mohar, A. (2019). The materiality of textbooks. *Logos, 30*(2), 26–33.

Kepic Mohar, A., & Kovač, M. (2021). Digitalna učna gradiva po pandemiji kot del šolske rutine? [Digital learning materials after pandemic as part of the school routine?]. *Sodobna pedagogika*, *72*(Special Issue), 28–43.

Kovač, M., & van der Weel, A. (2018). Reading in a post-textual era. *First Monday*, *23*(10). http://firstmonday.org/ojs/index.php/fm/article/view/9416/7592 https://doi.org/10.5210/fm.v23i10.9416

Learning ally. Learning Ally - Together It's Possible. (n.d.). https://learningally.org/

Mangen, A., Olivier, G., & Velay, J.-L. (February 15, 2019). Comparing comprehension of a long text read in print book and on Kindle: Where in the text and when in the story? *Frontiers in Psychology*, 10. https://doi.org/10.3389/fpsyg.2019.00038

Mayer, R. (2020). *Multimedia learning*. Cambridge University Press.

OECD. (2019). *Pisa 2021 creative thinking framework - OECD*. https://www.oecd.org/pisa/publications/PISA-2021-creative-thinking-framework.pdf

OECD. (2021). *21st-century readers: Developing literacy skills in a digital world*. OECD. https://doi.org/10.1787/a83d84cb-en

Pfost, M., Dörfler, T., & Artelt, C. (2013). Students' extracurricular reading behavior and the development of vocabulary and reading comprehension. *Learning and Individual Differences*, *26*, 89–102. https://doi.org/10.1016/j.lindif.2013.04.008

Rubery, M. (2016). *The untold story of the talking book*. Harvard University Press.

Sidi, Y., Shpigelman, M., Zalmanov, H., & Ackerman, R. (2017). Understanding metacognitive inferiority on screen by exposing cues for depth of processing. Learning and Instruction, *51*, 61–73. https://doi.org/10.1016/j.learninstruc.2017.01.002

Singer, L., &Alexander, P. (2017). Reading on paper and digitally: What the past decades of empirical research reveal. *Review of Educational Research*, *87*(6), 1007–1041.

Spitzer, M. (2019). *Die smarthphone pandemie* [The smartphone pandemic]. Klett-Cota Verlag.

Salmerón, L., Gil, L., & Bråten, I. (2018). Effects of reading real versus print-out versions of multiple documents on students' sourcing and Integrated Understanding. *Contemporary Educational Psychology*, *52*, 25–35. https://doi.org/10.1016/j.cedpsych.2017.12.002

The Swedish Book Market. (n.d.). https://www.kulturradet.se/en/swedishliterature/learn-more-swedish-book-market/

Thornton, B., Faires, A., Robbins, M., & Rollins, E. (2014). The mere presence of a cell phone may be distracting: Implications for attention and task performance. *Social Psychology 45*(6), 479–488.

Takacs, Z. K., Swart, E. K., & Bus, A. G. (2015). Benefits and pitfalls of Multimedia and interactive features in technology-enhanced storybooks. *Review of Educational Research*, *85*(4), 698–739. https://doi.org/10.3102/0034654314566989

Torkar, G. (2021). Interview with Richard E. Mayer about multimedia materials and textbooks. *CEPS Journal*. https://doi.org/10.26529/cepsj.1238

Ward, A. F., Duke K., Gneezy, A., & Boss, M.W. (2017). Brain drain: The mere presence of one's own smartphone reduces available cognitive capacity. *Journal of the Association for consumer research,* 2(2), 140–154.

Wikimedia Foundation. (2021, May 27). *Textbook*. Wikipedia. https://en.wikipedia.org/wiki/Textbook

Willingham, D. T. (2017). The *reading mind: A cognitive approach to understanding how the mind reads.* Jossey-Bass.

Wolf, M., & Barzillai, M. (2009). *The importance of deep reading. Challenging the whole child: Reflections on best practices in learning, teaching, and leadership*. ASCD.

Wolf, M. (2008). *Proust and the squid*. Harper Collins Publishers.

Wolf, M. (2016). *Tales of literacy for the 21st century: The literary agenda*. Oxford University Press.

Wolf, M. (2018). *Reader, come home. The reading brain in a digital world*. HarperCollins.

Biographical note

Miha Kovač, PhD, is professor at the Department of Library and Information Science and Book Studies at the University of Ljubljana. In his publishing career in late eighties and nineties, he worked as editorial director in the two largest Slovene publishing houses, DZS and Mladinska knjiga and as an editorial director of Slovene edition of National Geographic Magazine. He also worked as a consultant to textbook publishers in variety of Central and East European countries.

Alenka Kepic Mohar, PhD, is Editor-In-Chief at Mladinska knjiga Publishing House. Her research interests include reading and learning in the digital era and the role of textbooks in the digitalization. She combines practical experiences from educational publishing with scholarly research to encourage dialogue between academic researchers, publishing companies, and educational policy involved in the production of digital reading materials.

doi: 10.26529/cepsj.1283

Textbooks and Students' Knowledge

Barbara Japelj Pavešić*[1] and Gašper Cankar[2]

In Slovenia, textbooks are an integral part of the curriculum. Nationally certified textbooks guarantee both teachers and students that they provide all of the necessary knowledge in each subject. There are many available certified textbooks for each subject and teachers must decide which will be the source of instruction for their students. Our research question is whether groups of students who use different textbooks as their mandatory learning resource differ in their knowledge and their attitudes towards learning. We linked existing data from several sources and explored the scope of the use of different textbooks for mathematics and science subjects in primary schools. Data on student knowledge measured independently by National Assessments (and the Trends in International Mathematics and Science Study were used, and differences in knowledge and attitudes to learning between students who are taught using different textbooks were explored. Although the study has considerable limitations due to missing data, the results of the analyses indicate some profound differences in knowledge and attitudes between groups of students using different textbooks. These findings could serve as a guide for teachers when choosing the optimal available textbook for their students and, even more so, as support for improving the criteria in the national system of validation of textbooks in the future. The link between the use of textbooks and student learning outcomes also highlights the need to systematically collect information on the use of textbooks among students and follow the effects on achievement in order to improve the quality of future textbooks.

Keywords: certified textbooks, knowledge, attitudes, mathematics, science

1 *Corresponding Author. Educational Research Institute, Slovenia; barbara.japelj@pei.si.
2 National Examinations Centre, Slovenia.

Učbeniki in znanje učencev

Barbara Japelj Pavešić in Gašper Cankar

V Sloveniji so učbeniki sestavni del kurikuluma. Nacionalno potrjeni učbeniki učiteljem in učencem zagotavljajo, da bodo z njimi dosegli vsebinske cilje učnega načrta pri vsakem predmetu. Potrjenih učbenikov za določene predmete je na voljo več in učitelji se odločijo, kateri bodo obvezni vir pouka za njihove učence. Vprašanje je, ali se razlikujejo znanje in stališča učencev, ki uporabljajo različne učbenike kot svoj obvezni učni vir. Združili smo obstoječe podatke iz več virov ter raziskali obseg uporabe različnih učbenikov za matematiko in naravoslovne predmete v osnovni šoli. Uporabili smo obstoječe in medsebojno neodvisne podatke o znanju učencev z nacionalnih preizkusov znanja ter mednarodne raziskave znanja matematike in naravoslovja, TIMSS. Preverili smo razlike v stališčih do učenja med učenci, ki se učijo z različnimi učbeniki. Čeprav ima študija precej omejitev zaradi nekaterih manjkajočih podatkov, so rezultati analiz nakazali nekatere bistvene razlike med skupinami učencev, ki uporabljajo različne učbenike, v znanju in stališčih. Ugotavljamo, da bi bile lahko te opora učiteljem pri izbiri najboljšega mogočega razpoložljivega učbenika za svoje učence, še bolj pa podpora za dvig meril v nacionalnem sistemu potrjevanja učbenikov v prihodnje. Povezanost med uporabo učbenikov in učenjem učencev vzpodbuja tudi potrebo po sistematičnem zbiranju informacij o uporabi nacionalno potrjenih učbenikov med učenci ter povratnih informacijah o njihovih učinkih na pouk za izboljšanje kakovosti prihodnjih učbeniških gradiv.

Ključne besede: potrjeni učbeniki, znanje, stališča, matematika, naravoslovje

Introduction

Textbooks are a fundamental written source of knowledge and often represent a factor in educational research that explains differences in students' knowledge at different levels of the education system (Oakes & Saunders, 2004). As an 'embodiment' of the prescribed national curriculum for an individual school subject and year of schooling, they transfer the requirements and expectations regarding student outcomes at the national level into teaching practice (Oates, 2014). As a resource for teaching, textbooks present teachers with a set of learning objectives and their transformation into didactic presentations of the material intended for teaching. As a learning tool, they are a fundamental and reliable source of information for students as they acquire new knowledge, while offering the teacher the opportunity to focus attention on improving pedagogy and effective learning (Oates, 2014). The role of textbooks is multifaceted and crucial for pedagogical practice.

Textbooks link the expected, implemented and achieved curriculum as different curricular levels are understood by international large scale assessments (Johansson, 2003). The expected curriculum represents a national consensus on the content to be presented and taught to students, the implemented curriculum describes the process of teaching this material to students in the classroom, and the achieved curriculum is students' knowledge as demonstrated by their achievements. In the present study, we were interested in how the use of a textbook affects students' knowledge. Specifically, we investigate how the attitude towards textbooks at three levels in the curriculum in Slovenia compares to the situation in other countries, and whether it is possible to describe a relationship between students' achievement and textbooks, and, consequently, certain characteristics of teaching practices.

In general, the research literature suggests that textbooks should be considered as an important parameter in education studies and textbook choice a relevant factor for education practice. However, comparative studies of the relationship between different textbooks and student achievement are rare, with most of the available reports concerning learning mathematics. Researchers observe that teachers' choice of textbooks substantially impacts students' mathematics achievement, while individual textbooks differ in their effects. This impact was found to be cumulative over the school years and significantly linked to the problem-solving strategies students use from as early as Grade 1 (Sievert et al., 2021; Van den Ham & Heinze, 2018).

Researchers observing relationships between achievement, curriculum and textbooks in mathematics found little variation in achievement between different textbooks, but some variation between the curriculum that was implemented before and after a school reform in the US (Blazar et al., 2020). They assumed that textbooks became increasingly similar due to the reform, as it changed local curricula to the unified National Core Curriculum in many states. However, researchers suggest that excellent teaching and high student achievement cannot be achieved by the improvement of textbooks alone. A similar message about small differences between students' knowledge according to the use of different textbooks could be concluded for the case of Slovenia, where the national curriculum is already prescribed for all students.

Textbooks should also be seen as a pedagogical tool, as they can lend support to teachers when they are choosing teaching strategies. Examining how teachers' subject knowledge relates to their use of textbooks, Mili and Winch (2019) claim that textbooks can be a powerful pedagogical tool for highly trained teachers, as opposed to being a teaching script in the hands of poorly qualified teachers. Regarding open educational resources, which are increasingly attractive for facilitating distance learning and reducing education costs, research on their effect on student achievement shows that teachers matter more than textbooks. The results of an analysis by Hardin et al. (2019) reveal that traditional textbooks do not necessarily ensure greater academic success, even in the hands of well-trained instructors, thus suggesting a need to devote resources to teacher training and support in the use of different textbooks.

Comparisons within Slovenian mathematics education have shown that teachers in Slovenia lag behind other countries in teaching decimals and fractions, as these topics are covered later (Mullis et al., 2016). Another very important message emerges from recent research on the relationships between textbook content and mathematics learning, focusing on how children learn about rational numbers (i.e., fractions, decimals, percentages) (Siegler & Oppenzato, 2021). It is known from previous research that fraction knowledge in the fifth grade uniquely predicts mathematics achievement five years later in secondary school, even after controlling for socioeconomic status, IQ, reading comprehension and whole number arithmetic knowledge. However, recent findings suggest that children's learning depends on items and problems that are presented (or absent) in their textbooks. Differences in achievement were strongly associated with differences in the composition of different types of exercises in the textbooks used. More demanding textbooks with a diverse range of different problems seem to elicit higher achievement. In national and international assessments of mathematics in Slovenia, extremely low achievement in items that are not presented in textbooks in a similar form, or not presented at all, is traditionally observed, as we will demonstrate later.

The research problem

The importance of textbooks in educational research is growing as the research literature shows a trend towards the increased use of textbooks as the primary source of teaching, especially in mathematics at Grade 8 (Mullis et al., 2012). This could mean that the content taught in class is increasingly limited to the presentations and explanations of concepts given in selected students' textbooks. In other countries, the intensive use of textbooks as a primary sources of teaching has not lead to increased effectiveness in learning, especially not in science. Data from the Trends in International Mathematics and Science Study (TIMSS) show that mathematics textbooks are rarely used or read from in Slovenian classrooms. From this we assume that mathematics teachers in Slovenia view the textbook as the basic catalogue of expected knowledge students should acquire, both during lessons and through independent learning at home. This leads to our basic research question of whether student outcomes, as measured by the TIMSS and National Assessment, are related to textbook choice. Another problem of mathematics and science education in Slovenia is the decline in motivation to learn. Since publishers put a lot of effort into making textbooks attractive for students – and since they compete to encourage teachers, parents and students to choose their textbooks for compulsory learning material – we will explore whether more attractive textbooks actually motivate students to learn mathematics and science more, and whether mathematics and science teachers achieve greater motivation to learn among their students by choosing a specific textbook. The research question is: how does the choice of textbook relate to students' motivation for learning mathematics and science (whether they like learning these subjects), their self-confidence in both subjects, and their values associated with mathematics and science?

Data source

The study has two parts. In the first, we review reports about use of textbooks from international data provided by successive editions of the TIMSS, in which Slovenia has participated for several decades. These international large-scale assessments measure students' knowledge and teaching factors among representative samples of national populations of Grade 4 and Grade 8 students every four years. The samples for all TIMSS studies in each participating country follow a two-stage random sample design, with a sample of schools drawn as the first stage and one or more intact classes of students enrolled in a specific grade selected from each of the sampled schools as the second stage (Laroche et al., 2016). For example, in 2015, the Slovenian sample included 4,800 Grade 4 students and 4,600 Grade 8 students from 150 primary schools. The students were assessed in mathematics and science, and reported about themselves and their learning by answering questionnaires, while questionnaires were also answered by the students' school principals and teachers (257 fourth-grade class teachers, 471 mathematics teachers and 859 science teachers). The data were linked into a database for each study cycle and appropriate weightings were created to accommodate complex sampling design. Statistical analyses of the weighted student data enable us to calculate nationally representative population estimates. With the questionnaires, TIMSS studies over the years have collected a variety of comparative data about textbooks.

We combined the TIMSS data with a national database on use of textbooks in each Slovenian school from the national portal Trubar. Until 2019, this portal was intended to inform students and parents about the printed or paper textbooks that were selected as compulsory for each grade and subject at each school. Schools were obliged to enter the chosen list of teaching material into the central national database Trubar. The chosen textbooks were provided to schools nationally and were available in the school textbook library for students. In the present paper, we used data on the use of textbooks from the Trubar web portal for the 2016 school year. The data lacks validity after 2016, as schools were no longer obliged to report data.

The other part of the present study includes National Assessment (NA) data. Held in Grades 6 and 9 of primary school, NA is a low-stakes external assessment of Slovenian language, mathematics and a third school subject that is by default the first foreign language in the sixth grade or one of the selected third subjects in the ninth grade. NA is based on the curriculum and provides students, parents, teachers, schools and others with additional information about each student's knowledge in addition to the teacher's school grades.

Data on the use of textbooks from the Trubar database were linked to anonymised NA achievements in Mathematics (2016), Chemistry (2015), Biology (2017) and Physics (2016), again to explore differences in achievement due to the different textbooks used.

Method

Our report starts with a summary of TIMSS studies conducted in 2003, 2007, 2011 and 2015 on use of textbooks, as reported by either students or their teachers and principals. This gives us a good overview of the use of textbooks in Slovenia.

Next we selected variables and indices from the international database of TIMSS 2015 that reflect the use of textbooks, as well as data on student achievement, and linked them to the national database on the use of textbooks (Trubar). At the level of the individual sample unit (the student), we obtained information on which textbooks were prescribed at the student's school in 2015 for use in mathematics and science subjects, together with all of the data from the international study TIMSS 2015. This enabled the calculation of predictive models and differences in achievement in mathematics and science of eighth graders according to the use of different textbooks. For Grade 4, the Trubar portal contained much less data on the prescribed printed textbooks. This is due to the fact that for lower grades, a large selection of workbooks and digital textbooks or materials are freely available, and these were therefore most frequently prescribed as compulsory textbooks by schools. Since they were not lent by school libraries, the selections were not reported nationally. The results of the analysis for Grade 4 are therefore limited to the available data.

Similar analyses were repeated for National Assessment (NA) data from Grade 9.

The aim of studying the relationships between textbooks and knowledge was to determine whether the research could identify some effects; it was not to measure the effectiveness of individual textbooks. Therefore, the textbooks are shown under codes in order to prevent the disclosure of publishers. Correlation and regression analyses were performed using specific software developed for complex TIMSS assessment data (IDB Analyzer and SPSS) or using the program R in case of NA data.

Results

Overview of the use of textbooks in Slovenia until 2015

First assessment during the stepwise implementation of the extensive school reform of compulsory schooling, 2003

In 2003, an extensive school reform was implemented in the lower grades of most primary schools, but only in a sample of schools for Grades 6 to 9. The TIMSS 2003 study provides Slovenia with an opportunity to examine differences between the old and the reformed school system, which prolonged compulsory schooling by one year and renewed curricula and didactic approaches. In the TIMSS, principals of national representative samples of primary schools in more than 30 participating countries were asked about the lack of textbooks in their schools for students in mathematics and science classes. Slovenian principals of the vast majority (87%) of eighth-grade students confirmed that their schools do not feel limited in teaching due to a lack of textbooks; principals of 11% of the students judged that classes suffer little, and principals of only 2% of the students thought that classes suffer seriously from a lack of textbooks. The mean international percentage of students in schools where classes do not suffer from a lack of textbooks was 41%, less than half of that in Slovenia (TIMSS Almanacs 2003, 2005a, p. 183). This shows that Slovenian students had ample access to new textbooks for the reformed curricula immediately after the reform was implemented in practice. Differences in achievement between students from schools suffering from a shortage of textbooks and those not suffering from such a shortage cannot be estimated, as there are too few students in the last group, but the mean mathematics and science achievement across other countries decreases slightly with an increased shortage of textbooks for school instruction.

In Slovenia, mathematics and science teachers of all students confirmed that they use a textbook for teaching (TIMSS Almanacs 2003, 2005b, p. 151). The international average shows higher achievement when teachers use textbooks for teaching, but there are differences between countries. It was not possible to determine the differences in achievement in Slovenia, as there were not enough students with teachers who did not use textbooks for teaching. However, in some lower achieving countries, the students whose teachers did not use a textbook achieved lower scores than students who were taught by teachers using textbooks. In Belgium and England, both high-achieving countries, teaching with textbooks was associated with higher student knowledge.

In Slovenia, mathematics teachers of 68% of the students taking the TIMSS stated that the lack of textbooks was not an obstacle to teaching mathematics at all, teachers of 14% of the students regarded the lack of textbooks as a minor obstacle to teaching, teachers of another 14% of the students perceived the lack of textbooks as significantly or severely limiting teaching, and teachers of the remaining 3% of students did not use textbooks. Student achievement did not differ between the first three groups according to the teacher responses, but it did decrease in the group of students whose teachers felt severely limited in their teaching by the lack of textbooks. Across all countries, the last group of students whose teachers felt severely limited by the lack of textbooks is more than twice as large as in Slovenia and the first group is half as small, while mathematics achievement does not differ substantially between groups of students experiencing more or less limited teaching due to the lack of textbooks (TIMSS Almanacs 2003, 2005b, p. 199).

In science (TIMSS Almanacs 2003, 2005c, p. 226), teachers felt least limited by the shortage of textbooks in Hungary and the Netherlands, and most limited in the poorer countries of Ghana, Indonesia and Morocco. In Slovenia, 63% of the students had science teachers who reported that the lack of textbooks for science did not limit them in teaching. Despite the high share, teaching science was at least slightly limited for 14% of the students, quite limited for 8% and very limited for 3% due to lack of textbooks. In total, a substantial share of the participating students (a quarter of all of the Slovenian students) experienced limited teaching of science due to a lack of textbooks.

Reformed compulsory school before the first systematic review, 2007

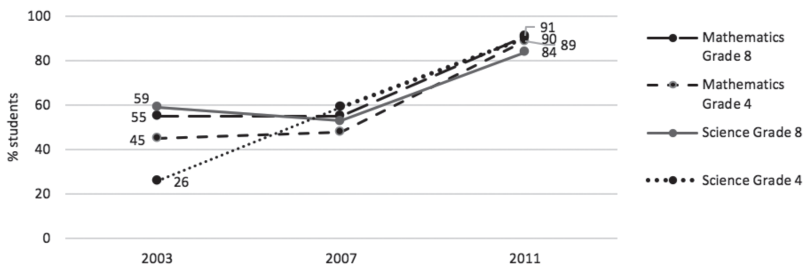
The TIMSS study in 2007 was the first reflection of the fully implemented newly reformed school system in Slovenia. The principals of the schools at which 81% of all of the eighth graders participating in the study were enrolled confirmed that they did not have any problems with a shortage of teaching materials, which was an extremely high proportion compared to other countries. Internationally, the average share of students in schools without a shortage of teaching materials was 50% (TIMSS Almanacs 2007, 2009a, p. 79). According to these data, the availability of textbooks increased during the full implementation of nine-year reformed primary schooling in Slovenia.

Science teachers reported how often they asked students to read about the taught content from a textbook (TIMSS Almanacs 2007, 2009b). The share of students asked by teachers to read from the textbook each lesson was less than 10% in three countries: Norway, Sweden and Slovenia (4.2%). Judging by the science achievements in Slovenia, more frequent reading from a textbook did not increase student achievement. On the contrary, slightly higher achievement was shown by students from whom teachers never required reading from a textbook than by students who read from a textbook during at least half of the lessons. This result can be explained by the national practice of teaching science and the nature of textbooks for these subjects, which are intended more to explain the content to students when they learn individually after school than as a working source for following the lessons in school. Unlike mathematics textbooks, science students mostly do not need to practise solving problems from textbooks, neither in school nor after school.

Textbooks in schools ten years after the reform, 2011

From the trends in the proportions of students whose teachers confirmed that they use a textbook as their primary source of instruction and not as a secondary source (Figure 1), it can be seen that teaching in Slovenia has increasingly relied on textbooks since the reform of compulsory school. In 2003, teachers of only about half of the students based teaching primarily on textbooks, whereas in 2011, in all of the observed subjects, teachers of 90% of the students or more viewed the textbook as the primary source of their teaching.

Figure 1
Trends in the share of students whose teachers use a textbook as their primary source of teaching, 2003–2011



Mathematics teachers of 90% of the participating students in Slovenia stated that a textbook is their primary source of preparation and implementation of lessons (TIMSS Almanacs 2011, Mathematics teacher background data with mathematics achievement). Comparable international data on the use of textbooks and the measured knowledge of students do not show that teaching strictly linked to the prescribed textbook is more effective. There was no difference between the achievements of students whose teachers use a textbook as a basic or secondary source in Slovenia. The results suggest that the way the textbook is used does not in itself lead to large differences in students' knowledge.

In science, a textbook was used as a secondary source by 15% of students (TIMSS Almanacs 2011, Science teacher background data with science achievement). Students' knowledge was higher when teachers used textbooks as a secondary source of teaching than when students were taught strictly from a textbook in several high-achieving countries, including Taiwan, Finland, Hong Kong and Hungary, and a slight trend is also observed in Slovenia.

The frequency of reading from textbooks and other materials in science lessons has increased in Slovenia since 2007. The share of students whose teachers require reading every lesson or every other lesson has more than doubled, and the share of students who have never been asked to read something from a textbook in class has halved. Contrary to expectations, the gap in knowledge between students who never or at least sometimes read in class has widened. Those who never read showed a trend towards higher achievements. The explanation could be that these students experience active learning in school and learn from textbooks individually at home to review the content and deepen their understanding.

After implementation of the reviewed changes in curricula, 2015

In the TIMSS conducted in 2015, Slovenian students demonstrated increasing knowledge in all areas, especially in science subjects. At the same time, the international comparison of curricula confirmed that a larger amount of content was expected to be taught to Slovenian students in science lessons than in most other countries. For the first time, mathematics achievement in the eighth grade, which had previously been largely stagnant, increased significantly. The problematic part of the research findings refers to the trend of negative student attitudes towards learning, with a further decline in liking learning mathematics and science to less than a quarter of all students, and to internationally comparable extremely low self-confidence of students, their valuing of knowledge, their sense of belonging to school, and their perception of engaging teaching of their teachers (Japelj Pavešić & Svetlik, 2016).

General relationships between textbooks, teaching and student knowledge

The relationships between the frequency of reading from textbooks, achievement and science teaching strategies were first checked by correlation analyses. The results showed insignificant or weak correlations between the frequency of reading from textbooks and other material in lessons and student achievement, the frequency of the teacher's linking new content to everyday life, the teacher asking students to solve challenging problems, the teacher encouraging discussion among students, or the teacher connecting subject knowledge with other areas, as well as the school's commitment to the academic success of students and the teachers' commitment to teaching by experimenting (correlation coefficients are less than .18) (Table 1)

Table 1

Correlation between reading textbooks in lessons and characteristics of teaching science, Grade 8

(s.e.) → Correlation coefficients ↓		A	B	C	D	E	F	G	H	I
How often the teacher\ Relates the lesson to students' daily lives	**A**		(.04)	(.04)	(.04)	(.04)	(.05)	(.05)	(.04)	(.02)
How often the teacher\ Asks students to explain their answers	**B**	.36		(.04)	(.04)	(.04)	(.05)	(.04)	(.04)	(.01)
How often the teacher\ Asks students to complete challenging exercises that require them to go beyond the instruction	**C**	.20	.35		(.04)	(.04)	(.04)	(.04)	(.04)	(.02)
How often the teacher\ Links new content to students' prior knowledge	**D**	.37	.41	.18		(.04)	(.06)	(.04)	(.04)	(.01)
How often the teacher\ Encourages students to express their ideas in class	**E**	.33	.41	.28	.36		(.04)	(.04)	(.04)	(.02)
School emphasis on academic success	**F**	-.16	-.13	-.24	-.09	-.16		(.04)	(.05)	(.03)
How often the teacher\ Asks students to read their textbooks or other resources	**G**	.06	.15	.18	.15	.11	-.10		(.05)	(.02)
Teachers emphasis on science investigation	**H**	-.24	-.26	-.29	-.22	-.29	.10	-.14		(.01)
Science achievement	**I**	.01	-.01	-.05	.00	.01	.07	-.02	.01	

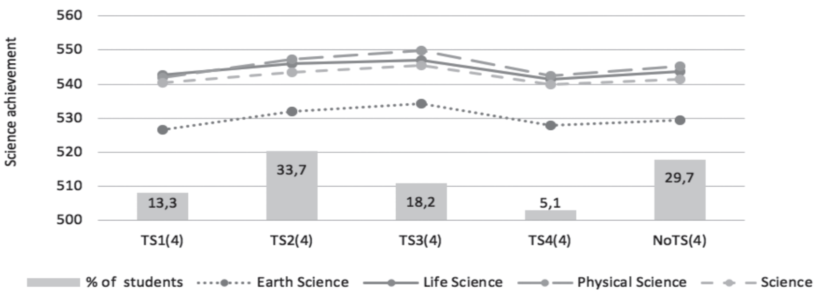
Differences in achievement in Grade 4 according to textbooks

An analysis of data from the Trubar database for Grade 4 in 2015 reveals that there are no data on prescribed printed textbooks for science lessons for about 30% of students, and that there are no data on prescribed printed textbooks for mathematics for a very large share, 81% of students. This could be attributed to the large number of digital or electronic textbooks already in use in 2015, as well as to the fact that no records of the prevalence or prescribed use among students were kept. According to school policy, all students should have prescribed textbooks. Therefore, students without data in the Trubar database most likely used online textbooks. Among the reported printed textbooks there were four for the subject science and seven for mathematics. The results of the analyses of textbook use in Grade 4 should be read with caution, taking into account the fact that printed textbooks were used by a decreasing share of students.

A third of students were learning science from one textbook and a fifth from another, while about a third were learning from several different sources, each of which was much less widespread. Two facts were revealed by the analysis: (a) differences in the knowledge of students learning from different textbooks were not significant, and (b) the knowledge of students without a printed textbook did not differ from the knowledge of students who had a printed textbook. The knowledge of different content sets of science was similar among students with different textbooks. It is therefore not possible to identify any textbook as more effective in learning either individual science content or science content in general (Figure 2).

Figure 2

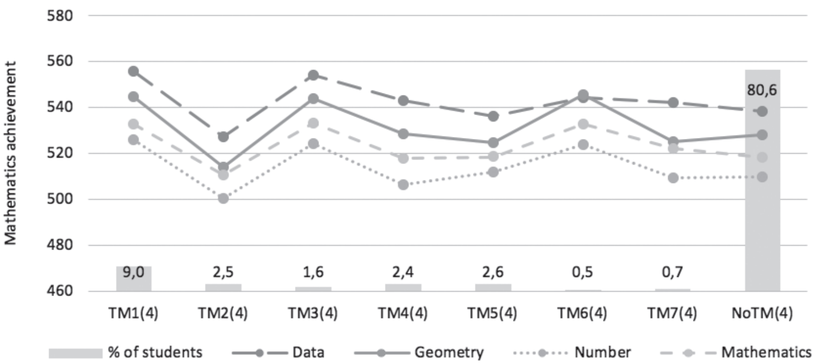
Percentages of students and science achievement by prescribed textbooks, TIMSS 2015 Grade 4



In mathematics, the results are limited to one-fifth of the entire student population. Of these students, half had one particular textbook and the other half one of six other different textbooks. The differences in knowledge are greater than in science. They show a favourable picture of higher knowledge among students who use the most frequently used textbook, but also a significant decrease in knowledge among students with another particular textbook TM2(4). In the area of numbers, the difference between students having one of other textbooks is 22 points (on a scale with an average of 500 and a deviation of 100 points). Following the recommendations for assessing differences by calculating the effect size d (Cohen, 1992), which is the quotient between the difference and the total standard error, we find that the above difference is large (d = 3.3; d > .8 is considered a large effect). However, this does not present a problem, because the textbook that proved to be less effective was only used by a very small group of students (2.5%). There is a similarly large difference between the achievement of students who learn from the predominant printed textbook and the achievement of 81% of students who learn from other learning sources. The latter group scored 14 points lower and the effect size was again defined as high (d = 3.4) (Figure 3).

Figure 3

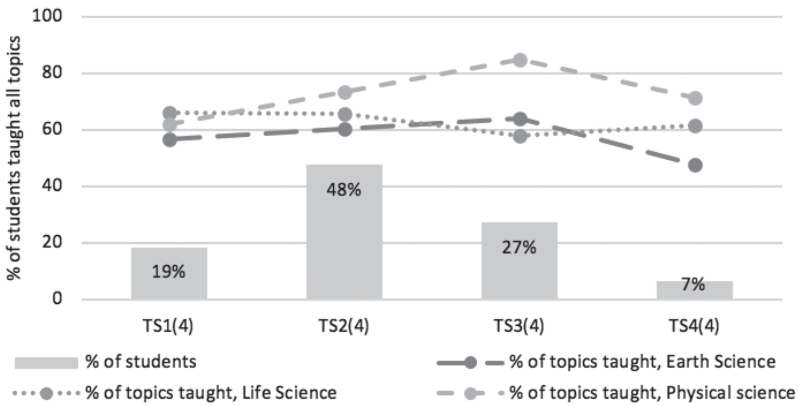
Percentages of students and mathematics achievement by prescribed textbooks, Grade 4



Due to the large differences in knowledge, we checked the number of topics taught based on teachers' reports in the TIMSS. The teachers indicated whether the students had already been taught the list of topics covered by the international TIMSS assessment, which is internationally agreed to reflect curricula from the participating countries. The reports partly limit our interpretation, as some topics were not covered by the Slovenian curriculum, while there were some topics that were not included in the international tests but were taught to students in Slovenia. Nevertheless, from the share of students whose teachers reported which topics had been taught to these students, it is possible to estimate the total level of taught content. In science (Figure 4), the share of students who had been taught all of the topics ranges from about half to 85% and varies considerably between textbooks. With the textbook that also showed the highest level of student knowledge (TS3(4)), most students learned all of the content, whereas the lowest number of students had the opportunity to learn all of the listed content with the textbook that is associated with the lowest achievement. It is problematic that the predominant science textbook does not demonstrate the maximum amount of material learned.

Figure 4

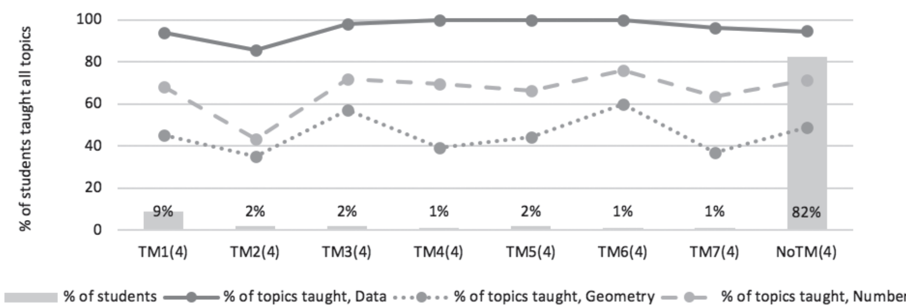
Proportions of students taught science content from the TIMSS test by prescribed textbooks, TIMSS 2015 Grade 4



The pattern is different in mathematics (Figure 5). All of the textbooks except one show comparable shares of students taught all mathematical topics included in the TIMSS assessments.

Figure 5

Proportions of students taught mathematics content from the TIMSS test by the prescribed textbook, TIMSS 2015 Grade 4



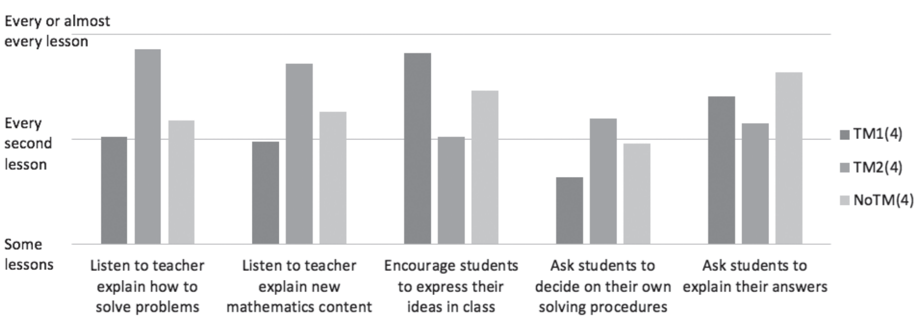
A side result of this comparison is the large difference between the individual content areas of mathematics. Only between 30% and 60% of the students had learned the geometry content expected by the TIMSS test, whereas almost all of the students had learned the content from the area of data. This large difference was not expected, as the same national curricula is requested to be taught to all students. The lower coverage could be explained by the fact that in the TIMSS tests, the geometry items included calculations and measurements with units as well as an understanding of the concept of area, which are taught in Slovenia later, from the fifth grade onwards. Similarly, the area of numbers also covers some concepts in the fourth grade that are not taught to Slovenian students, such as calculation with simple fractions and decimal numbers. Nevertheless, for mathematics, we can conclude that the number of topics taught to students using sources other than printed textbooks does not deviate from the number taught with a printed textbook.

Teaching fourth-grade students using different textbooks

From the comparison of the frequency of certain activities required by the teacher in the classrooms of fourth-grade students in Slovenia in 2015, we find significant differences in teaching with different textbooks. Teachers' data on their teaching of students were linked with information on prescribed textbooks at the student level and averaged for Slovenia. In mathematics, students more often state or explain their opinions if they have the first textbook, whereas teachers talk more often if they use the second textbook (Figure 6).

Figure 6

Frequency of activities in mathematics lessons by prescribed textbooks, TIMSS 2015 Grade 4



In science, the differences are smaller but still noticeable, especially in the frequency of activities associated with experimentation (Table 2). One science textbook (TS3(4)) stands out by the highest frequency of most teaching activities, especially planning experiments and other experimental activities, but also for the more frequent presentation of new content by the teacher. A second textbook (TS4(4)) stands out for the lowest frequency of all activities, especially the teacher's explanations, while teachers also reported that the students were least likely to have to memorise facts and procedures when they used this textbook.

Table 2

Frequency of approaches to teaching science by prescribed textbooks, Grade 4

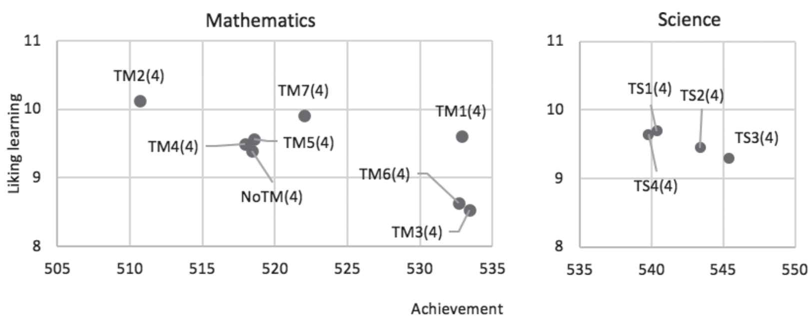
Frequency of activities (some lessons or half of the lessons)	TS1(4)	TS2(4)	TS3(4)	TS4(4)
Listen to me explain new science content	some	some	some	some
Have students memorise facts and principles	some	some	some	some
Read their textbooks or other resource materials	some	some	some	half
Observe natural phenomena such as the weather and describe what they see	some	some	some	some
Watch me demonstrate an experiment or investigation	some	half	half	half
Use evidence from experiments or investigations to support conclusions	half	half	half	half
Present data from experiments or investigations	half	half	half	half
Conduct experiments or investigations	half	half	some	half
Interpret data from experiments or investigations	half	half	half	half
Design or plan experiments or investigations	half	half	half	half

Differences in fourth graders' attitudes towards learning according to the textbook

As part of the TIMSS survey, the students were also asked to rate their attitudes towards learning mathematics and science, reporting their level of agreement with statements about learning each subject. Four scales of attitudes were modelled from the students' responses for each subject: liking learning mathematics and science, self-confidence, valuing mathematics and science, and perception of engaging teaching by their subject teachers. The students were assigned values on each scale that had international means of 10 points and standard deviations of 2 points. Higher values on the scale describe a more positive attitude.

An analysis of the comparison of mean scores on the attitudinal scales by the prescribed textbooks follows the opposite pattern as demonstrated knowledge (Figure 7). The use of the mathematics textbook TM2(4) shows the highest student motivation to learn, and TM3(4) and TM6(4) the lowest. We already know that the textbook TM2(4) is associated with the lowest level of knowledge among students, and textbooks TM3(4) and TM6(4) with the highest. We can thus confirm the presence of an internationally known paradox between achievement and attitudes (the attitudes-achievement paradox, Min et al., 2016), which shows higher achievement in groups of students with lower attitudes and, vice versa, lower achievement in groups of students with higher attitudes. This is significant in the across country comparisons and has been measured and reported several times for the international large scale assessments TIMSS and PISA. It is strongly present in the countries of East Asia, where they try to explain it to some extent with a culture of modesty in reporting personal attitudes. It is also clearly seen in the case of the Slovenian comparison of student attitudes towards the use of a mathematical textbook in the fourth grade. Roughly speaking, students with higher achievement develop less positive attitudes towards learning mathematics and science than students who are less successful in learning.

Figure 7
Liking learning vs. achievement by prescribed textbooks, TIMSS 2015 Grade 4



The paradox of the conflicting results in achievement and motivation for learning is repeated in fourth-grade science. The trends fall from the highest motivation to learn when using the textbook TS4(4), where achievement is lowest, to the lowest motivation for learning with the textbook TS3(4), which shows the highest achievement (Figure 7). Since the graphs for both subjects are drawn in the same axis ratio, we also see a direct comparison between the large differences in attitudes towards learning mathematics and the small differences in attitudes towards learning science between the textbooks for both subjects. Note, however, that direct comparisons of absolute differences on the scales of both attitudes and achievement are not possible, as each scale is modelled separately.

Differences in the knowledge of Grade 8 students using different textbooks

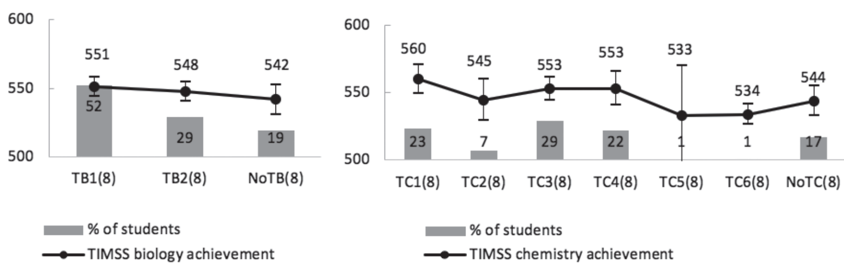
The TIMSS collected data for mathematics and natural science, which in the Slovenian education system includes biology, physics, chemistry and geography. Below we limit ourselves to mathematics and the first three typical natural science subjects.

In the first step, we observed the shares of students who studied according to a particular textbook and their achievements in the TIMSS for biology, chemistry, physics and mathematics. The overall science achievement should not be used, as textbooks are created for each school subject separately. In mathematics, we were able to examine overall knowledge and knowledge in individual fields of mathematics, which was reflected in the scales of individual types of achievements. The differences between achievements seem large, but are not statistically significant due to the large differences in the proportion of students using textbooks. The presented figures (8 and 9) include self-confidence intervals that single out some achievement differences as significant.

The differences in achievement in biology are not statistically significant among students who were using different textbooks (Figure 8). Nor is there any difference between students who have or do not have a prescribed printed textbook. In chemistry, where more textbooks (six) were used, three of which were used by most students, the achievements differed between students more than in biology. There are no statistically significant differences between the three most common textbooks, but there is a significant difference in achievement between TC1(8) and TC6(8) in favour of TC1(8), which also shows the highest average achievement of students.

Figure 8

Achievements in biology and chemistry by different textbooks, TIMSS 2015 Grade 8



Even in physics (Figure 9), where two thirds of the students used one textbook, two other textbooks were used by 13% and 1% of the students, respectively, and almost a fifth of the students do not use any printed textbook, the average achievement does not differ significantly among all of these four groups of students. In mathematics, two thirds of all of the students used one of the textbooks. Students achievement does not differ significantly because the groups using the other five textbooks are small and the confidence intervals wide (Table 3). Nevertheless, there is a trend for the achievements in each mathematics topic to follow the pattern of differences in the overall achievement among the students who learn with the aid of different textbooks. The data do not show that the students achieved a noticeably higher or lower level of knowledge in individual mathematics topics with one of the textbooks. It is, however, interesting to note that the achievement of the fifth of the students without a prescribed printed textbook was not the lowest.

Figure 9

Achievements in physics and mathematics by different textbooks, TIMSS 2015 Grade 8

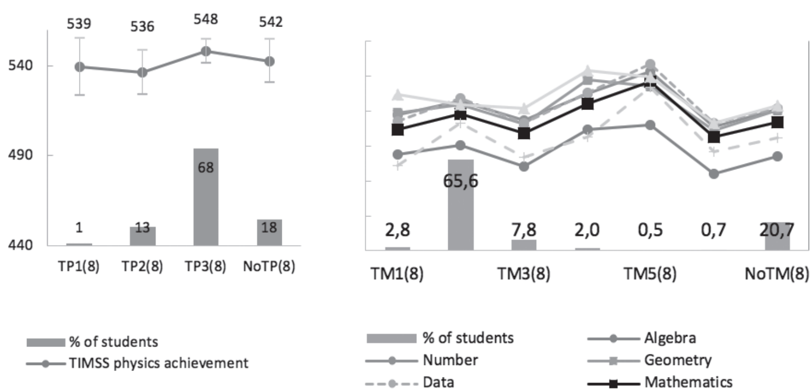


Table 3

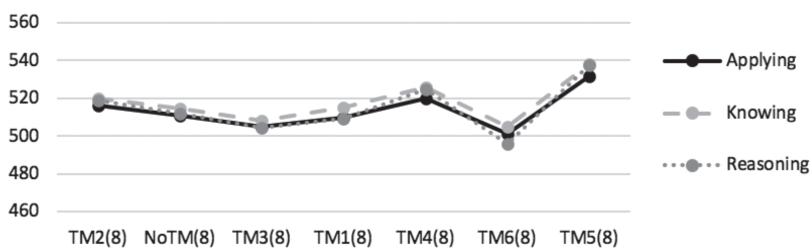
Proportions of students and mean mathematics achievement by content areas in 2015 by different textbooks, Grade 8

Prescribed textbooks	% of students	Mathematics (st. n.)		Algebra (st. n.)		Number (st. n.)		Geometry (st. n.)		Data (st. n.)	
TM1(8)	2.8	509.3	(10.4)	495.0	(7.9)	517.9	(8.6)	519.1	(9.5)	514.0	(12.4)
TM2(8)	65.6	518.3	(2.6)	500.7	(2.9)	525.8	(3.0)	523.6	(3.3)	527.2	(3.2)
TM3(8)	7.8	507.4	(7.1)	488.2	(7.6)	514.5	(8.6)	512.5	(7.4)	513.0	(9.7)
TM4(8)	2.0	524.2	(9.7)	509.3	(9.8)	530.0	(11.8)	537.8	(11.7)	530.3	(8.9)
TM5(8)	.5	536.8	(1.4)	512.2	(4.7)	542.7	(4.2)	534.1	(4.0)	546.9	(1.3)
TM6(8)	.7	504.9	(4.2)	484.3	(3.5)	511.0	(4.3)	508.9	(6.5)	513.1	(6.4)
NoTM(8)	20.7	513.7	(4.7)	494.0	(4.6)	521.2	(5.1)	520.4	(5.0)	521.7	(5.4)

We were able to compare mathematical achievement according to the cognitive aspect. For science, achievement on cognitive scales is only available for the overall science score and is therefore not useful for our analysis. Figure 10 shows the average achievement on three cognitive scales – knowledge of facts, use of knowledge and reasoning – according to the textbook used in the classroom. The average achievement in these scales again follows a common pattern of differences between textbooks. Since most of the differences are not significant, we are unable to confirm that the textbook TM2(8), which is used by two thirds of the students, is a more or less successful tool for acquiring mathematical knowledge than other textbooks.

Figure 10

Achieved scores on cognitive scales of mathematics by textbooks, TIMSS 2015 Grade 8

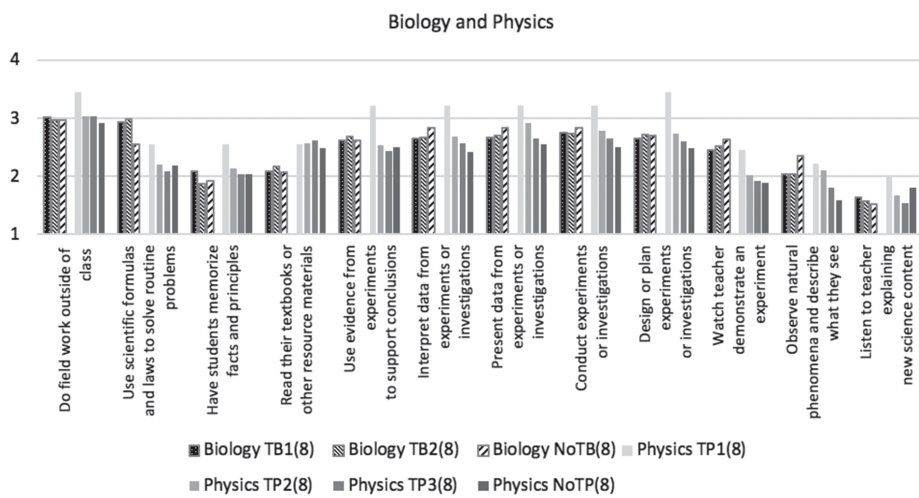


Differences in teaching students with different textbooks in Grade 8

Teaching with different textbooks was compared with the teacher's report on the frequency of certain activities with students in classes on a scale of 1 to 4 (4 = every or almost every lesson, 3 = about half of the lessons, 2 = at some lessons, or 1 = never). There are noticeable differences between lessons in all science subjects (Figure 11). In biology, the classes differ most for students who have a textbook and those who do not. In physics, these activities are least often carried out by students using the textbook TP1(8), and most often by those who use TP3(8) or are without a textbook. In chemistry, one textbook stands out, as students are very often (every lesson) asked to read from it and remember the facts. When observing data, we should remember that students may have their own textbooks or online textbooks that are not reported on the school level.

Figure 11

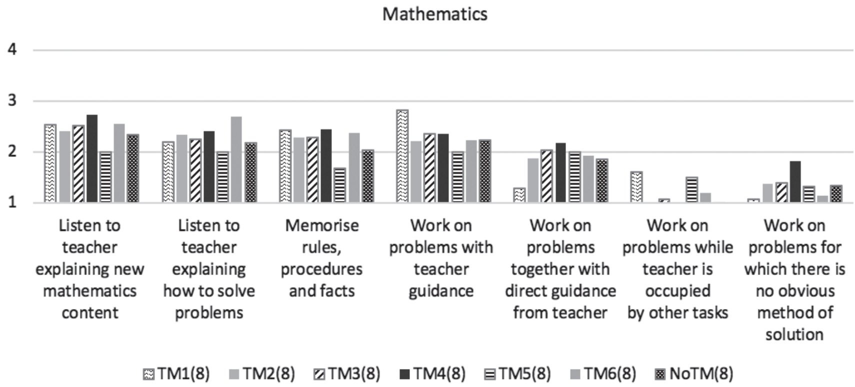
Frequencies of teaching activities in biology and physics by prescribed textbooks, Grade 8



Note. 1 = never, 2 = some lessons, 3 = every second lesson, 4 = every or almost every lesson.

Figure 12

Frequencies of teaching activities in mathematics by prescribed textbooks, Grade 8



Note. 1 = never, 2 = some lessons, 3 = every second lesson, 4 = every or almost every lesson.

In mathematics, the pattern of differences is different. Textbooks TM1(8) and TM5(8) deviate the most, while the textbook TM2(8), which is used by the vast majority of the students, does not stand out in any activity (Figure 12). Although the shares of students using other textbooks are small, the analysis clearly shows that the teaching methods vary according to the textbook used.

Differences in attitudes of the students towards learning and textbooks used

The students' views are measured by a scale with an international average of 10 points and higher values for more positive views. An analysis of differences in the students' attitudes towards mathematics and specific science subjects broken down by the textbooks used shows that the self-confidence of students has the highest values across all textbooks (Figure 13). Other attitudes do not follow the same pattern and differences are not always statistically significant. Across most of the textbooks, self-confidence is followed by the students' perception of the engaging teaching of mathematics, which also measures the teacher's direct attitude towards the student. Surprisingly, the textbook TM1(8) achieved almost the same high score in the students' perception of the teacher's commitment to teaching as the self-confidence of the students, which is much higher than the average of the students using other textbooks. Liking learning of mathematics shows a pronounced decline in the textbook TM6(8).

The data about students who do not have a printed textbook are also interesting. The values of all of the students' attitudes, except for somewhat higher self-confidence, are very similar and reach the mean values of the students who have textbooks. Therefore, the data do not show that the use of specific textbooks increases motivating factors for students (see t-test statistics in the Appendix, Table 1), but only that the textbooks differ in their influence on attitudes.

Figure 13

TIMSS attitudes toward mathematics by prescribed textbooks, TIMSS 2015 Grade 8

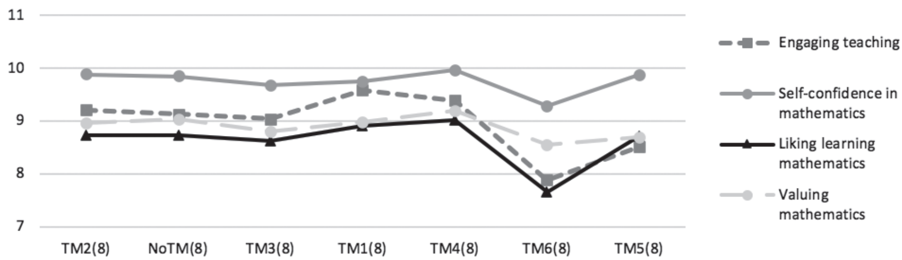
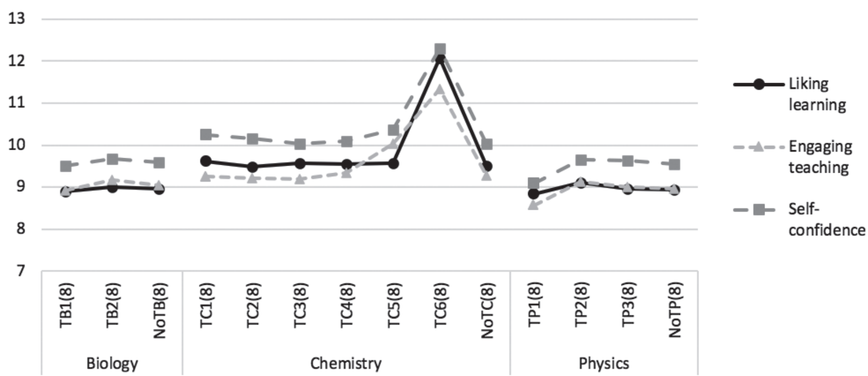


Figure 14

TIMSS attitudes toward science subjects by prescribed textbooks, TIMSS 2015 Grade 8



The differences in liking learning science subjects and self-confidence in science subjects between students with different textbooks are very small, with the exception of chemistry. Students prefer to learn chemistry and are confident in their chemical knowledge when using the textbook TC6(8) (see t-test statistics in the Appendix, Tables 2–4).

Use of different textbooks and National Assessment

National Assessment (NA) items are derived from the curriculum and reflect the requirements and expectations of the curriculum even more accurately than TIMSS. Therefore, differences in the performance of students using different textbooks could point even more directly to the unequal coverage of the curriculum in the textbook used by the students.

As mentioned above, the data on the use of textbooks were linked to anonymised achievements of NA in Mathematics (2016), Chemistry (2015), Biology (2017) and Physics (2016). For the most widely used textbooks, we looked at achievement in each item and checked whether there were differences in achievements between groups. Table 4 shows an example of items for physics with correct proportions for different textbooks.

Table 4
National Assessment results by four textbooks on selected items in Physics 2016 (Grade 9)

	Proportion correct (PC)						Content information		
Item	Total	TP1(9)	TP2(9)	TP3(9)	TP4(9)	Max diff.	Content	Taxonomy*	Item type**
01	.53	.58	.40	.54	.45	.18	Light	UA	MC
02	.88	.90	.82	.89	.82	.08	Force	KR	MC
03	.59	.59	.63	.59	.56	.06	Force	UA	MC
04	.30	.33	.19	.29	.37	.18	Force	UA	SA
05	.67	.73	.61	.67	.67	.12	Pressure, density, buoyancy	KR	SA
06	.14	.14	.11	.14	.15	.05	Pressure, density, buoyancy	UA	SA
07	.46	.47	.44	.47	.42	.05	Work, energy, heat	UA	MC
08	.60	.63	.52	.60	.61	.11	Motion	UA	MC
09	.54	.59	.46	.54	.42	.17	Motion	UA	MC
10	.46	.39	.39	.47	.53	.14	Space	UA	MC
11	.95	.96	.93	.95	.92	.04	Space	KR	MC

	Proportion correct (PC)						Content information		
Item	Total	TP1(9)	TP2(9)	TP3(9)	TP4(9)	Max diff.	Content	Taxonomy*	Item type**
12.1	.50	.52	.50	.51	.46	.07	Sound and waves	KR	SA
12.2	.52	.54	.48	.52	.57	.10	Sound and waves	KR	SA
13.a	.37	.46	.34	.37	.28	.18	Light	KR	MC
13.b	.28	.50	.14	.27	.22	.36	Light	KR	MC
14.1	.32	.34	.20	.32	.33	.14	Force	UA	SA
14.2	.40	.42	.27	.40	.36	.15	Force	UA	SA
15.a	.70	.73	.65	.71	.67	.08	Pressure, density, buoyancy	PS	MC
15.b	.41	.48	.46	.41	.32	.16	Pressure, density, buoyancy	PS	MC

Note. *UA – understanding and application; KR – knowledge and recognition; PS – problem solving.**MC – multiple choice; SA – short answer

The results for all of the school subjects are presented in Figures 15–20 by distributions of national proportions of correct answers for individual items (each vertical line), according to the textbooks prescribed to students by schools.

Figure 15

Proportions of correct answers in NA in Mathematics (2016) for students using three different textbooks, Grade 9

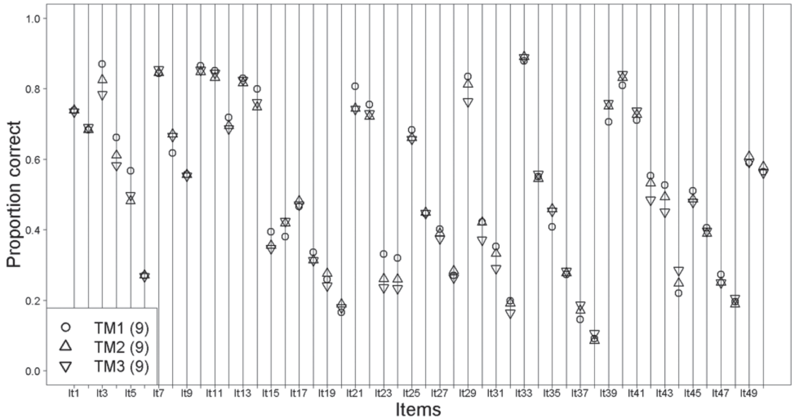


Figure 16

Proportions of correct answers in NA in Biology (2017) for students using three different textbooks, Grade 9

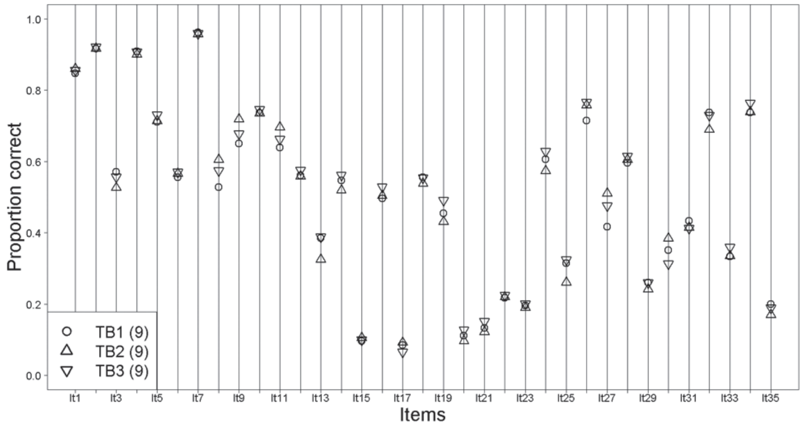


Figure 17

Proportions of correct answers in NA in Chemistry (2015) for students using three different textbooks, Grade 9

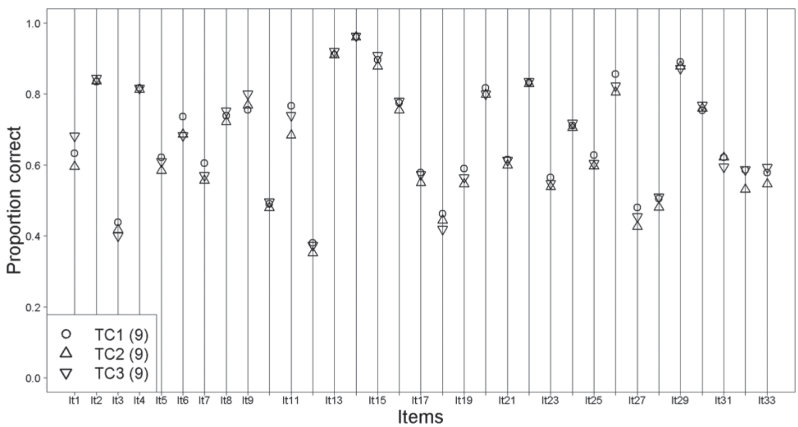
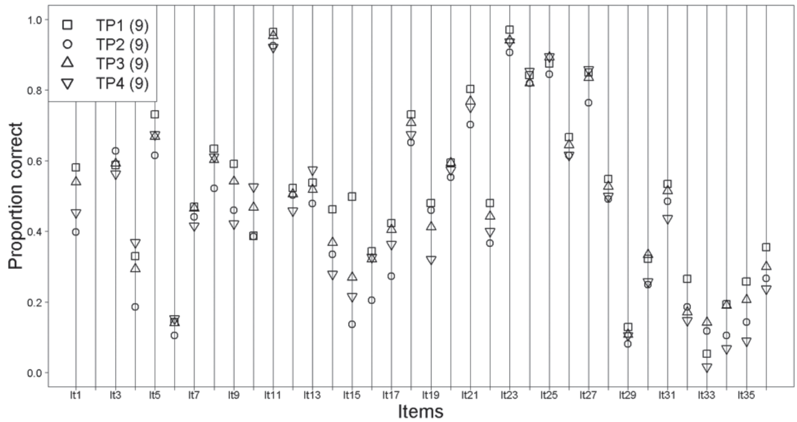


Figure 18

Proportions of correct answers in NA in Physics (2016) for students using four different textbooks, Grade 9



An analysis of the proportion of correct answers in each subject shows that the differences between groups are usually not large and that more difficult items appear more difficult in all groups. The differences between groups are significantly smaller than the differences between the difficulties of different items. For chemistry and biology, the differences between the textbooks are negligible in most cases: the group labels on the graphs are virtually overlapping. In mathematics, the differences are small, but systematic. The data for physics show larger variations from item to item, but it is more difficult to extract a systematic pattern. It should be noted that there could be several reasons for the differences between groups. Since they were not selected at random, it cannot be excluded that the groups are different in their initial abilities and these differences have been reflected in the observed deviations. On the other hand, some of the differences may also be due to actual differences in the treatment of specific content in textbooks and, consequently, to the impact of the textbook in achieving the knowledge associated with particular content. For example, there is a very large difference in Item IT15 (source: NEC, 2016) in physics (Figure 16).

Example IT15:
How is an image displayed in an eye and how is it displayed in a camera obscura? Select the correct answer.
A If it is flipped in the eye, it is flipped in the camera obscura.
B If it is not flipped in the eye, it is flipped in the camera obscura.
C If it is flipped in the eye, it is not flipped in the camera obscura.
D If it is not flipped in the eye , it is not flipped in the camera obscura.

IT15 was a multiple choice item, so we can examine the frequency of the correct answers (answer A) as well as the frequencies of the three incorrect answers (B, C and D) (Table 5).

Table 5

Proportion correct in Item IT15 by prescribed textbooks

Textbook	A	B	C	D	NR
TP1 (9)	49.8%	19.7%	28.3%	1.8%	0.4%
TP2 (9)	13.7%	28.6%	39.8%	16.8%	1.2%
TP3 (9)	26.9%	29.8%	37.7%	4.9%	0.6%
TP4 (9)	21.6%	27.9%	41.1%	9.5%	0.0%

Note. Adapted from NA Physics, 2016.

The task was difficult for the students and the correct answer was not even the most common. We see quite a lot of variation from the response patterns. For the first textbook TP1, the proportion of correct student answers is much higher than for the other textbooks and the correct answer is the most common answer. TP3 and TP4 have similar patterns of answers. Special attention is given to TP2, where the students were more likely to choose answer D, which was a very rare choice in the other textbooks. Although the factors influencing the specific response of each group cannot be determined without substantive analysis of the textbooks, the differences are clear and suggest that, at least in the case of Item IT15, such analyses could be successfully implemented and conclusions reached.

Discussion

Comparisons between groups of students who learn mathematics and science with the aid of different textbooks show that in the fourth and eighth grades, textbooks are an essential element of teaching that shapes the teacher's teaching and knowledge as well as the attitudes towards the subject and the knowledge of the students. Although the comparison, especially for Grade 4, was limited by incomplete data and, particularly in mathematics, restricted to the small part of the population that uses a printed textbook for learning, the results give rise to two dilemmas.

The use of different textbooks can lead to differences in student outcomes. Teachers should be aware of the relationships between achievement, attitudes and a specific textbook in order to select the best textbook for their students. With this in mind, we cannot overlook the hypothesis that influence is actually bi-directional. A particular textbook can lead to greater achievement among students, but it is also possible that highly skilled teachers choose a particular textbook more often than teachers who are less experienced in teaching the subject. This is also reflected in the comparison of teaching, where teaching differs in the frequency of learning opportunities between textbooks. Once again, the relationship is not causal, but most probably mutual. Teachers of students who demonstrate greater autonomy in learning mathematics are more likely to choose a textbook that supports and facilitates this method. If a particular textbook does not allow so much independent learning or has a poorer presentation of content, the teacher must devote more time to his/her interpretation of the content and less time to the active participation of students.

Different textbooks encourage a positive attitude towards learning differently. This information is essential for the Slovenian education system, where there is a trend of growing negative attitudes towards knowledge, learning and school. For teachers, it is essential that this information is not taken out of context, thus neglecting the connection between textbooks and student achievement, as it has been demonstrated that textbooks with higher enjoyment result in lower mathematics achievement among the same students. Teachers who choose a textbook that motivates students to learn best and is the most popular among students could unwittingly choose a textbook that does not enable the highest knowledge to be achieved. This prompts publishers and authors of textbooks to try to design effective textbooks that facilitate high knowledge acquisition together with more joy in learning. Until this is achieved, the dilemma for teachers remains, as some textbooks will in themselves inspire more joy, while others will probably lead to higher knowledge. The most recent data for 2015 suggests that greater access to online textbooks and materials leads to generally higher knowledge among learners. In 2015, this was true for science subjects, but less true for mathematics.

If we combine the findings on textbooks with information about teaching in Slovenia, we can see that teaching is increasingly based on the selected textbook, especially in the case of mathematics in the upper grades. This in itself would not be problematic if textbooks reliably presented all of the curricular content and objectives. In 2011, there was a significant update of the curriculum in Slovenia, at least for mathematics and science subjects. New editions of textbooks were supposed to replace old copies in the school library rental system. However, such changes take time and, at least in 2015, the available textbooks for students were still not all fully aligned with the new changes. Given that lessons are primarily based on textbooks, this means that students with the earlier editions of textbooks were taught according to a slightly outdated curriculum. In fact, lower achievement in topics introduced in the eighth grade during the 2011 curriculum revision was observed in 2015 among Grade 8 students, especially high achievers, based on the example of an extremely poor knowledge of algebraic expressions in mathematics. These topics received a great deal of attention in the revision process and were included in the revised curriculum. However, a workshop with teachers participating in the project KAUČ (Quality of Slovenian Textbooks) revealed that the content was not taught in schools, mainly because it was not presented in the editions of textbooks mostly still used in schools even in 2018. The new editions of textbooks did not replace the old ones also due to the economic crisis and a lack of funding from 2012 on. As a consequence of these problems on the systemic level, eighth-grade students missed the opportunity to learn the concept of formula, which is much needed for their age. According to the TIMSS in 2015 (TIMSS Almanacs, 2017), only 14% of Slovenian students were able to substitute variables with numbers and calculate the value of a formula, compared with 20% in some European countries (Malta, Lithuania), 40% in Russia and more than 60% in Asian countries.

National Assessment results provide an independent additional source of information. Comparing results between groups of students using different textbooks on specific items in most cases showed small differences. This is an indication that different textbooks performed similarly on most items. However, not all of the differences were small. We presented the example of an item in physics and the results suggest that it is possible to search for meaningful explanations in the content of the textbooks.

Conclusion

Since the last extensive reform of compulsory school, we have experienced the evolution of textbooks and their use. Data from international studies tell us that teaching is increasingly based on textbooks, as more and more teachers view textbooks as the primary source of their teaching. The relatively small differences in student achievement suggest that we cannot determine any of the textbooks used in this research as being more effective in achieving a higher level of knowledge. However, we did find differences between textbooks. The proportions of students whose teachers have chosen certain textbooks for their instruction vary between subjects. In chemistry, for example, three textbooks are roughly equally popular, while a single textbook is most popular in mathematics. Further research could be undertaken to better understand the relationships between textbooks, teaching factors and students' attitudes, as we observed statistically significant differences in the attitudes of students using different textbooks towards learning and knowledge. This information can be important in the process of the teacher's assessment and selection of a textbook for his/her students. Further qualitative analyses of each textbook would be necessary in order to determine which characteristics could lead to an improvement in students' attitudes. In Slovenia, where we have struggled for many years with negative attitudes and resistance to learning, the message of the present study that there are differences in the liking of learning across different textbooks is good news and should be explored further.

When analysing the link between achievement and textbooks we were limited by the available data. International research is gathering less and less information on the precise use of printed textbooks, as their use has proved to be a universal condition for teaching and, on the other hand, the increasingly diverse use of different sources for teaching and learning cannot highlight the most effective approach at the international level. The textbook is primarily a reflection of the national curriculum and the orientation of each education system towards its specific objectives, which vary greatly between countries; therefore, the use of textbooks is not always comparable between countries. Data spanning more cohorts would also help to confirm our observations and generalise our findings. Despite the limitations imposed by the data, we found that teaching with different textbooks determines the way students are taught, students' attitudes and achievement in different ways. Research into the impact of textbooks could benefit from more accurate data, with additions on the use of online textbooks, as well as qualitative reports from teachers regarding which textbooks are actually used in lessons with students. We expect a greater overview of the use of textbooks from an improved record of the use of materials resulting from the renovation of the Trubar portal.

Acknowledgement

The article was produced as part of the project "Za kakovost slovenskih učbenikov" (For the Quality of Slovenian Textbooks, https://kauc.splet.arnes.si), which is co-funded by the Republic of Slovenia and the European Union from the European Social Fund.

References

Blazar, D., Heller, B., Kane, T., J., Polikoff, M., Staiger, D. O., Carrell, S., Goldhaber, D., Harris, D. N., Hitch, R., Holden, K., L., & Kurlaender, M. (2020). Curriculum reform in the common core era: Evaluating elementary math textbooks across six U.S. states. *Journal of Policy Analysis and Management*, *39*(4), 1019. https://doi.org/10.1002/pam.22257

Foy, P., & Olson, J.F. (2009a). *TIMSS 2007 user guide for the international database*: *School background data almanac by mathematics achievement– 8th grade.* TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. https://timss.bc.edu/TIMSS2007/PDF/T07_Almanacs.zip

Foy, P., & Olson, J.F. (Eds.)(2009b). *TIMSS 2007 user guide for the international database: Science teacher background data almanac by science achievement – 8th grade.* TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. https://timss.bc.edu/TIMSS2007/PDF/T07_Almanacs.zip

Foy, P., Arora, A., & Stanco, G. M. (Eds.) (2013). *TIMSS 2011 user guide for the international database: TIMSS Almanacs 2011.* TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA). https://timss.bc.edu/TIMSS2011/downloads/T11_G8_Almanacs.zip

Foy, P. (Ed.) (2017). *TIMSS 2015 user guide for the international database: TIMSS Almanacs 2015.* TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA). https://timssandpirls.bc.edu/timss2015/international-database/downloads/T15_G8_Almanacs.zip

Foy, P. (Ed.) (2017). *TIMSS 2015 user guide for the international database:TIMSS international database 2015.* TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA). https://timssandpirls.bc.edu/timss2015/international-database/downloads/T15_G8_SPSSData_pt4.zip

Hadar, L. L. (2017). Opportunities to learn: Mathematics textbooks and students' achievements. *Studies in Educational Evaluation, 55*, 153–166. https://doi.org/10.1016/j.stueduc.2017.10.002

Van den Ham, A. K., & Heinze, A. (2018). Does the textbook matter? Longitudinal effects of textbook choice on primary school students' achievement in mathematics. *Studies in Educational Evaluation*, 59, 133–140. https://doi.org/10.1016/j.stueduc.2018.07.005

Hardin, E. E., Eschman, B., Spengler, E. S., Grizzell, J. A., Taiyib Moody, A., Ross-Sheehy, S., & Kevin M. Fry. (2019). What happens when trained graduate student instructors switch to an open textbook? A controlled study of the impact on student learning outcomes, *Psychology Learning & Teaching, 18*(1), 48–64. https://doi.org/10.1177/1475725718810909

IDB Analyzer. (n.d.). http://www.iea.nl/data

Japelj Pavešič, B., & Svetlik, K. (2016). *Znanje matematike in naravoslovja med osmošolci v Sloveniji in po svetu. Izsledki raziskave TIMSS 2015* [Knowledge of mathematics and science among eighth-grade pupils in Slovenia and worldwide. Results from the TIMSS 2015 survey]. Pedagoški inštitut. http://timsspei.splet.arnes.si/?page_id=714

Johansson, M. (2003). *Textbooks in mathematics education: A study of textbooks as the potentially implemented curriculum.* Luleå University of Technology. https://www.diva-portal.org/smash/get/diva2:991466/FULLTEXT01.pdf

LaRoche, S., Joncas, M., & Foy, P. (2016). Sample design in TIMSS 2015. In M. O. Martin, I. V. S. Mullis, & M. Hooper (Eds.), Methods and procedures in TIMSS 2015. TIMSS & PIRLS International Study Center, Boston College. http://timssandpirls.bc.edu/publications/timss/2015-methods.html

Martin, M. O. (Ed.) (2005a). *TIMSS 2003 User guide for the international database: TIMSS Almanacs 2003. School background data almanac by SCIENCE achievement – 8th grade.* TIMSS & PIRLS International Study Center, Boston College. https://timss.bc.edu/timss2003i/userguide.html

Martin, M. O. (Ed.) (2005b). *TIMSS 2003 user guide for the international database: TIMSS Almanacs 2003. Mathematics teacher background data almanac - 8th grade.* TIMSS & PIRLS International Study Center, Boston College. TIMSS Almanacs 2003. https://timss.bc.edu/timss2003i/userguide.html

Martin, M. O. (Ed.) (2005c). *TIMSS 2003 user guide for the international database:* TIMSS Almanacs 2003. *Science teacher background data almanac - 8th grade*. TIMSS & PIRLS International Study Center, Boston College. TIMSS Almanacs 2003. https://timss.bc.edu/timss2003i/userguide.html

Martin, M. O., Mullis, I. V. S., Foy, P., & Hooper, M. (2016). *TIMSS 2015 international results in science.* Boston College. http://timssandpirls.bc.edu/timss2015/international-results/

Mili, & Winch, C. (2019). Teaching through textbooks: Teachers as practitioners of a discipline? *Theory and Research in Education, 17*(2), 1–21. https://doi.org/10.1177/1477878519862547

Min, I., Cortina, K. S., & Miller, K. F. (2016). Modesty bias and the attitude-achievement paradox across nations: A reanalysis of TIMSS. *Learning and Individual Differences*, *51*, 359–366. https://doi.org/10.1016/j.lindif.2016.09.008

Mullis, I. V. S., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 international results in mathematics*. TIMSS & PIRLS International Study Center.

Mullis, I. V. S, Martin, M. O., Foy, P., & Hooper, M. (2016). *TIMSS 2015 international results in mathematics*. TIMSS & PIRLS International Study Center. http://timssandpirls.bc.edu/timss2015/international-results/

Oakes, J., & Saunders, M. (2004). Education's most basic tools: Access to textbooks and instructional materials in California's public schools. *Teachers College Record*, *106*(10), 1967–1988. https://doi.org/10.1111/j.1467-9620.2004.00423.x

Oates, T. (2014). *Why textbooks count*. University of Cambridge, Local Examinations Syndicate. https://www.cambridgeassessment.org.uk/Images/181744-why-textbooks-count-tim-oates.pdf

Portal Trubar (n.d.). https://paka3.mss.edus.si/Trubar/Javno/default.aspx

Siegler, R. S., & Oppenzato, C. O. (2021). Missing input: How imbalanced distributions of textbook problems affect mathematics learning. *Child Development Perspectives*, *15*(2), 76–82. https://doi.org/10.1111/cdep.12402

Sievert, H., van den Ham, A. K., & Heinze, A. (2021). Are first graders' arithmetic skills related to the quality of mathematics textbooks? A study on students' use of arithmetic principles. *Learning and Instruction*, *71*, 101401. https://doi.org/10.1016/j.learninstruc.2020.101401

Appendix

Table 1

*Mathematics textbooks and attitudes – t-test results**

Mathematics Textbook		Student value mathematics		Student like mathematics		Student confidence in mathematics		Engaging teaching	
reference group	compared group	diff_se	diff_t	diff_se	diff_t	diff_se	diff_t	diff_se	diff_t
1	2	0.39	-0.07	0.41	-0.45	0.39	0.36	0.46	-0.81
1	3	0.39	-0.49	0.43	-0.69	0.41	-0.17	0.48	-1.14
1	4	0.44	0.45	0.45	0.23	0.46	0.47	0.52	-0.38
1	No MT	0.39	0.12	0.42	-0.46	0.40	0.26	0.47	-0.97
2	3	0.08	-2.16	0.15	-0.77	0.11	-1.92	0.14	-1.24
2	4	0.24	0.96	0.21	1.35	0.25	0.31	0.25	0.71
2	No MT	0.08	0.98	0.11	-0.09	0.08	-0.46	0.11	-0.79
3	4	0.24	1.59	0.25	1.59	0.29	0.99	0.28	1.23
3	No MT	0.08	2.81	0.13	0.84	0.10	1.67	0.14	0.63
4	1	0.44	-0.45	0.45	-0.23	0.46	-0.47	0.52	0.38
4	2	0.24	-0.96	0.21	-1.35	0.25	-0.31	0.25	-0.71
4	3	0.24	-1.59	0.25	-1.59	0.29	-0.99	0.28	-1.23
4	No MT	0.24	-0.63	0.22	-1.35	0.25	-0.44	0.25	-1.02

Note. No MT = Group of students without prescribed Mathematics textbook.*Textbooks TM5 and TM6 were used by a very small share of the students, so a t-test is not available.

Table 2

Physics textbooks and attitudes – t-test results

Physics Textbook		Student like physics		Student confidence in physics		Engaging physics teaching	
reference group	compared group	diff_se	diff_t	diff_se	diff_t	diff_se	diff_t
1	2	0.86	0.30	0.74	0.75	0.81	0.69
1	3	0.86	0.14	0.74	0.74	0.80	0.54
1	No PT	0.87	0.11	0.75	0.61	0.82	0.50
2	3	0.15	-0.96	0.13	-0.12	0.16	-0.81
2	No PT	0.20	-0.81	0.18	-0.54	0.21	-0.75
3	1	0.86	-0.14	0.74	-0.74	0.80	-0.54
3	No PT	0.17	-0.10	0.17	-0.48	0.19	-0.14

Note. No PT = Group of students without prescribed Physics textbook.

Table 3

Chemistry textbooks and attitudes – t-test results

Chemistry Textbook		Student like chemistry		Student confidence in chemistry		Engaging chemistry teaching	
reference group	compared group	diff_se	diff_t	diff_se	diff_t	diff_se	diff_t
1	2	0.18	-0.73	0.15	-0.63	0.23	-0.12
1	3	0.16	-0.34	0.15	-1.34	0.18	-0.30
1	4	0.14	-0,46	0.16	-0.98	0.14	0.67
1	No CT	0.16	-0.72	0.18	-1.15	0.18	0.17
2	3	0.19	0.41	0.17	-0.65	0.25	-0.09
2	4	0.18	0.40	0.16	-0.33	0.23	0.54
2	No CT	0.19	0.08	0.19	-0.60	0.25	0.24
3	2	0.19	-0.41	0.17	0.65	0.25	0.09
3	4	0.16	-0.06	0.17	0.34	0.18	0.83
3	No CT	0.16	-0.39	0.18	-0.01	0.19	0.43
4	3	0.16	0.06	0.17	-0.34	0.18	-0.83
4	No CT	0.16	-0.34	0.19	-0.30	0.18	-0.36

Note. No CT = Group of students without prescribed Chemistry textbook.

Table 4

Biology textbooks and attitudes – t-test results

Biology Textbook		Student like biology		Student confidence in biology		Engaging biology teaching	
reference group	compared group	diff_se	diff_t	diff_se	diff_t	diff_se	diff_t
1	2	0.15	0.81	0.12	1.34	0.15	1.72
1	No BT	0.24	0.35	0.23	0.31	0.23	0.54
2	No BT	0.24	-0.16	0.23	-0.41	0.22	-0.59

Note. No BT = Group of students without prescribed Biology textbook.

Biographical note

Barbara Japelj Pavešić is a researcher at the Educational Research Institute, involved in international large-scale assessments in education, nationally coordinating Trends in International Mathematics and Science Study, TIMSS, and The OECD Teaching and Learning International Survey, TALIS. With the background in mathematics and statistics, her field of research is the statistical modelling of complex data to explain the knowledge and learning of mathematics and science of students K-13.

Gašper Cankar, PhD, is a researcher and methodology expert working on problems of knowledge assessment in a National Examinations Centre (NEC, Slovenia) methodologically supporting general and vocational Matura examinations and national assessment in Grades 6 and 9. Some of the topics he works on include error of rating process, value added models of knowledge, interplay between socioeconomic background and educational outcomes and Simpson's paradox.

doi: 10.26529/cepsj.1288

Textbooks and Teaching Materials in Rural Schools: A Systematic Review

Núria Carrete-Marín[1] and Laura Domingo-Peñafiel*[2]

This paper presents the results of a research project whose main purpose is to analyse the concept of multigrade teaching resources and the teaching materials used by teachers in rural schools, in particular the role of textbooks. The use and dimensions of teaching materials are studied in order to promote inclusion and learning in multigrade classrooms with children of different ages mixed together. The present systematic review aims to identify and analyse all of the research papers published internationally on teaching resources in rural schools for the Web of Science and Scopus databases (from 1992 to 2021) and Google Scholar (between 2010 and February 2021). Due to the dearth of publications focused on the topic of study, the reviewed articles have broad inclusion and exclusion criteria. This gives relevance and an innovative character to the research, allowing us to objectify the state of the question on multigrade didactic materials and their relation to teaching-learning processes. From a total of 332 research papers in the field of rural multigrade teaching identified for further analysis, only papers that met the inclusion and exclusion criteria and passed all phases of the PRISMA flow diagram were used ($N = 33$). Some research publications contributed to identifying opportunities and needs, and to suggesting criteria to be taken into account for the selection and creation of materials to promote inclusion and active learning methodologies. The first results show the need to create one's own materials that analyse the reality of these schools, as well as the need to personalise and adapt printed or digital textbooks and other teaching materials in order to involve the students actively in the learning process and to respond to the needs of rural students in multigrade classrooms.

Keywords: multigrade classrooms, rural school, teaching material, textbooks, systematic review

1 PhD Student at Faculty of Education, Translation, Sport and Psychology, University of Vic-Central University of Catalonia, Spain.

2 *Corresponding Author. Faculty of Education, Translation, Sport and Psychology, University of Vic-Central University of Catalonia, Spain; laura.domingo@uvic.cat.

Učbeniki in učna gradiva podeželskih šol: sistematični pregled

Núria Carrete-Marín in Laura Domingo-Peñafiel

V prispevku predstavljamo rezultate raziskovalnega projekta, katerega glavni namen je bil analiza koncepta kombiniranih učnih sredstev in učnih gradiv, ki jih uporabljajo učitelji podeželskih šol, zlasti vloge učbenikov. Uporaba in razsežnosti učbenikov so preučene z vidika spodbujanja inkluzije in učenja v kombiniranih oddelkih z otroki različnih starosti. Sistematični pregled skuša prepoznati in preiskati vse mednarodno objavljene prispevke o učnih gradivih v podeželskih šolah z uporabo podatkovnih baz *Web of Science* in *Scopus* (1992–2021) ter *Google Učenjaka* (2010–2021). Zaradi pomanjkanja publikacij, ki bi se osredinjale na izbrano tematiko, imajo pregledani članki blaga merila glede vključenosti oz. izključenosti. To daje raziskavi relevantnost in inovativnost, kar nam dopušča, da objektiviziramo stanje glede kombiniranega didaktičnega gradiva in njihovo razmerje do procesov poučevanja/učenja. Od skupno 332 člankov s področja poučevanja v kombiniranih oddelkih podeželskih šol, ki smo jih vključili v obdelavo, so bili nazadnje uporabljeni samo prispevki, ki so zadostili merilom in prestali vse faze diagrama PRISMA (N = 33). Nekatere raziskovalne publikacije so prispevale k prepoznanju priložnosti in potreb ter k predlaganju meril, ki bi jih bilo treba upoštevati pri izboru in ustvarjanju gradiv za spodbujanje inkluzije in metod aktivnega učenja. Prvi izsledki kažejo potrebo po oblikovanju lastnih gradiv, pri čemer bi se preučilo samo realnost teh šol, pri tem pa tudi potrebo po personalizaciji in prilagoditvi tiskanih ali digitalnih učbenikov ter drugega učnega gradiva, da bi lahko učence aktivno vključevali v učni proces in se odzivali na potrebe podeželskih učencev v kombiniranih oddelkih.

Ključne besede: kombinirani oddelki, podeželska šola, učno gradivo, učbenik, sistematični pregled

Introduction

One of the main characteristics of rural schools is the grouping of students into multigrade classrooms, either for structural reasons or by choice. This configuration requires a response adjusted to the diversity of the classrooms and the rural context in order to take advantage of this context and the differences between students, thus enriching joint learning (Boix & Bustos, 2014; Domingo-Peñafiel & Boix, 2015; Santos, 2011) by considering different aspects of multigrade didactics, such as teaching resources to support the methodology used. The objective of the present research is to deepen the conceptualisation and the state of the art of teaching materials in the multigrade classroom of the rural school. The research is necessary due to the lack of publications focused on this topic despite its importance, as manifested in various publications (Coladarci, 2007; Fargas-Malet & Bagley, 2021). In order to help fill the research gap on this topic, the study presents a systematic review of the main findings of the existing literature on the subject, defines the term "multigrade teaching resources", and provides relevant information on the typology of resources used in the multigrade classroom, as well as on the needs and opportunities in this regard and the criteria for its development, thus making the study relevant.

Given the heterogeneity of multigrade classrooms in rural schools, working together with students of different ages and characteristics and taking advantage of their pedagogical and inclusive value involves changes in methodology and requires teaching-learning processes that are different from those used in single-grade classrooms. Teachers have a very important role to play in 'multigrade didactics' (Bustos, 2007), implementing innovative and globalised teaching strategies that take into account the contextual particularities of the rural school. These methodological changes must be accompanied by didactic materials that allow the development of multigrade teaching, with adaptation of the teaching response to the diversity of the classrooms being essential (Boix & Bustos, 2014). It is crucial to avoid the hegemony of actions carried out in graded classrooms (Terigi, 2009) that do not fit this school model. In many cases, this is what happens with textbooks, which are one of the most used resources. Textbooks are created from a graded and homogeneous school model based on the curriculum and education legislation. Furthermore, they condition the teaching methodology according to the commercial activity of the publishing sector (Fernández Palop & Caballero García, 2017). This does, however, depend on the national education policy in some countries, as there is evidence of the creation of some textbooks for teachers and students in rural schools, as shown in certain studies in South America and Finland (Da Silva & Cardioso

Filho, 2017; Juárez, 2012; Santamaría Pérez, 2021) indicating that governmental guidelines have had an impact on school. Although such textbooks provide guidance, they do not fully reinforce the pedagogical value of the multigrade classroom, nor are they perceived by teachers as a useful resource adapted to the context and the multigrade classroom (Boix & Bustos, 2014; Bustos, 2007). For teaching material to be useful, it requires reflection by teachers in order to adapt it to the methodology used and the characteristics of the students in the classroom (Boix & Bustos, 2014; Santos, 2011). It also requires teacher training that is not based on urbanised patterns, so as not to import a standardised and graded teaching model. Rural classrooms require a specific pedagogical model due to their organisational peculiarities, whereby teaching and learning processes, as well as the educational response, have to be adapted to the maturity and learning level of each student in confluence with other students of different ages (Terigi, 2009).

The lack of specific and adequate didactic resources to address teaching and learning in multigrade classrooms, together with the lack of knowledge of their nature, is a recurrent problem that is highlighted in research and publications in the field of rural education and multigrade didactics (Brown, 2010; Juvane, 2005; Msimanga, 2019). The creation and adaptation of materials is therefore a challenge for teachers in rural schools. The invisibility of the rural school is evidenced by the scarcity of published studies on the subject, despite its obvious importance (Bagley & Fargas-Malet, 2021). There are no research activities focusing exclusively on the field of multigrade resources (Coladarci, 2007; Fargas-Malet & Bagley, 2021; OECD, 2019). The main focus of the present article is to review internationally published research in order to define the concept of multigrade materials and examine their reality in practice, as well as to identify opportunities and needs, and to contribute to the generation of knowledge on the subject. The theoretical review enables us to make proposals for the future improvement of teaching-learning processes and the task of teaching based on a critical documentary analysis of the reviewed publications.

The present paper complements previous publications in this journal about printed and digital textbooks (Ivic, 2019), as well as another systematic review (Devetak & Ferk Savec, 2020), by presenting unpublished research from a systematic international review on teaching materials and textbooks in the context of multigrade classrooms in rural schools. According to a recent scientific literature review about research on rural schools (Fargas-Malet & Bagley, 2021), this is a relevant issue that needs to be addressed.

The relevance and role of textbooks and teaching materials in multigrade rural classrooms

The organisation of children of different ages in the same classroom involves a change in the methodological approach to teaching and learning. Didactic proposals must be multigrade and contextualised, must consider the rural school environment and the territorial dimension of the curricular proposals, and must be interdisciplinary and meaningful (Boix & Domingo-Peñafiel, 2019). This methodological change means that teachers also have to create and adapt resources to this way of teaching, based on the proposals developed and the specific didactics of multigrade classrooms (Bustos, 2007). This requires a firm conviction on the part of teachers regarding the pedagogical value of multigrade classrooms and the positive interactions that take place in such classrooms. In the teaching and learning process in the multigrade classroom there are two components that can be considered basic: learning activities and teaching resources (Boix & Bustos, 2014). Special emphasis is placed on teaching materials and their variety in order to be able to address and provide an adequate educational response to the full diversity of students in the multigrade classroom. However, the mere presence of teaching materials does not guarantee the quality of teaching and learning. The materials must support didactic strategies in an effective way in order to guide learning. In any case, it is necessary to emphasise the role of teachers in adapting them to the educational situation and to the didactic objectives. This aspect is even more relevant in the case of rural schools when, through teacher reflection, it comes to adapting, creating or designing materials that promote the pedagogical value of the multigrade classroom and its rural context, allowing them to respond to the diversity of students in terms of equity (Cornish, 2006; Domingo-Peñafiel & Boix, 2015). Materials must be adapted to the specific contextual requirements of each classroom. Textbooks and other multigrade materials need to be varied and diversified to help ensure educational success in multigrade classes (Santos, 2011). Moreover, they should be properly organised and arranged in the classroom so as to facilitate learning and ensure interaction between students (Joubert, 2010).

Other studies highlight the inadequacy of existing multigrade resources as a challenge for teachers in multigrade classrooms. Such resources represent one of the needs of rural schools in their commitment to multigrade learning (Boix & Bustos, 2014; Fargas-Malet & Bagley, 2021; Juvane, 2005; McEwan, 2008) and indicate a line of research still to be investigated. This has been taken into account in the approach of the present research, which also highlights training needs and the shortcomings of initial training and throughout professional development. Teacher training should contribute to a belief in the pedagogical value of multigrade classrooms and their pedagogical possibilities, taking multigrade resources as an indispensable aspect of multigrade didactics. Little (2005) argued that materials and resources should be created specifically for multigrade classrooms. It is important to teach using materials adapted and contextualised to the existing reality in these schools, to the multigrade classroom and the rural context, and not to simply use single-grade materials and textbooks that are not adapted to the these conditions. In general terms, resources are focused on the achievement of curricular learning and do not usually take into account the particularities of the contexts. It is necessary to review these resources with all schools and their educational needs in mind in order to provide support to teachers for their adaptation or development.

Interest in the rural school as an object of study has re-emerged in the twenty-first century, recognising the value of its school model and didactics. Nonetheless, there is still a lack of existing repositories of textbooks, multigrade materials and publications on the subject. There is a need to disseminate and analyse materials that have been created in this context and to guide teachers in overcoming the difficulties associated with the creation of teaching materials for multigrade classrooms (Brown, 2010). Teachers should be encouraged to reflect on the existing materials, to assess their suitability and to create their own materials adapted to the characteristics of their multigrade classroom and rural context. It is difficult to find materials that are adapted (Genc, 2016) and having guidelines to assist their creation or adaptation would also facilitate their development (Petrie & Darragh, 2018). A more detailed analysis of textbooks and teaching materials is needed in order to shed more light on the current state of this issue. The present review therefore provides a detailed examination of how the meaning of multigrade teaching resources is currently conceived in the literature and considers the types of materials used in rural schools, especially textbooks, and their role in teaching and learning processes. It analyses the needs and opportunities related to the subject and proposes criteria for the creation, adaptation or selection of textbooks and other materials to suit multigrade education and the context. The study also aims to highlight the importance of multigrade resources as a significant aspect to be considered by teachers when making innovative classroom proposals. The aim is not only to describe the state of the art, but also to implement future improvements in teaching practice.

The present study was conducted in order to examine all of the papers published on the subject by recognised scientific journals in Web of Science

(WOS) and Scopus up to the time of the search (1 February 2021). After searching the databases, the results obtained on teaching materials in rural schools dated from 1992 to 2021. Due the scarcity of studies on the topic, and in order to expand the references and information from popular sources, search results from Google Scholar were added in order to examine the most recent publications, appearing from 2010 to February 2021. The time range for Google Scholar had to be narrowed due to the large number of results, so that only those from the last ten years are selected. The aim was to examine the evidence on multigrade teaching materials in rural schools in order to establish a definition of materials for rural multigrade classrooms and to examine the state of the art.

Based on this research focus, the following research questions can be formed:

- What are the characteristics of the research published on textbooks and teaching materials used in rural schools? What are the main findings that could help to determine how multigrade teaching resources are conceptualised? What is the state of the art?

  Our research objectives are the following: (1) to describe and analyse the concept of multigrade resources; (2) to review and analyse the types of didactic materials, especially textbooks, used by teachers in rural schools, as reflected in the selected publications; (3) to identify opportunities and needs regarding didactic materials; and (4) to identify the dimensions and criteria that a resource must have in order to be multigrade and pedagogically significant.

Method

A systematic review was undertaken to describe, through a conceptual theoretical investigation, the nature of multigrade resources and the typologies that can be found in multigrade classrooms according to studies published internationally in the selected databases. Aspects to be considered in future research related to innovative proposals on teaching resources and their dimensions are also highlighted. The systematic review configures a qualitative research methodology that is appropriately adapted to the achievement of the proposed objectives. It systematically and objectively summarises the results of various qualitative and quantitative studies in order to answer the research questions and objectives posed (Sánchez-Meca & Botella, 2010; Smith et al., 2011).

Search strategy and parameters

In the first step of the study, the article selection criteria and search parameters were determined. The inclusion criteria for papers were as follows: (1) the studies were related to multigrade teaching materials or resources in rural schools; (2) all of the studies were selected for WOS and Scopus and, from the period from 2010 to February 2021, for Google Scholar; (3) the studies were reported in languages that are intelligible to the researchers, i.e., English, Spanish, Portuguese, Catalan or French; 4) due of the international nature of the study, which investigates all research focused on the types of teaching materials used by rural teachers, their use, problems and needs, as well as the conceptualisation of such materials, all of the relevant studies were included regardless of their geographical origin; the papers included were open access articles listed in the databases used and accessible to the researchers; and (5) all relevant primary studies on the subject were included, but systematic reviews were excluded. All publications in the scientific field of social sciences, the area of multigrade teaching or education, and the subarea of materials or resources in the rural education context were included in the research.

The WOS, Scopus and Google Scholar databases were reviewed with the same search parameters. Google Scholar was included in order to broaden the range of publications of a more informative nature and thus provide more information. Due to the scarcity of more recent results, the last ten years were used in Google Scholar, with particular attention devoted to the most current results.

A systematic search was carried out by combining Boolean indicators using the main concepts that allow definition of the object under study and are linked to the object of analysis. With the aim of finding the most suitable articles for the object of study, the strategy of searching with more specific terms was used, also taking into account possible orthographic or synonymous variations. Thus, AND/OR operators and the wildcard (*) were used to find possible endings of the same word. Finally, the following search string was set up: multigrade AND (teach* OR class* OR education*) AND (material* OR resource*). This combination of terms related to the object of study was chosen as it provided the best and most accurate results. The search was done for abstract, title and keywords. This search allowed us to find articles in the field of social sciences, and specifically rural education, that relate to teaching materials in the multigrade classroom.

A total of 332 papers were identified. After deleting duplicates, 294 sources remained. The final selection was conducted using the PRISMA (Preferred

Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram (Moher et al., 2009), which ensures the quality treatment and selection of research articles and data, as well as enhancing the rigour and reliability of the obtained results (Figure 1). Content analysis was performed on the final 33 selected studies related to teaching materials in the multigrade rural school context.

Filtering in the paper selection process and data extraction

For the paper selection process, the phases of the PRISMA flow chart were followed (Moher et al., 2009). The research procedure was characterised by an initial heuristic phase based on the identification of all publications from the selected databases. In the screening phase, the results were filtered out excluding duplicates. Results that did not meet the search objectives were also removed by the authors after reading the title, abstract or keywords. In the subsequent adaptation phase, the full text of the resulting articles was reviewed based on predetermined selection criteria for their possible relevance. Finally, in the last phase of the process, only the publications to be analysed in the systematic review were included.

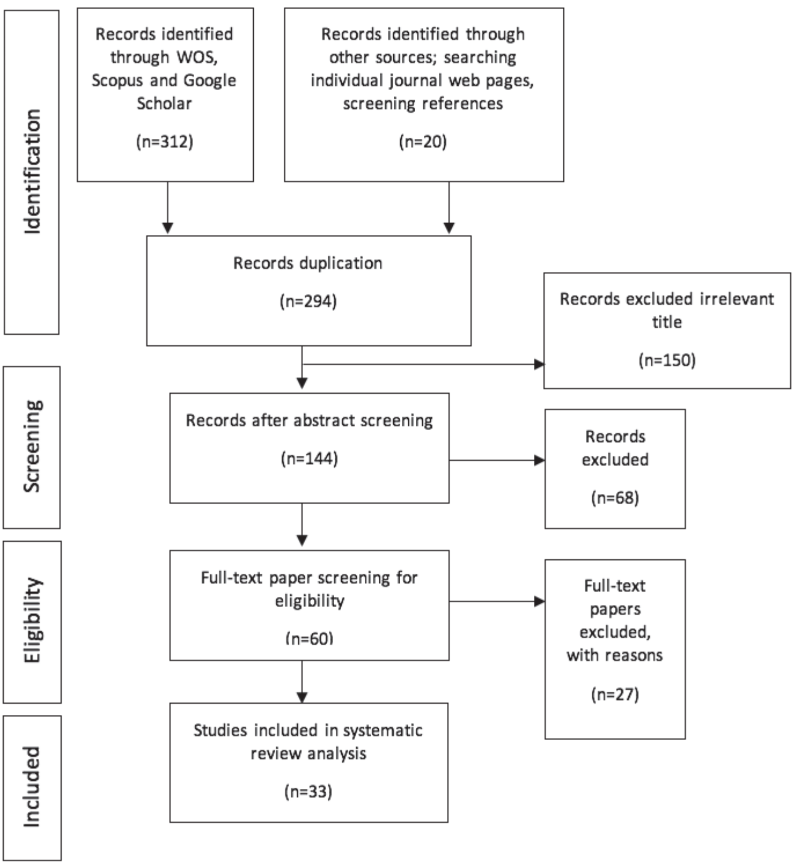
The resulting articles were recorded in descriptive tables or guides for qualitative synthesis. From there, the information was managed and classified with Mendeley's programming and Excel tables. An initial phase of data reduction was carried out with Atlas Ti through a system of codes and categories, performing various descriptive analysis actions (reporting of codes, frequency of events, citations, memos and word clouds) in order to organise and describe the information obtained. In the second phase, actions were carried out to critically interpret the results (analysis of code networks, analysis of interpretation of citations or tables of co-occurrences) in order to enable theorisation. A critical and rigorous analysis of the 33 publications was carried out, taking into account ethics and respecting the positions of the different authors.

Results and discussion

As reflected in the flow diagram (Figure 1), only 294 of the 332 articles initially identified in the databases passed the screening phase, eliminating duplicates (38 records). A further 150 articles were excluded due to irrelevant titles. After reading the abstract, and based on the inclusion/exclusion criteria, 68 studies were withdrawn from complete analysis, leaving a total of 144 publications. The main reason for rejecting the aforementioned 68 papers was that their content was repetitive or unspecific, or they contained information linked to the terminology of the subject but unrelated to the objectives of the present study. In addition, some of them were not in open access in the databases and the full text could not be obtained by the researchers. Consequently, only 60 articles passed the eligibility phase. Of these, a further 27 papers were excluded after reading the full text, mainly because their content deviated from the central question and were not aligned with the objectives and focus of the research. Finally, in the last phase of the process, a total of 33 published papers were included in the systematic review analysis.

Figure 1

PRISMA flow diagram for the present study



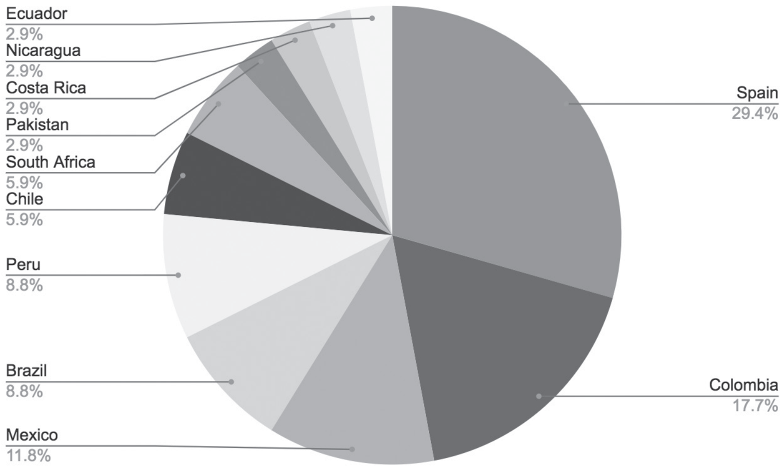
Note. Adapted from Moher et al., 2009.

Characteristics of the analysed studies

Due to the international focus of the systematic review, articles were included regardless of their geographical origin, as noted above. The final studies selected present contextual characteristics that must be taken into account. The international articles included were mostly based on research conducted in Spain, Colombia, Brazil, Peru and South Africa, in that order (Figure 2). This indicates which countries have devoted more attention to the topic of multigrade resources in their research on rural education and multigrade didactics.

Figure 2

Country of origin of the reviewed articles.



The languages of the articles coincide with the majority language of the countries of origin shown above. Most are written in Spanish (N = 22), followed by English (N = 9) and Portuguese (N = 2). There are no selected records in other languages in the check result, which shows the appropriate choice of selection criteria in terms of the language and country of origin of the articles.

Most of the papers reviewed belong to education journals (N = 21), followed by doctoral theses or university projects, and finally lectures and books (Table 1). This theoretically supports the research, as most of the studies belong to peer-reviewed education journals.

Table 1
Type of papers included in the sample

Total sample of papers = 33	
Source	***N***
Education journals	21
Theses	5
Bachelor's degree final theses	3
Postgraduate final theses	2
Conferences	1
Books	1

The majority of the papers reviewed deal with the subject of multigrade teaching resources in a superficial manner, so an exhaustive review of each of them was required (Figure 3). Most of them are related to multigrade didactics and teachers (48.6%), followed by those referring to resources in general (20%) and those of a specific nature (25.7%). The publications highlight the use of Information and Communication Technologies, followed by textbooks (Figure 4), either printed or digital. These are the resources most frequently used in multigrade classrooms. In the present review, special emphasis is therefore placed on textbooks, both printed and digital. The analysis of these data highlights the lack of specificity in the treatment of resources in rural schools, with even less attention being devoted to multigrade schools.

Figure 3
Focus of the papers reviewed

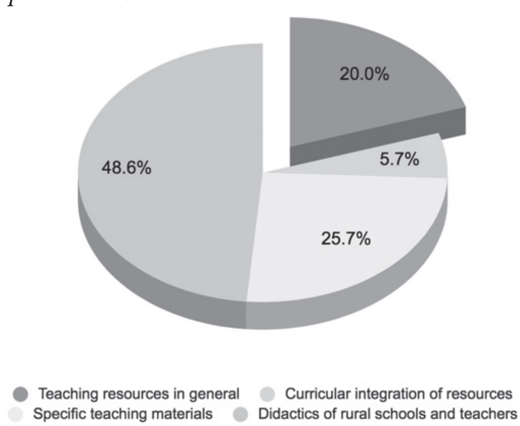
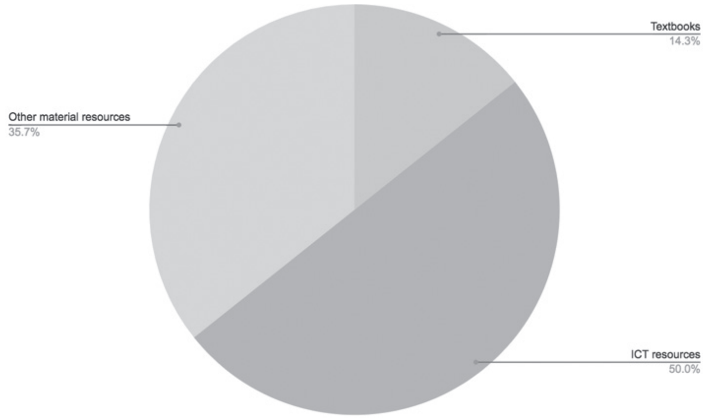


Figure 4
Main teaching resources in the papers



Regarding the design of the studies, most of them are qualitative (75.5%), others use mixed methods (15.5%) and a minority are quantitative (9.1%). This reflects not only the great variety of studies and methodologies on the topic, but also the need for more studies of a mixed type, including both quantitative and quantitative data, in order to obtain data of a diverse nature that complement each other and enable an increase of knowledge on the subject.

Characteristics of the studies analysed in terms of content

The analysed studies were divided into four sub-sections of the main research fields: (1) the concept of multigrade teaching resources, theoretical bases and the state of the art; (2) the types of didactic materials, especially textbooks; (3) the opportunities and needs based on multigrade resources; and (4) the dimensions or criteria that multigrade material and textbooks must have.

The concept and the state of the art of textbooks and multigrade teaching resources

Despite the importance given to multigrade teaching materials for educational improvement and the success of multigrade teaching (Bustos, 2007) in the reviewed publications, the review highlights the fact that there is no explicit and consensual definition of the concept of multigrade teaching resources in the context of rural schools. According to some studies (Brown, 2010; Msimanga, 2019), there is a lack of knowledge about the nature of these resources. Although this terminology is not used, most papers nonetheless present various proposals to define all of the didactic resources in the framework of multigrade didactics (Coaquira, 2018; López, 2011; Msimanga, 2019; Rahim, 2018). Various terms are used, such as learning resources, support materials like textbooks, learning media, teaching aids, equipment and teaching tools (Juárez, 2012; Msimanga, 2019). From the outset, it is necessary to analyse the terms used to designate learning materials, as they are often referred to as instructional materials or self-learning materials, definitions that are more related to one-way, self-directed learning by the learner. It is necessary to use terminologies that relate the materials to discovery-related and active-participatory learning (Newby et al., 1998 in Drinkwater, 2002). The materials must be coherent with the work of the multigrade classroom, avoiding connotations far removed from the work with active methodologies in the multigrade classroom. More contextualised learning in line with the multigrade classroom also implies changes in the type of materials and their connotation.

The review also highlights the fact that textbooks and teaching materials must be as diverse as possible (López, 2011), taking into account multiple intelligences, the diversification of learning styles (Msimanga, 2019), different levels of complexity and forms of presentation (Block, et al., 2017). They can be specifically created for multigrade learning or used in the classroom context for this purpose (Silva et al., 2020). From the information in the various articles, it can be seen how an initial distinction can be made between specifically multigrade materials, such as textbooks or other curricular materials created with the multigrade classroom in mind, and materials that help to promote, improve or facilitate multigrade work through their use and teaching action, even though they are not specifically created for the multigrade classroom; for example, manipulative materials created with elements from the environment. The latter are conditioned by the methodological strategies applied and by teaching planning (Silva et al., 2020). In addition, research indicates that the boundary between didactic materials or resource, activity and strategy is not clear. In some cases, although reference is made to teaching materials, it is evident from the examples that the studies actually refer to classroom methodology (Dávila, 2018). This demonstrates the need to conceptualise and delimit the term teaching materials, so that it is clear what it refers to. Moreover, most of the studies conclude that textbooks and teaching materials should facilitate the interpretation of content and encourage cooperation and interaction between learners of different ages; they must be specific and appropriate to the configuration of multigrade classrooms and their context (Marhot, 2001). The importance of the existence of materials or textbooks in line with the multigrade classroom and the contextual reality of each rural school is highlighted (Boix & Bustos, 2014).

In fact, several authors (García Prieto & Pozuelos Estrada, 2017; Guerra, et al., 2019) refer to textbooks as the main material used in rural school classrooms, classifying children by levels, which is very different from the perspective of multigrade work. More creative use of these materials, based on teacher reflection, is important in order to support classroom projects (Domingo-Peñafiel & Boix, 2015; García Prieto et al., 2017); for example, the adaption of existing materials so that they can respond to all of the students regardless of their characteristics and be useful for supporting active methodologies.

Finally, based on the results and the discussion of the data, an approximation of the definition of multigrade teaching resources in the context of rural schools can be made. Multigrade teaching resources are all those diversified and accessible teaching aids designed to support learners and teachers, and which facilitate and improve teaching and learning processes in multigrade classrooms. They contribute to providing an educational response to diversity by creatively adapting to children, regardless of their age, level of knowledge or competence. They are created or adapted for educational benefit. They must be able to successfully support specific multigrade teaching. They can be curricular in nature, like textbooks, and can either address multigrade teaching itself or have a specific application that contributes to meaningful work according to the methodology used. These are materials that allow us to give a useful, joint and inclusive educational response to all of the students in the multigrade classroom, and that do not respond to homogeneous graded classrooms working separately by grades. They are a concretisation of the general didactic resources adapted to the idiosyncrasies of rural schools, as they must be useful and in line with the nature of the didactic strategies used, which are multigrade, contextualised, interdisciplinary and significant. In this regard, specific dimensions have to be taken into account in the elaboration or use of these resources, so that they are adapted and in accordance with the context, the needs of the classroom and the methodological strategies used to positively promote multigrade. Furthermore, as García Prieto (2015) mentions, it is necessary to move towards deliberate curricular materials. Teaching materials and textbooks should be a resource that is open to decisive teacher intervention and that does not allow for automatic transfer to the classroom. Such materials need adaptations that depend on the context and the group of learners. They must be flexible with regard to the different ages and characteristics of the students in the multigrade classroom, so that they can be revised with the understanding that not all rural schools or multigrade classrooms are the same, thus preventing homogenisation (Domingo-Peñafiel & Boix, 2015; Tomlinson, 2001). In fact, all didactic materials must be flexible so that the teacher can adapt them to each classroom, to the students and to the circumstances. This aspect is especially relevant in the case of didactic materials referring to rural versus urban schools, where it is also necessary to face and adapt to the diversity of ages that interact in the same classroom, the strategies used in multigrade teaching and the management of time and space to enhance the pedagogical value of the multigrade classroom and the joint circulation of knowledge.

The research also indicates other important considerations for defining the state of the art. Two papers (Msimanga, 2019; Torres et al,. 2019) highlight Montessori educational theory as a point of reference for the construction of multigrade materials. This theory attaches particular importance to teaching materials to promote meaningful learning, that is, the materials should encourage learner autonomy and active learning through discovery, and should be flexible and include manipulatives. Most of the studies reviewed find that although teachers strive to develop their own teaching materials, many of them do not know how to design materials that are appropriate in this context (Brown, 2010; Bustos, 2014). The works consulted highlight the need for teaching materials and textbooks to be specific to multigrade teaching, taking into account the inclusive and contextual dimension, and improving knowledge of the territory and the children's sense of identity. They also emphasise the need for student and community participation in the development of classroom materials (Bustos, 2009; Säg, 2009), while the need for a multigrade and cyclical curriculum that allows for interdisciplinary work is also highlighted (Juárez, 2012). Finally, from the studies reviewed it can be concluded that there is a lack of specific teacher training to introduce changes in teaching and to be able to design curricular materials in accordance with the idiosyncrasies of rural schools.

Types of teaching materials

The analysis of publications focusing on the types of teaching materials used by teachers in rural multigrade classrooms shows various classifications. These refer, for instance, to tangible, intangible, rural, digital, printed, tactile, auditory, visual or audiovisual, didactic-technological or programmable resources (García Prieto, 2015; Guevara et al., 2012; Msimanga, 2019; Rahim, 2018; Fraser et al., 1993). Moreover, materials can be adapted, edited or self-created (Msimanga, 2019). It is observed that the use of technological resources for didactic purposes must be suitable for multigrade work (Miranda et al., 2018).

The human resources of the educational community, families or students themselves are also considered (García Prieto, 2015; Mathot, 2001). Physical resources related to school infrastructure, spaces and the rural context itself are also mentioned (Rahim, 2018). It can be concluded from the publications analysed that teaching materials can be of a curricular nature, such as textbooks that take the multigrade into account in their development, or unspecific materials for multigrade work, where their appropriateness is determined by the methodology used. Both can be intended for individual or group work (Afzalnia, 2005; Msimanga, 2019). Most of the reviewed publications refer to textbooks, whether printed or digital, and the need to adapt them to multigrade work. In addition, the systematic review shows that textbooks and digital resources are most frequently mentioned in the articles as the main teaching resources in multigrade rural school classrooms (Boix & Bustos, 2014; Boix & Domingo-Peñafiel, 2014; Bustos, 2014; Brown, 2010; Forero et al., 2016; García Prieto, 2015; Guerra, et al., 2019; Juárez, 2012; López, 2011; Msimanga, 2019; Rodríguez & Saavedra, 2018; San Pedro & López, 2017; Severiano, 2013; Valderrama, et al., 2011). Other resources described in the works and used in multigrade classrooms are Tika workbooks (López, 2011), multimodal resources and textbooks (Cardona, 2020), games, video games (Giraldo, et al. 2018), the curriculum itself as a support document for teachers, and the school garden or eco-garden (Dávila, 2018). In general, it can be surmised from the publications that graded work with didactic materials prevails in most cases, despite efforts to teach in a multigrade manner. Existing research highlights the lack of suitable classroom materials created for this purpose (Torres et al., 2019). In addition, there is no material in the publications that is *a priori* called multigrade. This aspect allows us to determine the need, as mentioned in the papers, to explore the existing resources in rural schools in order to provide help and guidance to teachers for their own creation and adaptation of resources, as well as improving their knowledge by giving them more support (Chaves et al., 2013). Finally, in the context of the present article, an in-depth analysis is carried out of the studies on textbooks as one of the most prevalent didactic resources in rural schools.

Textbooks in multigrade rural classrooms

Most of the papers surveyed conclude that textbooks do not contribute to multigrade didactics and therefore would not be a useful resource for this purpose (García Prieto et al., 2017; Bustos, 2007). Nonetheless, they are still the most widely used resources (Boix & Bustos, 2014). Despite considering textbooks inadequate for the contextual reality, teachers use them to gain security when faced with the difficulty of planning or applying existing materials to the multigrade reality of rural schools (Brown, 2010). This is associated with the lack of teacher training related to rural schooling, multigrade teaching and the acquisition of competencies to serve students regardless of their characteristics in the same classroom in a rural school (García & Machado, 2017). The research revealed that most textbooks are created to work in a graded way; they are decontextualised and contain standard content (García Prieto et al., 2017) that is not adapted to the organisation of multigrade teaching. All of this can also be related to what is happening in the publishing industry, where there is no contact with educational experience or teachers at the time when textbooks and printed teaching materials are developed, and the needs of the different types of schools are not considered (Fernández Palop & Caballero García, 2017).

Several studies show that rural schools cannot adopt materials from graded urban education (García Prieto et al., 2017; Boix & Bustos, 2014). Textbooks, whether printed or digital, are not designed for the multigrade classroom, nor do they take into account the rural context of the students, instead taking graded schools as a point of reference (Bustos, 2007). Some textbooks are elaborated as guides for teachers, but this does not contribute to obtaining adequate guidelines for managing a multigrade classroom or implementing methodological changes (Little, 2005). The articles reviewed highlight the need for textbooks or edited materials that are focused on the reality of schools and the joint work of students of different ages in the same classroom (Msimanga, 2019). Some of the papers value the need for the creation of printed materials that help teachers to carry out work that enhances the inclusive nature of the multigrade classroom, thus facilitating the adaption of strategies and projects. The need for a textbook-based support curriculum that is cyclical and works on different subjects at different levels of complexity is highlighted. Moreover, the need for textbooks to have content that responds to a globalised, interdisciplinary and contextualised approach to learning, including activities with different levels of complexity, is valued in various publications. Such textbooks should promote autonomous individual or group work and cooperation, and should have creative content and a resourceful format that facilitates student discovery, learning and creative thinking. According to the data collected, many textbooks encourage repetition and work separated by school year, which is something that needs to be reconsidered and changed because is it not in line with multigrade didactics (Chaves et al., 2013; Msimanga, 2019). The incorporation of these textbooks would not imply methodological changes and would not be adjusted to the context and idiosyncrasies of rural schools. Furthermore, the importance of self-learning guides or orientation bases to guide students' work with different levels of complexity is also mentioned (Juela & Matailo, 2015). As far as digital books are concerned, the playful aspect of the complementary activities they present, their attractive format and their flexibility are valued in many cases versus printed books, contributing to the completion of network tasks and student motivation. However, digital support does not contribute to methodological change either; most digital resources follow the same line of content and work by levels as the printed textbooks (Forero et al., 2016; Severiano, 2013). The different studies reviewed shed light on how the role of textbooks and their prevalent use in rural schools influences the methodology used by teachers in rural schools. This calls for both a review of their format and content and an assessment of their suitability or necessity in this type of context.

Opportunities and needs based on multigrade resources in rural schools

Some of the studies analyse the opportunities and needs of multigrade teaching materials in rural schools. Two of the papers highlight the barrier of invisibility of some of the multigrade materials developed by teachers in rural schools in different countries, pointing to the need for platforms to share these resources and support teachers in their processing and development. They also highlight the difficulties for teachers in creating such materials, as well as the need for more training to adapt existing materials or create new ones to support the active teaching methodologies implemented in the multigrade classroom (García Prieto, 2015; Pozuelos & Travé, 2004).

Another important aspect that some authors highlight is Montessori Theory as a starting point for the development of manipulative materials that allow students of different ages to work together in the same classroom (Msimanga, 2019). This theory comprises specialised and flexible educational materials that allow it to cater for learners of mixed ages. These materials can support meaningful and discovery learning (Msimanga, 2019; Torres et al.; 2019). This highlights the need to review existing resources along these lines and seek opportunities for the creation of multigrade materials. Different studies also emphasise the fact that in many cases the problem not a lack of materials, but the use of existing ones, as in the case of digital resources. There is also a need to review these materials in order to move towards cooperative work in networks, opening the school to the environment and to the global territory, including working with online resources and making use of new learning modalities (Bustos, 2014; Forero et al., 2016; Mathot; 2001), thus helping to reduce the isolation of rural territories (Carrete-Marín & Domingo-Peñafiel, 2021). In this regard, Severiano (2013) concluded that just having certain resources does not help to generate methodological changes; everything has to be coordinated along the same lines.

In line with earlier remarks on the importance of context, the appearance of the Place Based Education concept in one of the studies analysed (San

Pedro & López, 2017) not only allows us to take into account the didactic possibilities of the environment, but also the need to move towards a new curricular approach that seeks to link classrooms with communities or regions (Smith, 2002, 2007). This idea could be related to the concept of Biogeographic Regions (Rivas-Martínez et al., 2011), making it possible to extrapolate textbooks and teaching materials from one area to another at an international level based on their similarities in terms of contextual characteristics. This would allow the creation of new networks of teachers around the world to share materials and meaningful proposals.

Criteria for multigrade teaching materials: Proposals for improvement

Another important aspect of the research is the detection and proposal of criteria to be considered for the creation or adaptation of teaching resources suitable for the multigrade classroom. The analysis of the information in most of the papers has allowed us to establish a proposal for the dimensions that must be taken into account in teaching materials in order to ensure their suitability (Table 2).

Table 2

Criteria to develop multigrade teaching materials

Interdisciplinary and globalised work	Materials must be complete and well organised with a clear design
Use of educational specialities such as music or art as a transversal axis of learning	Materials must take into account multiple intelligences and different ways of presenting information
Use of an alternative curriculum system with a common curricular basis and topics as a reference	Inclusion of manipulative elements (in accordance with the Montessori Theory)
Curricular flexibility. Materials not divided by courses, but by abilities and skills. Inclusion of different levels of complexity.	Focused on problem-based and discovery-based learning
Inclusion of guides or bases of multilevel orientation	Creative and motivating
Flexible and open materials, adapted to multigrade classrooms and rural contexts	Contextualised materials. Inclusion of the territorial dimension and referents of the rural environment.
Fostering individual or group autonomy and cooperative work	Effective and specific in relation to the multigrade methodology

Teaching materials must allow interdisciplinary and globalised work, in line with the active methodology used in multigrade didactics, integrating the specialities in a transversal way. Their content must be adapted to the different age groups, taking into account different levels of complexity, making the curriculum more flexible and adapting flexibly to the context and to each multigrade classroom, including elements of the territorial dimension. The materials should allow for individual and cooperative work among students, and should therefore be well organised in order to understand how to work at any given moment. They must consider not only the diversity of ages, but also the diversity of learning styles, presenting information in different ways and therefore providing manipulative elements, as well. Finally, they should be motivating for students and encourage creative thinking and discovery learning, such as problem-based learning. We will briefly consider those materials that are related to the methodologies used in multigrade teaching and that are effective in their context of use.

In addition to these criteria, it is necessary to highlight the important role of teachers in the adaptation of materials according to their specific context and the multigrade classroom (Bustos, 2007; Santos, 2011). Some criteria are included that are applicable not only to rural schools, but could also be taken as a point of reference for the creation of resources and materials from an inclusive perspective; for example, the fact that materials should be flexible and open. It should, however, be emphasised that these aspects are especially relevant in the case of this type of school, including the complexity of the treatment of multigrade classrooms and the inclusion of the rural environment.

Most of the materials would be applicable to creating or adapting curricular resources, where the textbook, whether digital or printed, would come in, as well as other multigrade teaching materials, such as including different levels of complexity of the activities. The research highlights the fact that there is a lack of materials that are suitable for multigrade classrooms (Torres et al., 2019) and that fulfil the majority of these criteria as well as being deliberative, that is, materials need to be revised and adjusted by the teachers of rural schools, thus guaranteeing their effectiveness in every classroom and context (García Prieto, 2015).

Conclusions

The main focus of the present study was to review the concept of multigrade resources, that is, the types of teaching materials, specifically textbooks, used by teachers in rural multigrade schools, to identify opportunities and needs with respect to all teaching materials, and to identify criteria for multigrade teachers to be pedagogically meaningful. Papers were reviewed identifying the main findings that could help determine the current situation of multigrade textbooks and materials in rural education. It is important to highlight the fact that all of the research objectives have been achieved to a high degree, which has enabled the central research question to be answered. Through the systematic review it has been possible to determine the main characteristics of the works related to teaching resources in multigrade classrooms. English, Spanish and Portuguese stand out as the main languages in relation to the predominant countries in the scientific production on the subject. In addition, most of the articles reviewed come from peer-reviewed research journals in the field of social sciences. Most of the studies selected for the systematic review are qualitative in nature, demonstrating the need for more studies that provide quantitative, mixed method data in addition to qualitative data in order to expand knowledge on the topic. Furthermore, there is still a lack of research focused on multigrade teaching materials, and the present review contributes to summarising and expanding knowledge, as well as putting forward proposals for educational improvement. The analysis of the results of the studies that form part of the systematic review shows that there is a lack of a consensual definition of the concept of didactic material in the context of rural multigrade classrooms. This research has made it possible to analyse the concept and the state of play of multigrade teaching materials, proposing a definition and thus establishing a basis for responding to the difficulties that teachers have in understanding the nature of these resources and being able to create them. It is necessary to review education policies as well as the initial and ongoing training of teachers to enable them to face the challenges of diversified attention in multigrade classrooms and the selection, adaptation or creation of didactic resources. Many publications refer to multigrade resources as textbooks, which shows that, despite the efforts of teachers, working with printed or digital textbooks is not adequate for multigrade classrooms with different levels and ages of students. There is a need to revise existing materials and textbooks so that they are adapted to the context and to multigrade teaching. The importance of considering digital teaching materials should also be emphasised, as they allow for greater adaptability, flexibility and interactivity compared to print materials. Digital materials are particularly important in rural territories in order to open the local context to the global territory. In addition, the present study highlights the need for future research on the topic and for networking among teachers and sharing existing multigrade materials through digital platforms. Moreover, the results show that improvements in teacher training should also be implemented, taking into account the specific configuration of rural schools and multigrade work, due to the difficulties teachers encounter in developing their own teaching materials and applying them to multigrade teaching. It is necessary to foster the creation or adaptation of textbooks or other teaching materials to facilitate and promote multigrade work in the classroom. This should be done in cooperation with the publishing industry and education policy in general, taking into account the existing multigrade school model and the educational reality of classrooms through teachers. Finally, the present research enables us to establish some criteria that should be considered in the creation of textbooks and other curricular materials, allowing their adaptation to multigrade classrooms and promoting the inclusion of all students. Knowledge of these criteria and the considerations involved can help teachers to improve their teaching and the adaptation of teaching materials. The present research project therefore contributes to reflecting on the state of play of textbooks and teaching materials, and to envisioning possible challenges and solutions to improve their creation, thus facilitating multigrade classroom teaching and situated learning. In addition, it contributes to the necessary discussion on whether materials should be adapted to contexts and be more flexible in order to take into account other realities.

Acknowledgements

This paper is part of a doctoral thesis of the Educational Innovation and Intervention programme of the University of Vic-Central University of Catalonia (UVic-UCC). The research has been made possible thanks to the UVic-UCC Predoctoral Research Grant.

References

Afzalnia, M. (2005). *Designing and familiarity with the centers of learning materials and resources.* Samt.

Block, D., Carrillo, J., & Reséndiz, L. (2017). Una clase de matemáticas sobre problemas de aplicación, en una escuela multigrado unitaria. Un estudio de caso [A mathematics lesson on application problems in a multigrade unitary school. A case study]. *Educación matemática*, *29*(2), 99–123.

Boix, R., & Bustos, A. (2014). La enseñanza en las aulas multigrado: Una aproximación a las actividades escolares y los recursos didácticos desde la perspectiva del profesorado [Teaching in multigrade classrooms: An approach to school activities and teaching resources from a teacher's perspective]. *Revista Iberoamericana de Evaluación Educativa, 7*(3), 29–43.

Boix, R., & Domingo-Peñafiel, L. (2019). Escuela graduada y escuela multigrado: una dicotomía a superar en el siglo XXI [Graded school and multigrade school: A dichotomy to overcome in the 21st century]. *Cuadernos de pedagogía*, *496*, 84–89.

Brown, B. (2010). Multigrade teaching. *A Review of Issues, Trends and Practices: Implications for*

Teacher Education in South Africa. Centre for Education Policy Development (CEPD). Johannesburg, South Africa.

Bustos, A. (2007). Enseñar en la escuela rural aprendiendo a hacerlo. Evolución de la identidad profesional en las aulas multigrado [Teaching in rural schools by learning to teach. Evolution of professional identity in multi-grade classrooms]. *Profesorado. Revista de currículum y formación de profesorado*, *11*(3), 1–26.

Bustos, A. (2009). Valoraciones del profesorado de escuela rural sobre el entorno presente [Rural school teachers' assessments of the present environment]. *Revista Iberoamericana de Educación*, *48*(6), 1–11.

Bustos, A. (2014). La didáctica multigrado y las aulas rurales: perspectivas y datos para su análisis [Multigrade didactics and rural classrooms: Perspectives and data for analysis]. *Innovación educativa*, *24*. https://doi.org/10.15304/ie.24.1994

Carrete-Marín, N., & Domingo-Peñafiel, L. (2021). Los recursos tecnológicos en las aulas multigrado de la escuela rural: Una revisión sistemática [Technological resources in rural school multigrade classrooms: A systematic review]. Revista Brasilera de Educaçao do Campo, *6*, e13452, 1–31. https://doi.org/10.20873/uft.rbec.e13452

Cardona, L. A. (2020). *Influencia del uso de textos multimodales en las prácticas de escritura en los estudiantes de cuarto y quinto del Centro Educativo Rural Jesús Aníbal Gómez, sede Patio Bonito* [Influence of the use of multimodal texts on the writing practices of fourth- and fifth-grade students at the Jesús Aníbal Gómez Rural Education Centre, Patio Bonito campus]. http://bibliotecadigital.udea.edu.co/handle/10495/15859

Chaves. L., García, J., & Alvarado, R. (2013). Las escuelas unidocentes en Costa Rica: fortalezas y limitaciones [One-teacher schools in Costa Rica: Strengths and limitations. *Revista Educación*, *37*(1), 1–27.

Coaquira, P. (2018). Mejorar uso de recursos y materiales educativos en el desarrollo de sesiones de aprendizaje [Improving the use of educational resources and materials in the development of learning sessions]. http://repositorio.usil.edu.pe/handle/USIL/8286

Coladarci, T. (2007). Improving the yield of rural education research: An editor's swan song. *Journal on Research in Rural Education, 22*(3), 1–9.

Cornish, L. (2006). Reaching EFA through multi-grade teaching: Issues, contexts and practices. Kardoorair Press.

Da Silva, C., & Cardoso Filho, I. D. (2017). Rural education and teaching material: An analysis of history textbooks. *Revista Brasileira De Educação Do Campo*, *3*(1), 76–103. https://doi.org/10.20873/uft.2525-4863.2018v3n1p76

Dávila, M. C. (2018). *Eco invernadero escolar didáctico como recurso para lograr aprendizajes significativos* [Didactic school eco-greenhouse as a resource to achieve significant learning]. University San Ignacio de Loyola. http://repositorio.usil.edu.pe/handle/USIL/7304

Devetak, I., & Ferk Savec, V. (2020). Chemical education research in Slovenia after 1991: A systematic review. *CEPS Journal*, *10*(1), 9–36. https://doi.org/10.26529/cepsj.818

Domingo-Peñafiel, L., & Boix, R. (2015). What can be learned from Spanish rural schools? Conclusions from an international project. *International Journal of Educational Research, 74,* 114–126. https://doi.org/10.1016/j.ijer.2015.11.002

Drinkwater, M. (2002). Improve your classroom climate: A workbook for Bed Honours students. Keurkopie.

Fargas-Malet, M., & Bagley, C. (2021). Is small beautiful? A scoping review of 21st-century research on small rural schools in Europe. *European Educational Research Journal*, 14749041211022202. https://doi.org/10.1177/14749041211022202

Ferguson, D., & Jeanchild, L. (2004). Cómo poner en práctica las decisiones curriculares. In S. Stainback, and W. Stainback, (Eds., *Aulas inclusivas: un nuevo modo de enfocar i vivir el currículo* [Inclusive classrooms: A new way of approaching and experiencing the curriculum] (pp. 21–35). Narcea.

Fernández Palop, M. P., & Caballero García, P. A. (2017). El libro de texto como objeto de estudio y recurso didáctico para el aprendizaje: fortalezas y debilidades [The textbook as an object of study and a didactic resource for learning: Strengths and weaknesses]. *Revista Electrónica Interuniversitaria de Formación del Profesorado, 20*(1), 201–217.

Forero, A., Alemán, L., & Gómez. G. (2016). Experiencias de los docentes en la implementación de las TIC en escuelas rurales multigrado [Teachers' experiences of ICT implementation in rural multigrade schools]. *Edmetic*, *5*(1), 52–72.

Fraser, J., Loubser, C., & Van Rooy, M. (1993). Didactics for the undergraduate student (2nd ed.). Butterworths.

Garcia, L., & Machado, I. (2017). Recursos didáticos utilizados por professores de escolas rurais de um município mato-grossense nas décadas de 1980 e 1990 [Didactic resources used by teachers in rural schools in a municipality of Mato Grosso in the 1980s and 1990s]. *Educação & Formação*, *2*(3), 139–152.

García Prieto, F. J. (2015). *Escuela, medio rural y diversidad cultural en un contexto global: currículum, materiales didácticos y práctica docente de Conocimiento del Medio: situación, límites y posibilidades en centros onubenses* [School, the rural environment and cultural diversity in a global context: Curriculum, didactic materials and teaching practice in Environmental Knowledge: Situation, limits and possibilities in schools in Huelva] (Doctoral thesis). University of Huelva. http://rabida.uhu.es/dspace/handle/10272/11440

García Prieto, F. J., & Pozuelos Estrada, F. J. (2017). El currículum integrado: los proyectos de trabajo como propuesta global para una escuela rural alternativa [The integrated curriculum: Project work as a global proposal for an alternative rural school. *Aula abierta, 45*(1), 7–14.

García Prieto, F. J., Pozuelos F. J., & Álvarez Álvarez, C. (2017). Uso de los libros de texto en la educación rural en España [Use of textbooks in rural education in Spain]. *Sinéctica*, (49), 1–18.

Genc, Z. (2016). The opinions of primary school teachers' candidates towards material preparation and usage. *Cypriot Journal of Educational Sciences*, *11* (2), 70–76.

Giraldo, J., Rosales, A., & Lozano, Y. (2018). 21st Century competencies: How to develop them through the use of video games in a multigrade context? *Panorama*, *12*(23), 7–18.

Guerra, M., Leyva, Y., & Conzuelo, S. (2019). Factores contextuales que afectan la práctica docente desde la perspectiva de docentes de México [Contextual factors that affect teaching practice from the perspective of teachers in Mexico]. *Publicaciones: Facultad de Educación y Humanidades del Campus de Melilla*, *49*(1), 137–149.

Guevara, M. A., Maltez, M. S., & Martínez, G. E. (2012). *Aplicación de recursos didácticos en la clase de Ciencias Naturales en el sexto grado de educación primaria en la Escuela Rural Los Laureles del turno matutino del Departamento de Chinandega Municipio de Chichigalpa 2011* [Application of didactic resources in the Natural Sciences class in the sixth grade of primary education in the Los Laureles Rural School, morning shift, Department of Chinandega, Municipality of Chichigalpa 2011]. National Autonomous University of Nicaragua.
http://riul.unanleon.edu.ni:8080/jspui/bitstream/123456789/5638/1/220757.pdf

Ivić, I. (2019). Printed and digital media: Printed and digital textbooks. *CEPS Journal*, *9*(3), 25–49. https://doi.org/10.26529/cepsj.694

Joubert, J. (2010). Multi-grade teaching in South Africa. *Commonwealth Education Partnership*, 58–62. https://www.cedol.org/wp-content/uploads/2012/02/58-62-2010.pdf

Juárez, D. (2012). Educación rural en Finlandia: experiencias para México [Rural education in Finland: Experiences for Mexico]. *Revista CPU-e*, 15, 140–154.

Juela, L. F., & Matailo, B. R. (2015) *El auto aprendizaje como metodología alternativa en escuelas multigrado* [Self-learning as an alternative methodology in multi-grade schools] [Doctoral thesis] University of Cuenca, Ecuador. https://repositorioslatinoamericanos.uchile.cl/handle/2250/1130732

Juvane, V. (2005). Redefining the role of multi-grade teaching. In *Working document prepared for the Ministerial seminar on education for rural people in Africa: Policy lessons, options and priorities. Addis Ababa, Ethiopia* (pp. 7–9).

Little, A. (2005). Learning and teaching in multigrade settings. UNESCO.
http://multigrade.ioe.ac.uk/fulltext/fulltextLittle.pdf

López, I. (2011). *Aprendizaje del Castellano Utilizando en Cuaderno de Trabajo T'ika Básico de Niños y Niñas del 2° y 3° Grado* [Spanish language learning using the T'ika Básico Workbook for 2nd and 3rd grade boys and girls] (Doctoral thesis). National University of Altiplano Puno.
http://repositorio.unap.edu.pe/handle/UNAP/159

Mathot, G. B. (2001). A handbook for teachers of multigrade classes (Vol. 1). UNESCO.

McEwan, P. (2008). Evaluating multigrade school reform in Latin America. *Comparative Education*, *44*(4), 465–483. https://doi.org/10.1080/03050060802481504

Miranda, L. A., Angulo-Hernández, L. M., & Román-López, G. (2018). El programa perfiles, dinámicas y desafíos de la educación costarricense: una propuesta para la innovación pedagógica y la producción de materiales y recursos didácticos tecnológicos en la Universidad Nacional, Costa Rica [The programme profiles, dynamics and challenges of Costa Rican education: A proposal for pedagogical innovation and the production of technological teaching materials and resources at the National University, Costa Rica]. *Revista Electrónica Educare, 22*(1), 201–224.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS medicine*, *6*(7), 264–269.

Msimanga, M. R. (2019). Managing the use of resources in multi-grade classrooms. *South African Journal of Education*, *39*(3), 1–9.

OECD. (2019). *Does attending a rural school make a difference in how and what you learn? In PISA in Focus* (Issue 94). OECD. https://doi.org/https://doi.org/10.1787/d076ecc3-en

Petrie, G. M., & Darragh, J. J. (2018). Educative curriculum materials: A promising option for independent professional development. *English Teaching Forum, 56*(4), 2–15.

Pozuelos, F. J., & Travé, G. (2004). Aprender investigando, investigar para aprender: el punto de vista de los futuros docentes [Learning by researching, researching to learn: The perspective of future teachers]. *Investigación en la Escuela, 54,* 5–25.

Rodríguez, Y., & Saavedra, E. (2018). Las TIC: Una oportunidad para fortalecer procesos de enseñanza en el marco del modelo escuela nueva [ICTs: An opportunity to strengthen teaching processes in the framework of the new school model]. *Espacios, 39*(49), 1–11.

Rahim, B. (2018). School resources and learning outcomes: An empirical study of rural public schools. *Excellence in Education Journal*, *7*(2), 3–41.

Rivas-Martínez, S., Navarro, G., Penas, A., & Costa, M. (2011). Biogeographic map of South America. A preliminary survey. *International Journal of Geobotanical Research, 1*(1), 21–40.

Sağ, R. (2009). Birleştirilmiş siniflarda öğretmen olmak [Becoming a teacher in multigrade classes]. *Elektronik Sosyal Bilimler Dergisi. Electronic Journal of Social Sciences*, *8*(28), 20–39.

San Pedro, M. B., & López, I. (2017). El profesorado de Ciencias Sociales en contextos rurales de Asturias: concepciones y uso de recursos educativos [Social Science teachers in rural contexts in Asturias: Conceptions and use of educational resources]. *Aula Abierta*, *45*(1), 33–40. https://doi.org/10.17811/rifie.45.1.2017.33-40

Sánchez-Meca, J., & Botella, J. (2010). Revisiones sistemáticas y meta-análisis: herramientas para la práctica profesional [Systematic reviews and meta-analyses: Tools for professional practice]. *Papeles del Psicólogo, 31*(1), 7–17.

Santamaría Pérez, J. G. (2021). Consideraciones sobre los textos escolares para la escuela rural: una revisión bibliográfica [Considerations on textbooks for rural schools: A literature review]. *Revista Espaço Do Currículo*, *14*(2), 1–20.

Santos, L. (2011). Aulas multigrado y circulación de los saberes: especificidades didácticas de la escuela rural [Multigrade classrooms and the circulation of knowledge: Didactic specificities of the rural school]. *Profesorado, Revista de currículum y formación del profesorado*, *15*(2), 71–91.

Severiano, J. (2013). *El uso de la Tecnología en el Aula Multigrado* [The use of technology in the multigrade classroom] (Doctoral thesis). Tecvirtual University- Tecnológico de Monterrey. https://repositorio.tec.mx/handle/11285/619532

Silva, N dos S., Silva, J. D. F. A., & Costa, R. T. (2020). The game in multigrade classes at rural schools: Aid to the adoption of essential rules to life. *Educação & Formação, 5*(14), 93–114.

Smith, G. (2007). Grounding learning in place. *World Watch, 20*(2), 20–24.

Smith, G. (2002). Place-based education: Learning to be where we are. *Phi Delta Kappan*, *83*(3), 584–594. https://doi.org/10.1177/003172170208300806

Smith, V., Devane, D., Begley, C. M., & Clarke, M. (2011). Methodology in conducting a systematic review of systematic reviews of healthcare interventions. *BMC Medical Research Methodology, 11*(1), 15–16.

Terigi, F. (2009). Organización de la enseñanza de los plurigrados de las escuelas rurales [Organisation of the teaching of multigrade rural schools] (Doctoral thesis). Latin American Faculty of Social Sciences (FLACSO). https://repositorio.flacsoandes.edu.ec/handle/10469/1266

Tomlinson, C. A. (2001). *Aula diversificada: dar respuesta a las necesidades de todos los alumnos* [The diversified classroom: Meeting the needs of all students]. Octaedro.

Torres, G. J., Castillo, E. E., & Torres, D. C. (2019) *Fortalecimiento de la enseñanza aprendizaje en el Centro Educativo Rural cumainde el paso de Barbacoas Nariño* [Strengthening teaching and learning in the rural educational centre cumainde el paso in Barbacoas Nariño]. University Santo Tomás. https://repository.usta.edu.co/handle/11634/16089

Valderrama, E., Winkler, C., & Schmidt, A. (2011). Utilizing multimedia capabilities of mobile phones to support teaching in schools in rural Panama. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 935–944). http://nixdell.com/classes/Tech-for-the-underserved/valderrama_chi2011.pdf

Biographical note

Núria Carrete-Marín, PhD student in the Educational Innovation and Intervention programme on the Faculty of Education, Translation, Sport and Psychology at the University of Vic-Central University of Catalonia (UVic-UCC). She is also a researcher of the GREUV - Educational Research Group at the UVic-UCC. Her main areas of research interest are: inclusion, rural education, multigrade teaching, the territorial dimension in rural schools and teaching resources and materials in rural education.

Laura Domingo-Peñafiel, PhD, is a lecturer in the field of Pedagogy on the Faculty of Education, Translation, Sports and Psychology at the University of Vic - Central University of Catalonia. She is also researcher of the GREUV - Educational Research Group at the UVic-UCC. Her main areas of research interests are: training teachers, inclusion, democratic and rural schools. It is focused on multigrade and place-based teaching in a rural education framework.

doi: 10.26529/cepsj.1285

Differences in the Requirements of Digital and Printed Mathematics Textbooks: Focus on Geometry Chapters

Dubravka Glasnović Gracin*[1] and Ana Krišto[2]

Textbooks have always played an important role in mathematics education. Textbook tasks are widely used by students, so it is important to examine their requirements in order to identify the opportunities students have to learn mathematics. Publishers now produce both printed and digital versions of textbooks. While the requirements of the tasks in printed textbooks have been well examined all over the world, the tasks in digital textbooks are yet to be analysed and systematically developed. The research presented in this paper encompasses the analysis and comparison of the tasks in the printed and digital versions of the same mathematics textbook set. The examined set covers Grades 1 to 4 of primary education in Croatia. The aim was to find what task requirements are predominant in the printed and the digital textbooks, and to determine whether these textbook versions provide a wide variety of task features. In addition, the features and capacities typical of digital tasks, such as interactivity and dynamics, are examined. These task features are particularly important in geometry education for comprehending visual and dynamic geometrical objects and relations. The results show that both the printed and the digital textbook tasks have traditional requirements, with an emphasis on closed answer forms. Moreover, the new opportunities afforded by digital tasks are not realised. These findings reveal the potential of digital tasks as a new area to be explored and developed.

Keywords: textbook tasks, requirements, printed textbook, digital textbook, geometry

1 *Corresponding Author. Faculty of Teacher Education, University of Zagreb, Croatia; dubravka.glasnovic@ufzg.hr.

2 Primary School Petar Preradović Zagreb, Croatia.

Razlike v zahtevah digitalnih in tiskanih matematičnih učbenikov: poudarek na poglavjih o geometriji

Dubravka Glasnović Gracin in Ana Krišto

Učbeniki so imeli od nekdaj pomembno vlogo pri poučevanju matematike. Naloge iz njih učenci pogosto uporabljajo, zato je bistveno, da preučimo njihove zahteve in prepoznamo priložnosti, ki jih imajo učenci za učenje matematike. Založniki zdaj izdajajo tiskane in digitalne različice učbenikov. Medtem ko so bile zahteve nalog tiskanih učbenikov dodobra analizirane po vsem svetu, morajo biti zahteve digitalnih še preiskane in sistematično razvite. Raziskava, predstavljena v tem prispevku, vključuje analizo in primerjavo nalog tiskane in digitalne različice istega učbeniškega kompleta za matematiko. Izbrano gradivo zajema pregled učbenikov od 1. do 4. razreda osnovnošolskega izobraževanja na Hrvaškem. Namen je bil ugotoviti, katere naloge so prevladujoče v tiskanih in digitalnih učbenikih, ob tem pa določiti, ali te različice ponujajo široko množico funkcij. Poleg tega so bile preučene značilnosti in zmožnosti, ki so značilne za digitalne naloge, kot sta na primer interaktivnosti in dinamičnost. Te so še posebej pomembne pri poučevanju geometrije za razumevanje vizualnih in dinamičnih geometrijskih likov in odnosov. Rezultati kažejo, da imajo naloge tiskanih in digitalnih učbenikov tradicionalne zahteve s poudarkom na odgovorih zaprtega tipa. Mimogrede, nove priložnosti, ki jih sicer ponujajo digitalne naloge, niso uresničene. Te ugotovitve razkrivajo potencialne digitalne naloge kot novo področje, ki bi ga bilo treba raziskati in razviti.

Ključne besede: učbeniške naloge, zahteve, tiskani učbeniki, digitalni učbeniki, geometrija

Introduction

Textbooks have always played an important role in mathematics education. Textbooks that are consistent with the curriculum facilitate teachers' work and guarantee mathematical knowledge and exercises for students (Johansson, 2006; Love & Pimm, 1996). They provide security and convenience for teachers, students and parents. Research has shown that textbooks are mainly used by students as a source of tasks, particularly practice exercises (Pepin & Haggarty, 2001). A task is considered as a request for initiating student activity (Markovac, 2001), and working on tasks is the most common student activity in mathematics education (Kurnik, 2000). Therefore, textbook tasks provide opportunities for learning mathematics (Sullivan et al., 2013), they "potentially influence and structure the way students think and can serve to limit or to broaden their views of the subject matter with which they are engaged" (Henningsen & Stein, 1997). Consequently, it is important to examine their features and requirements.

In the past decade, new digital curriculum materials and e-textbooks have been developed. They include a number of new features that are not found in traditional resources. Digital curriculum materials can be transformative, with many dynamic and interactive possibilities; they can easily provide customised instruction and formative assessment, and often have links to multimedia resources and viral communities (Choppin et al., 2014). Choppin et al. (2014) provided a typology of digital curriculum materials concerning: (a) students' interactions with these resources; (b) curriculum use and adaptation; and (c) assessment. Similarly, Pepin et al. (2017) conceptualised three features of digital resources in terms of: instruction; assessment and reporting; and management. The authors regard digital curriculum resources as opportunities for changing instruction because of their potential "to provide stimulating and meaningful learning experiences for students, and motivating opportunities for teacher collaborative learning, including the enhancement of teachers' design capacity" (Pepin et al., 2017, p. 646). These new features refer to the whole learning space of digital curriculum materials (presentation spaces, navigation spaces, platforms, etc.), also encompassing the *tasks* in digital textbooks. Producing digital tasks has become a real challenge for textbook authors because e-tasks "can extend and amplify pedagogical features present in non-digital environments" (Leung & Baccaligni-Franck, 2017, p. ix). These challenges include enhancing interactivity and customisability through tasks, feedback and formative assessment (Choppin et al., 2014; Rezat, 2021), thus influencing educational processes and bringing "new educational dynamics" (Pepin et al., 2017, p. 646).

The representational and visual potential of e-textbooks may help students better understand mathematical ideas (Usiskin, 2018). This is particularly important for geometry education, which is strongly connected to visualisation. Interactions and dynamics in digital tasks may bring new opportunities to geometry education, which has been reduced in quantity within mathematics curricula in recent times (Glasnović Gracin & Kuzle, 2018; Kuzle & Glasnović Gracin, 2020; Mamanna & Villani, 1998). Contemporary thinking on geometry education involves organising it around certain fundamental ideas (Mamanna & Villani, 1998). In line with this, Wittman (1999) proposed the organisation of school geometry around seven fundamental ideas: geometric forms and their construction, operations with forms, coordinates, measurement, geometric patterns, geometric forms in the environment, and geometrisation.

In 2019, within a comprehensive education reform in Croatia, a new curriculum for mathematics was published (Ministry of Science and Education [MZO], 2019) based on student outcomes, real-world orientation and a problem-solving approach. The support of high-quality resources, including textbooks and IT resources, was identified as one of the key reform factors. Geometry content within the new mathematics curriculum contains the following concepts. First-grade geometry encompasses basic 2D and 3D shapes and patterns, straight and curved surfaces and lines, and points. Second-grade content refers to the line segment and its measurement, its edge points, length units, the sides of the square, quadrilateral, and triangle, and the edges of geometric solids. Third-grade geometry contains the line, the ray, the line segment and its measurement, length units, circumference, intersecting and parallel lines, and the circle. Fourth-grade geometry refers to angles, triangle types, the circle, radius, measuring area, units of area, and the square grid. This content, which is incorporated in learning outcomes, is surely reflected in textbook tasks.

With these considerations in mind, the aim of the present study is to analyse tasks within the geometry chapters in both the printed and digital versions of the same mathematics textbook set. In comparison to the research on tasks in printed textbooks, analysis of tasks in digital mathematics textbooks is rare. Therefore, it was important to develop an instrument for task analysis in digital mathematics textbooks.

Theoretical background

Five-dimensional framework for analysing textbook tasks

In order to analyse textbook tasks, Glasnović Gracin (2018) developed a five-dimensional instrument consisting of the following categories: mathematical content, mathematical activity, complexity, answer type and context. The basis for this instrument is a combination developed from two theoretical sources: Austrian standards for mathematics (Institut für Didaktik der Mathematik [IDM], 2007) and the framework provided by Zhu and Fan (2006).

Content. The content requirements refer to "finding out what mathematical knowledge a student should possess in order to solve a particular textbook task" (Glasnović Gracin, 2018, p. 1009). In primary grades, it encompasses arithmetic, geometry, measurement, statistics and probability, and patterns. Since the present study refers to geometry, within the content dimension we took the framework based on Wittmann's (1999) aforementioned seven fundamental ideas. These ideas were further developed by Kuzle and Glasnović Gracin (2020) as follows. (1) Geometric forms and their construction refers to different shapes and forms (e.g., points, lines, 2- and 3-dimensional shapes), which can be constructed or produced in a variety of ways. (2) Operations with forms refers to different geometry operations, such as translation, rotation, mirroring, dilation, etc. (3) Coordinates and spatial visualisation refers to describing locations using coordinates, and may also encompass positional relationships and spatial visualisation. (4) Measurement means describing geometric forms using units of length, area or volume, and also contains angle measuring and formulae for perimeter, area and volume. (5) Geometric patterns are patterns in which geometric objects are used. (6) Geometric forms in the environment refers to real world objects described with the help of geometric forms. (7) Geometrisation means mathematical non-geometric properties translated into the language of geometry (e.g., triangular numbers). These fundamental ideas may overlap in tasks, i.e., a task that contains one fundamental idea may refer to another one, too.

Activities. Mathematical activities in tasks refer to the question of *what* should be done in a particular task; for example, does the textbook task require the activity of computation or maybe drawing a figure, or giving a mathematical explanation (Glasnović Gracin, 2018)? This field is divided into: representations and modelling; calculation and operation; interpretation; and argumentation and reasoning (IDM, 2007). These activities are not hierarchically ordered. (1) Representations refer to transmissions of the given mathematical data into another type of presentation, while modelling means recognising relevant mathematical relationships from the given situation and representing the same problem in a mathematical mode (symbolic, graphical, etc.). (2) Calculation refers to conducting elementary computations. Operation concerns the conducting of computational or constructional steps. (3) Interpretation concerns recognising relations and relevant data given through mathematical representations and their understanding in the given context. (4) Argumentation means the description of mathematical aspects that speak pro or contra a particular decision. Reasoning is the sequence of true arguments that lead to a conclusion.

Complexity. Tasks may be put on different levels of cognitive complexity (e.g., Organisation for Economic Co-operation and Development, 2003; Zhu & Fan, 2006). IDM (2007) distinguished between reproduction, connections and reflection. Some items are simple and require the direct application of basic knowledge and skills, while others are more complex and require constructing and dealing with connections between a variety of concepts and rules in order to solve the problem. Finally, some items require reflecting on ideas that are not directly apparent from the posed problem (IDM, 2007).

Answer type. Based on a study by Zhu and Fan (2006), Glasnović Gracin (2018) distinguished between closed answer form, open answer form and multiple choice. Open-ended tasks refer to tasks with several or many correct answers, while closed-ended tasks have only one answer and can be easily validated as correct or incorrect. Multiple choice tasks provide a limited number of response options.

Context. The contextual features of a task refer to the extent of real-world experiences that are present in a particular textbook task. Glasnović Gracin (2018) distinguishes between tasks with intra-mathematical situations, tasks with realistic context and tasks with authentic context. Within this categorisation, intra-mathematical problems are unrelated to the real world, authentic tasks contain genuine real-life situations and data, and realistic tasks have contexts that imitate authentic situations (using fictive names and data).

Categorisation of digital task features

We conceptualise here several categories of digital task features given emphasis in the literature, such as interactivity, dynamics, personalisation, response form, feedback and cooperation.

Types of digital textbooks and their interactivity. Usiskin (2018) distinguished between three types of e-textbook forms: minimal, hybrid and exclusive digital textbooks. The minimal platform is simply a digitalised version of a printed textbook (e.g., pdf version) with some additional links. The hybrid platform refers to both the paper and electronic form with built-in features to provide links with video explanations, hints, additional exercises, software for manipulating objects, etc. The exclusively digital platform is designed wholly as digital material with social media interactive objects that are linked and can be combined. A similar classification is given by Pepin et al. (2016), who distinguish between the integrative, evolving, and interactive e-textbook. These classifications imply different task features; for example, whether the student should solve the task in his/her notebook or in the digital answer space provided.

Dynamic diagrams for exploring mathematics. Unlike static printed textbooks, digital textbooks provide dynamics that may help in exploring and better understanding mathematical concepts and dependences. Dynamic representations clearly show the process of transformation of figures, the change of function values as the domain values change, the effects of parameters, etc. (Usiskin, 2018).

Personalisation in learning. Digital materials may be customised to meet the individual student's learning needs as s/he progresses through the mathematical topics. The programme may select tasks based on the user's performance in assessment (Pepin et al., 2017). In addition, digital resources have the possibility to generate additional tasks. Such adaptive learning software may be used in diagnostic tests and in formative evaluations of the student's progress (Usiskin, 2018). In this way, the learning path refers to the evident non-linear way of working through the textbook tasks.

Task form. Digital tasks and the associated responses required from students can be presented in a number of forms. Some tasks are given in multiple-choice forms, others in a fill-in-the-blank form, etc. (Pepin et al., 2017).

Feedback. Digital resources and tasks have the potential to provide feedback and performance data to students and other stakeholders (Choppin et al., 2014). Feedback is considered as a powerful and influential factor in learning and achievement (Cohen, 1985; Hattie & Timperley, 2007). Schute (2008) provides a categorisation of different feedback types: no feedback, verification (right/wrong), correct response, try again, error flagging, elaborated (why the answer is/isn't correct), attribute isolation, topic contingent, response contingent, hints, misconceptions, and informative tutoring.

Cooperation. Digital resources have the possibility for the user to share their workspace with others and to enable collaboration (Pepin et al., 2017). This may be a very important feature for improving the learning process. The e-textbook tasks may foster this feature.

These features of digital textbooks raise the question of whether this potential is actually used in the current digital tasks provided for students.

Research questions

The aim of the present study was to determine which task features predominate in printed and digital textbooks, and whether the two versions of the textbook provide a wide variety of task features. Therefore, the following research questions were posed:

1. What are the differences between task requirements in digital and printed textbooks?
2. What additional task features are presented in digital geometry tasks?

Method

Sampling. The study presented in this paper refers to an empirical study using textual analysis. It encompassed the paper and digital versions of the most frequently used mathematics textbook set in Croatia for Grades 1, 2, 3 and 4 (Miklec et al., 2021a, 2021b, 2021c, 2021d). The analysis referred to all of the textbook tasks provided for practice and revision in the geometry chapters. Altogether, the analysis encompassed 600 textbook geometry tasks: 267 from the printed textbooks and 333 from the digital textbooks. The study took into consideration only the digital tasks in which the student should give their answer in the digital space provided, because such tasks contain at least some of the aforementioned digital task features.

Instrument for textbook analysis. Based on the theoretical background, two instruments for task analysis were established. The first refers to the aforementioned five-dimensional framework for analysing textbook tasks (Table 1). This instrument was developed from the framework by Glasnović Gracin (2018), focusing on geometry tasks in primary grades. The content strand focused on geometry; specifically, on Wittmann's (1999) fundamental ideas, which were further developed by Kuzle and Glasnović Gracin (2020). For the purposes of this study, the codes F1 to F7 (Kuzle & Glasnović Gracin, 2020) were slightly modified further for primary education (Table 1). The mathematical activities (Glasnović Gracin, 2018) were developed for this study by separating calculation and operation, and adding estimation with measurement and comparison. In this way, we sought to get a better insight into the activities required in geometry tasks of primary grades. As in Glasnović Gracin (2018), code H is used for mathematical activities (German: Handlungen), K for complexity (Komplexität), A for answer form and C for context.

Table 1
Five-dimensional framework for geometry task analysis, developed for the study

Dimension	Details and codes
Content (fundamental ideas of geometry)	Points (F1.0), 1-dim objects (F1.1), 2-dim objects (F1.2), 3-dim objects (F1.3) Operations with forms (F2) Positional relationships and spatial visualisation (F3) Measurement (F4) Geometric patterns (F5) Forms in the environment (F6) Geometrisation (F7)
Mathematical activity	Representations (H1) Calculation (H2) Operation (H3) Interpretation (H4) Argumentation (H5) Estimation and measuring (H6) Comparison (H7)
Complexity level	Direct application of rules and definitions (K1) Simpler connections (K2) Complex connections and reflection (K3)
Answer form	Closed (A1) Open (A2)
Context	Intra-mathematical context (C1) Realistic context (C2) Authentic context (C3)

The second instrument was developed according to the literature review on the categorisation of digital task features. Each category and its subcategory is defined by a code given in Table 2. *Dynamic diagrams* refer to tasks that require exploring geometry properties using dynamics. *Personalisation* in an e-task means that the exercise reflects a concern for the student's individual progression in learning based on his/her performance. *Cooperation* means that the task provides the possibility for users to share their workspace with others in order to solve the task. According to Pepin et al. (2017), assessment can be incorporated into programs in different ways and forms, such as multiple-choice questions or fill-in-the-blank responses. For the purpose of the present study, we developed seven categories of *task form*, for the better categorisation of all of the examined tasks: multiple-choice questions, fill-in-the-blank responses, matching, true or false, put in order, other closed forms and open-ended questions. *Feedback* codes are established according to Schute's (2008) categorisation: no feedback refers to situations in which a student responds to the given question, but there is no way of knowing whether or not the answer is correct. Tasks with feedback are further divided into the following components. Verification feedback means returning a simple right/wrong response. Correct response means giving feedback just for correct answers. Try again refers to the possibility of trying until the correct answer is given, while error flagging highlights incorrect answers in the task, but does not offer the correct answer. Elaborated feedback means giving an explanation of why a specific response was correct or incorrect. Hints/cues/prompts are part of elaborated feedback: they guide students in the right direction by giving them examples or recommendations on what to do next, for instance, but the correct answer is not explicitly given.

Table 2

Framework developed for the analysis of digital tasks (special features)

Dimension	Details and codes
Dynamic diagrams for exploring mathematics	Yes – manipulating objects of dynamical geometry (D1) No (D2)
Personalisation in learning	Yes (P1) No (P2) What type of personalisation? ______________________
Task form	Multiple-choice questions (TF1) Fill-in-the-blank responses (TF2) Matching (TF3) True or false (TF4) Put in order (TF5) Other closed forms (TF6) Open-ended questions (TF7)
Feedback	No feedback (F1) Verification (F2) Correct response (IF1) Try again (IF2) Error flagging (IF3) Elaborated (IF4) Hints/cues/prompts (IF5) Other (IF6)
Cooperation	Yes (COO1) No (COO2)

Procedure and exemplary analysis. The analysis of tasks in the selected textbook set was conducted using the instruments given in Table 1 and Table 2: the printed textbooks were analysed according to the codes given in Table 1, and the digital tasks by the codes from tables 1 and 2. The five-dimensional framework (Table 1) was firstly applied to all exercise tasks in both the printed and the digital textbooks in order to find the (mathematical) features of these tasks. Altogether, 600 tasks were examined by this instrument: 267 in the printed textbooks and 333 in the digital textbooks. Each of the tasks was coded into the corresponding category. This analysis was followed by the application of the framework given in Table 2, which was applied to the 333 digital tasks. Both analyses referred to the qualitative textual analyses, because the meaning of the text led to the appropriate code. The accuracy and reliability of the coding in both instruments was ensured by checking task samples by both authors. Figure 1 presents a task from the digital textbook (translation: Match the picture to the term). In Table 3 and Table 4, the exemplary analysis of this task is given. The coded data were further analysed using quantitative methods; specifically, finding the relative frequencies of codes within a particular category.

Figure 1

An example of a task from the digital textbook

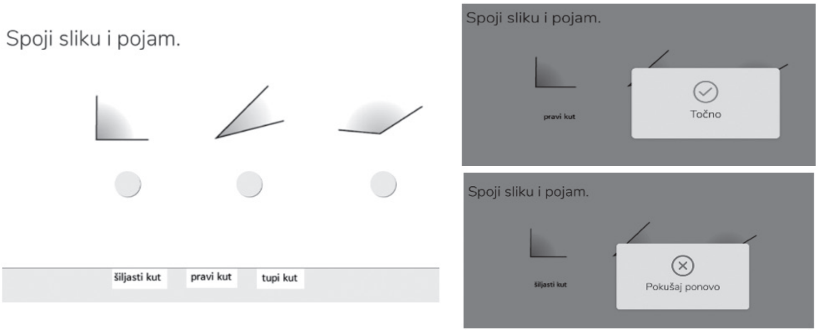


Table 3

Coding of the task in Figure 1 according to the framework given in Table 1

Category	Code	Description
Content	(F1.2)	2-dim objects
Mathematical activity	(H4)	Interpretation
Complexity level	(K1)	Direct application of rules and definitions
Answer form	(A1)	Closed
Context	(C1)	Intra-mathematical context

Table 4

Coding of the task in Figure 1 according to the framework given in Table 2

Category	Code	Description
Dynamic diagrams for exploring mathematics	(D2)	No
Personalisation in learning	(P2)	No
Task form	(TF3)	Matching
Feedback	(F1) (IF2)	Verification Try again
Cooperation	(COO2)	No

Results

The results section is divided into two parts. The first part, *Task features in printed and digital textbooks*, refers to the first research question, analysis of printed and digital tasks according to the codes given in Table 1, and their comparison. The second part, *Special features of digital tasks*, refers to the second research question and analysis of digital tasks according to the codes given in Table 2.

Task features in printed and digital textbooks

Printed textbooks. The results of task requirements in the printed textbooks are presented in Table 5. The overall findings indicate a lack of variety in task requirements.

In the *content* category, about a third of the analysed exercises referred to knowledge about 1-dimensional objects (straight and curved lines) and another third to 2-dimensional objects (F1.2). Tasks on geometric solids were underrepresented, with no such tasks at all in Grades 3 and 4. In line with this, examination of the curriculum (MZO, 2019) shows the dominance of two-dimensional objects and lines. Tasks that contain operations with geometric forms (F2), such as symmetry, and tasks with positional relationships (F3) and geometrisation requirements (F7) were also omitted altogether from the printed textbooks. These ideas are not highlighted in the geometry curriculum either. Measurement is present in about 20% of all of the geometry tasks in Grades 1–3, and 10% in Grade 4. However, length and area measurement are highlighted as important in the fourth-grade geometry curriculum (MZO, 2019). In line with the curriculum, geometric patterns (F5) are present only in the first grade, with a proportion of 19%. On the other hand, the curriculum highlights real situations (F6), while the findings from the textbooks show a lack of such tasks, particularly in Grades 3 and 4.

Table 5
Task requirements in the printed textbooks

Categories and codes		Grade 1 (n = 58)	Grade 2 (n = 36)	Grade 3 (n = 82)	Grade 4 (n = 91)	TOTAL (n = 267)
Content (fundamental ideas of geometry)	F1.0	8.62%	8.33%	0.00%	0.00%	3.00%
	F1.1	10.34%	41.67%	65.85%	20.88%	35.21%
	F1.2	15.52%	19.44%	17.07%	74.73%	36.70%
	F1.3	12.07%	8.33%	0.00%	0.00%	3.75%
	F2	0.00%	0.00%	0.00%	0.00%	0.00%
	F3	0.00%	0.00%	0.00%	0.00%	0.00%
	F4	20.69%	19.44%	19.51%	9.89%	16.48%
	F5	18.97%	0.00%	0.00%	0.00%	4.12%
	F6	15.52%	11.11%	1.22%	0.00%	5.24%
	F7	0.00%	0.00%	0.00%	0.00%	0.00%
Mathematical activity	H1	53.45%	50.00%	48.78%	56.04%	52.43%
	H2	0.00%	0.00%	17.07%	29.67%	15.36%
	H3	0.00%	0.00%	3.66%	0.00%	1.12%
	H4	44.83%	27.78%	2.73%	25.27%	28.46%
	H5	1.72%	5.56%	1.22%	6.59%	3.75%
	H6	0.00%	19.44%	12.20%	12.09%	10.49%
	H7	0.00%	0.00%	2.44%	2.20%	1.50%
Complexity level	K1	98.28%	77.78%	82.93%	63.74%	79.03%
	K2	1.72%	13.89%	17.07%	36.26%	19.85%
	K3	0.00%	8.33%	0.00%	1.10%	1.50%
Answer form	A1	93.10%	83.33%	92.68%	86.81%	89.51%
	A2	6.90%	16.67%	7.32%	13.19%	10.49%
Context	C1	53.45%	83.33%	90.24%	89.01%	80.90%
	C2	46.55%	5.56%	7.32%	5.49%	14.98%
	C3	0.00%	11.11%	2.44%	5.49%	4.12%

The results regarding the mathematical *activities* required reveal that half of the analysed geometry tasks in each grade require presentation activities (H1). The opposite activity, interpretation (H4), is present in 45% of the tasks in the first grade, while the tasks for other grades vary between 21% and 28%. Estimation and measuring (H6) are present in one fifth of items in the second grade, while argumentation (H5) is underrepresented in all of the examined grades.

The *complexity* category shows the predominance of simple tasks (K1) in all grades. Simpler connections (K2) increase in the printed textbook tasks from Grade 1 to Grade 4. The results also show the clear dominance of closed *answers* in the geometry chapters in all of the examined grades, as well as intra-mathematical tasks. Although 47% of all of the examined tasks in the first grade have a realistic *context*, this is underrepresented in all of the other grades (5–7%). There are barely any authentic tasks (C3) in the analysed textbook set.

Digital textbooks. The results of task requirements in the digital textbooks according to the five-dimensional framework are presented in Table 6. In terms of the *content category*, lines, two-dimensional objects and measurement are present in all of the grades, which is in line with curriculum requirements (MZO, 2019). However, points, which are highlighted in the first-grade curriculum, are minimally present in the e-tasks. As can be seen in Table 6, some content is not required in any of the examined e-tasks, such as operations with geometric forms (F2), tasks with positional relationships and spatial visualisation (F3), forms in the environment (F6), and geometrisation (F7). These ideas are not present in the Croatian curriculum for primary grades either, except for F6, which is highlighted as important (MZO, 2019). Geometric patterns (F5) are only present in the first grade (12%), in line with curricular requirements.

Interpretation (H4) is the most required *activity* in the digital tasks (from 45% to 81%). Representations and calculation are represented with 15% and 17%, respectively, in all of the e-tasks. Argumentation (H5) and operation (H3) are only present in Grade 4, and to a small extent. Estimation and measuring (H6) are found mostly in the second-grade e-tasks, but they are not present at all in the first and fourth grades (Table 6).

Table 6
Task requirements in the digital textbooks

Categories and codes		Grade 1 (n = 68)	Grade 2 (n = 20)	Grade 3 (n = 70)	Grade 4 (n = 175)	TOTAL (n = 333)
Content (fundamental ideas of geometry)	F1.0	1.47%	40.00%	17.14%	0.00%	6.31%
	F.1.1	8.82%	50.00%	54.29%	10.86%	21.92%
	F.1.2	23.53%	35.00%	8.57%	69.14%	45.05%
	F.1.3	33.82%	30.00%	2.86%	0.00%	9.31%
	F2	0.00%	0.00%	0.00%	0.00%	0.00%
	F3	0.00%	0.00%	0.00%	0.00%	0.00%
	F4	30.88%	250.00%	34.29%	21.14%	26.13%
	F5	11.76%	0.00%	0.00%	0.00%	2.40%
	F6	0.00%	0.00%	0.00%	0.00%	0.00%
	F7	0.00%	0.00%	0.00%	0.00%	0.00%
Mathematical activity	H1	19.12%	30.00%	12.86%	12.57%	15.02%
	H2	0.00%	5.00%	28.57%	20.57%	17.12%
	H3	0.00%	0.00%	0.00%	1.14%	0.60%
	H4	80.88%	45.00%	51.43%	63.43%	63.36%
	H5	0.00%	0.00%	0.00%	2.86%	1.50%
	H6	0.00%	20.00%	5.71%	.00%	2.40%
	H7	0.00%	0.00%	1.43%	0.57%	0.60%
Complexity level	K1	100.00%	100.00%	100.00%	94.86%	97.30%
	K2	0.00%	0.00%	0.00%	5.14%	2.70%
	K3	0.00%	0.00%	0.00%	0.00%	0.00%
Answer form	A1	100.00%	100.00%	100.00%	100.00%	100.00%
	A2	0.00%	0.00%	0.00%	0.00%	0.00%
Context	C1	44.12%	85.00%	87.14%	93.14%	81.38%
	C2	55.88%	15.00%	12.86%	3.43%	16.82%
	C3	0.00%	0.00%	0.00%	3.43%	1.80%

In terms of task *complexity*, the results revealed the dominance of simple e-tasks in all four grades, while connections (K2) are only represented by a very small percentage in Grade 4. Furthermore, the results show that the closed *answer* form (A1) is required in all of the examined digital tasks. Intra-mathematical *context* (C1) increases from the first to the fourth grade, while realistic context decreases. In the fourth grade, only a very small percentage of the digital tasks have an authentic context (C3).

Comparison of task requirements in the printed and digital textbooks. One of the aims of the study is to compare the task requirements in the printed and digital tasks. According to the total results (Tables 5 and 6), the *content* proportion of the printed tasks is greater for 1-dimensional objects, while there are more tasks with 2-dimensional objects among the digital tasks. Tasks containing the concept of points and solids are differently distributed over the grades in the printed textbooks and the e-textbooks. The printed textbooks contain points to a similar extent (8%) in Grades 1 and 2, while points are differently distributed in the e-tasks in Grades 1 to 3, with 40% in Grade 2. The proportion of geometric solids is greater in the digital than in the printed textbooks, but they are present mainly in Grades 1 and 2. Tasks with measurement (F4) are present to a greater extent in the digital than the printed tasks. Geometric patterns (F5) are present only in the first grade in both the printed and the digital tasks, as well as in the curriculum. It is interesting to note the similarity of ideas in both the printed and the digital tasks in the fourth grade: only lines, 2D objects and measurement are present in Grade 4. The curriculum covers these fundamental ideas, but also emphasises F6. Forms in the environment (F6) are present in only 5% of the tasks in the printed textbooks, while they are not present at all in the digital textbooks. There were no tasks that contain operations with geometric forms (F2), or tasks with positional relationships (F3) and geometrisation requirements (F7), in either the printed or the digital textbooks.

A comparison of the data (Table 5, Table 6) shows that the most required mathematical *activity* in the printed tasks is presentation (52%), while the same activity is present in only 15% of the e-tasks. On the other hand, interpretation (H4) is a frequent requirement in the digital textbook tasks (63%). Estimation and measuring (H6) is present more frequently in the printed tasks (10.5%) than in the digital tasks (2%). Calculation (H2) is present to a similar extent in the printed tasks (15%) and the digital tasks (17%), while argumentation (H5) is barely present in either the printed or the digital textbooks.

The *complexity*, *answer type* and *context* categories show similar proportions in the printed and the digital tasks analysed: the dominance of simple tasks (K1), closed answer forms (A1) and intra-mathematical tasks (C1). Open-ended questions (A1) are present only in the printed textbooks. Realistic context (C2) is present to a similar extent in the printed textbooks (15%) and the digital textbooks (17%), while authentic tasks (C3) are underrepresented in both sources.

Special features of digital tasks

The second research question refers to the special features and potential provided in the digital tasks. According to the results given in Table 7, almost half of the total number of e-tasks take the *task form* of a multiple-choice question (TF1), with the highest proportion in the third grade (60%). In almost one fifth of the digital tasks, the answer needs to be filled in (TF2), but there is a big difference in the proportion of these tasks in each grade: 30% in the fourth grade and less than 2% in the first grade. E-tasks with matching (TF3) are most common in the first grade, while true/false tasks are most common in the second grade. Put-in-order tasks (TF5) are present only in the first and third grades, but in less than 7% of the analysed tasks.

Verification *feedback* (F2) is given in 95% of all of the examined e-tasks. If the given answer is wrong, in half of the examined tasks, students can try to solve the task again (IF2). Try-again-feedback (IF2) is most represented in Grade 2 (65%), with about 50% in Grades 3 and 4, while in Grade 1 it is present in a smaller percentage of the tasks (30%). If the student's answer is wrong, some types of tasks give correct answers or error flagging. The correct answer (IF1) is most present in Grade 3 tasks, while error flagging (IF3) is most common in Grade 4. The results show that elaborated feedback (IF4), as well as hints/clues or prompts (IF5), are not present in any of the digital geometry tasks.

Table 7
Special features of the digital tasks

Categories and codes		Grade 1 (n = 68)	Grade 2 (n = 20)	Grade 3 (n = 70)	Grade 4 (n = 175)	TOTAL (n = 333)
Digital features	D1	10.29%	15.00%	0.00%	0.00%	3.00%
	D2	89.71%	85.00%	100.00%	100.00%	97.00%
Personalisation	P1	0.00%	0.00%	0.00%	0.00%	0.00%
	P2	100%	100%	100%	100%	100%
Task form	TF1	47.06%	45.00%	60.00%	44.00%	48.05%
	TF2	1.47%	20.00%	18.57%	30.86%	21.62%
	TF3	23.53%	5.00%	8.57%	18.86%	16.82%
	TF4	.00%	15.00%	8.57%	4.57%	5.11%
	TF5	7.35%	0.00%	2.86%	0.00%	2.10%
	TF6	20.59%	15.00%	1.43%	1.71%	6.31%
	TF7	0.00%	0.00%	0.00%	0.00%	0.00%

Categories and codes		Grade 1 (n = 68)	Grade 2 (n = 20)	Grade 3 (n = 70)	Grade 4 (n = 175)	TOTAL (n = 333)
Feedback	F1	10.29%	15.00%	.00%	2.29%	4.20%
	F2	89.71%	85.00%	100.00%	97.71%	95.80%
	IF1	16.18%	15.00%	31.43%	22.86%	22.82%
	IF2	32.35%	65.00%	54.29%	54.86%	50.75%
	IF3	16.18%	5.00%	25.71%	29.14%	24.32%
	IF4	0.00%	0.00%	0.00%	0.00%	0.00%
	IF5	0.00%	0.00%	0.00%	0.00%	0.00%
	IF6	7.35%	0.00%	0.00%	1.14%	2.10%
Cooperation	COO1	0.00%	0.00%	0.00%	0.00%	0.00%
	COO2	100%	100%	100%	100%	100%

Although *cooperation* and *personalisation* are presented in the literature as instances of the great potential of digital tasks, the present study showed that they are not utilised at all in the geometry tasks examined. Another advantage of digital textbooks is being able to provide *dynamic* diagrams for exploring for investigations and gaining a better understanding of mathematical concepts, particularly in geometry. However, dynamic diagrams for exploring (D1) are only present in Grades 1 and 2 in less than 15% of the digital tasks examined.

Discussion and conclusions

The study findings reveal the presence of traditional requirements in both the printed and the digital textbooks, with a predominance of simpler tasks with closed answer forms and intra-mathematical context. The results of the first research question, which points to the differences between task requirements in digital and printed tasks, suggest that presentation is the most frequent mathematical activity in the printed textbooks, while interpretation is most present in the digital textbooks. Furthermore, the proportion of measurement items is greater in the digital tasks than in the printed tasks. The second research question, which pointed to additional features of digital tasks, brought the following results. In the digital textbooks, a variety of answer forms are used and feedback is given in almost all of the digital tasks (95%). Nonetheless, the significant potential of digital tasks in areas such as cooperation, interactivity and dynamics is not represented at all in the analysed items. These results indicate that the examined textbook tasks do not provide a full range of features according to both instruments: the five-dimensional framework and the framework on digital features.

The analysis according to Wittmann's (1999) fundamental ideas shows that both sources provide a rather narrow picture of geometry through the given tasks. This finding is in line with the results presented in Glasnović Gracin and Kuzle (2018, 2019). It is also important to note that textbook tasks in Croatia are subject to curriculum requirements. Therefore, some fundamental ideas are not represented in the examined tasks because they are not required in the Croatian curriculum for mathematics (MZO, 2019); for example, symmetry, 3D objects, and geometric patterns. In this way, the content opportunities to learn geometry provided in textbook tasks are influenced by what is prescribed in the curriculum. These results call for a discussion on rethinking the Croatian geometry curriculum so that it is organised around more fundamental ideas. Consequently, printed and digital tasks would also provide a wider picture of geometry, helping students to create opportunities for a broader view of mathematics (Henningsen & Stein, 1997; Sullivan et al., 2013). Further in-depth research on the content requirements of Croatian mathematics curriculum is therefore needed.

Digital textbooks and the features they include may bring new opportunities to learning mathematics. According to Usiskin's (2018) conceptualisation of different types of e-textbooks, the digital textbook explored in the present study would be a hybrid resource because it contains both a paper and an electronic form, the latter with built-in digital activities. Ruthven (2017) claims that in such resources, the new digital media should provide different forms of interaction with the student and not "simply replicate the functionality" (p. 261) of traditional resources. However, the study presented in this paper revealed that the potential of digital geometry tasks in terms of interactivity, dynamics and collaboration is not being utilised. These findings are in line with the results presented in Pepin et al. (2017) about the features of digital curriculum resources: they are described as "still relatively rudimentary" (p. 652), with no emphasis on connectedness and lacking focus on the educative nature of digital resources. Therefore, generally speaking, the potential of digital textbook tasks have not yet been unlocked, and we can see here that the challenge in the future lies in developing the educational features of digital curriculum resources.

International implications

The study presented in this paper may be of interest to the international audience. It uses a five-dimensional framework for analysing the textbook tasks in primary grades, which is an extension of research conducted for middle grades (Glasnović Gracin, 2018). This framework has been used and cited worldwide, and the study shows that it can also be implemented for digital tasks. Furthermore, we developed a multi-dimensional framework for examining the special features of digital tasks, which can be implemented in different countries. This framework may contribute to raising awareness of the potential of e-tasks, as they have not yet been well investigated and developed (Pepin et al., 2017). This framework may be implemented for other mathematical disciplines, as well as for other school subjects and in different countries.

Limits of the study and future research

The study presented in this paper shows the features of printed and digital textbook geometry tasks; it also provides a framework for examining the special features of digital tasks. However, examining a larger number of textbooks (not just one textbook set, and not only from Croatia) may contribute to a better insight into the variety of task features provided in printed and e-textbooks. This study focuses on exercise tasks, but other parts of textbooks may also be examined, such as motivation or new content blocks. In addition, it is important to note that some parts of the instrument presented in Table 2 (special features of digital tasks) are more suitable for a larger set of tasks than tasks as individual items (e.g., the personalisation category). Future research would encompass an extension of the study to the role of teachers as mediators between tasks and students. The interplay between using printed and digital tasks opens new research spaces, particularly in relation to the untapped potential of digital resources. Moreover, the features of digital resources, such as dynamics, feedback, personalisation and cooperation, open up new ideas for students and research on student roles in learning mathematics.

It seems that the new technologies afford great possibilities for teaching and learning, but these possibilities have not yet been fully exploited. Task design is a well-developed branch of research in mathematics education; now is the time for deeper consideration of digital task design and its implementation in current resources.

References

Choppin, J., Carson, C., Borys, Z., Cerosaletti, C., & Gillis, R. (2014). A typology for analyzing digital curricula in mathematics education. *International Journal of Education in Mathematics, Science and Technology, 2*(1), 11–25.

Cohen, V. B. (1985). A reexamination of feedback in computer-based instruction: Implications for instructional design. *Educational Technology, 25*(1), 33–37.

Glasnović Gracin, D. (2018). Requirements in mathematics textbooks: A five-dimensional analysis of textbook exercises and examples. *International Journal of Mathematical Education in Science and Technology, 49*(7), 1003–1024.

Glasnović Gracin, D., & Kuzle, A. (2018). Drawings as external representations of children's mathematical ideas and emotions in geometry lessons. *CEPS – Center for Educational Policy Studies Journal, 8*(2), 31–53. https://doi.org/10.26529/cepsj.299

Glasnović Gracin, D., & Kuzle, A. (2019). Fundamentalne ideje za nastavu geometrije [Fundamental ideas for teaching geometry]. *Matematika i škola, 99*, 147–151.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Henningsen M., & Stein M. (1997). Mathematical tasks and student's cognition: Classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education, 28*(5), 524–549.

Institut für Didaktik der Mathematik. (2007). *Standards für die mathematischen Fähigkeiten österreichischer Schülerinnen und Schüler am Ende der 8. Schulstufe* [Standards for the mathematical competencies of Austrian students at the end of the 8th grade]. Alpen-Adria- Universität Klagnefurt.

Johansson, M. (2006). *Teaching mathematics with textbooks. A classroom and curricular perspective.* [Doctoral dissertation, Luleå University of Technology]. Luleå University of Technology. http://ltu.diva-portal.org/smash/record.jsf?pid=diva2%3A998959&dswid=-133

Kurnik, Z. (2000). Matematički zadatak [Mathematical task]. *Matematika i škola, 7*, 51–58. https://mis.element.hr/fajli/545/07-02.pdf

Kuzle, A., & Glasnović Gracin, D. (2020). Making sense of geometry education through the lens of fundamental ideas: An analysis of children's drawing. *The Mathematics Educator, 29*(1), 7–52.

Leung, A., & Baccaligni-Franck, A. (Eds). (2017). *Digital technologies in designing mathematics education tasks. Potential and pitfalls.* Springer.

Love, E., & Pimm, D. (1996). "This is so": A text on texts. In A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook of mathematics education* (Vol. 1., pp. 371–409). Kluwer International Handbooks of Education, Springer.

Mammana, C., & Villani, V. (Eds.). (1998). *Perspectives on the teaching of geometry for the 21st century: An ICMI study.* Kluwer Academic Publishers.

Markovac, J. (2001). *Metodika početne nastave matematike* [Methodology of initial mathematics teaching]. Školska knjiga.

Miklec, D., Jakovljević Rogić, S., & Prtajin, G. (2021a). *Moj sretni broj 1* [My lucky number 1]. Školska knjiga.

Miklec, D., Jakovljević Rogić, S., & Prtajin, G. (2021b). *Moj sretni broj 2* [My lucky number 2]. Školska knjiga.

Miklec, D., Jakovljević Rogić, S., & Prtajin, G. (2021c). *Moj sretni broj 3* [My lucky number 3]. Školska knjiga.

Miklec, D., Jakovljević Rogić, S., & Prtajin, G. (2021d). *Moj sretni broj 4* [My lucky number 4].

Školska knjiga.

Ministry of Science and Education. (2019). *Kurikulum predmeta Matematika* [Curriculum of the school subject Mathematics]. https://narodne-novine.nn.hr/clanci/sluzbeni/2019_01_7_146.html

Organisation for Economic Co-operation and Development. (2003). *The PISA 2003 assessment framework – Mathematics, reading, science and problem solving knowledge and skills.* OECD. https://www.oecd.org/education/school/programmeforinternationalstudentassessmentpisa/33694881.pdf

Pepin, B., & Haggarty, L. (2001). Mathematics textbooks and their use in English, French and German classrooms: A way to understand teaching and learning cultures. *ZDM – Mathematics Education, 33*(5), 158–175.

Pepin, B., Gueudet, G., Yerushalmy, M., Trouche, L., & Chazan, D. (2016). E-textbooks in/for teaching and learning mathematics: A potentially transformative educational technology. In L. English & D. Kirshner (Eds.), *Handbook of international research in mathematics education* (pp. 636–661). Taylor & Francis.

Pepin, B., Choppin, J., Ruthven, K., & Sinclair, N. (2017). Digital curriculum resources in mathematics education: Foundations for change. *ZDM – Mathematics Education*, *49*, 645–661

Rezat, S. (2021). How automated feedback from a digital mathematics textbook affects primary students' conceptual development: Two case studies. *ZDM – Mathematics Education*, *53*(6), 1433–1445. https://doi.org/10.1007/s11858-021-01263-0

Ruthven, K. (2018). Instructional activity and student interaction with digital resources. In L. Fan, L. Trouche, Ch. Qui, S. Rezat, & J. Visnovska (Eds.), *Research on mathematics textbooks and teachers' resources. ICME-13 monographs.* (pp. 261–275). Springer. https://doi.org/10.1007/978-3-319-73253-4_12

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153–189. https://doi.org/10.3102/0034654307313795

Sullivan, P., Clarke, D. M., & Clarke, B. A. (2013). *Teaching with tasks for effective mathematics learning*. Springer.

Usiskin, Z. (2018). Electronic vs paper textbook presentations of the various aspects of mathematics. *ZDM – Mathematics Education*, *50*(5), 849–861.

Wittmann, E. Ch. (1999). Konstruktion eines Geometriecurriculums ausgehend von Grund- ideen der Elementargeometrie [Construction of a geometry curriculum based on the basic ideas of elementary geometry]. In H. Henning (Ed.), *Mathematik lernen durch Handeln und Erfahrung. Festschrift zum 75. Geburtstag von Heinrich Besuden* (pp. 205–223). Bueltmann und Gerriets.

Zhu, Y., & Fan, L. (2006). Focus on the representation of problem types in intended curriculum: a comparison of selected mathematics textbooks from Mainland China and the United States. *International Journal of Science and Mathematics Education*, *4*(4), 609–626.

Biographical note

Dubravka Glasnović Gracin, PhD, is an associate professor in the field of mathematics education at the Faculty of Teacher Education, University of Zagreb. Her research interests include mathematics textbook research, task analysis and development, resources for mathematics teaching and learning, and studying the fundamental ideas of geometry within the tripartite model of curriculum.

Ana Krišto is a primary school teacher at Primary School Petar Preradović in Zagreb. Her areas of interest include mathematics textbook research with a focus on task analysis, geometry teaching in primary education and research related to elementary school curriculum.

doi: 10.26529/cepsj.1303

The Teaching of Initial Multiplication Concepts and Skills in Croatian Textbooks

Goran Trupčević*[1] and Anđa Valent[2]

The goal of this paper is to describe the teaching of initial multiplication concepts and skills, up to the multiplication table, in the Croatian educational system. As Stiegler and Hiebert (1999) concluded, teaching is a complex system rooted in a cultural script of a given society. To describe it without ignoring certain features of it that appear to be self-evident to an insider, it is necessary to step out of this cultural frame. For that reason, we study the teaching of initial multiplication in Croatia by comparing Croatian mathematical textbooks with textbooks from Singapore, Japan, and England. For the textbook analysis, we adapt the framework of Charalambous, Delaney, Hsu, and Mesa that examines a textbook as an environment for the construction of knowledge of a single mathematical concept. The analysis provides evidence that practice and automation are at the centre of the initial learning of multiplication in Croatia. The meaning of multiplication usually is not clear, and pupils are not provided additional tools for developing understanding, nor they are encouraged to use different calculation strategies in a flexible manner. The study also indicates that Croatian textbooks present mathematics as a practice that is closed and pre-given, restricted to the one and the only right way through it.

Keywords: concept construction, multiplication, textbook analysis

1 *Corresponding Author. Faculty of Teacher Education, University of Zagreb, Croatia; gtrupcevic@gmail.com.
2 Zagreb University of Applied Sciences, Croatia.

Pouk začetnih pojmov in spretnosti množenja v hrvaških učbenikih

Goran Trupčević in Anđa Valent

Cilj tega prispevka je opisati, kako se poučevanje začetnih pojmov in spretnosti množenja, vključujoče vse do poštevanke, udejanja znotraj hrvaškega vzgojno-izobraževalnega sistema. Kot sta sklenila Stiegler in Hiebert (1999), je poučevanje zapleten sistem, ukoreninjen v kulturnem zapisu dane družbe. Zato da ga opišemo, ne da bi zoperstavili določene značilnosti, ki se sicer njegovim notranjim članom zdijo samoumevne, moramo stopiti izven kulturnega okvirja. Iz tega razloga proučujemo začetno poučevanje množenja na Hrvaškem s primerjanjem hrvaških matematičnih učbenikov z učbeniki iz Singapura, Japonske in Anglije. Za analizo učbenika smo prevzeli ogrodje, ki ga predlagajo Charalambous, Delaney, Hsu in Mesa, ki presoja učbenik kot okolje za stvaritev znanja enega samega matematičnega koncepta.

Analiza priskrbi z dokazi, da sta praksa in avtomatizem v samem središču začetnega poučevanje množenja na Hrvaškem. Pomen množenja običajno ni jasen, pri čemer učencem niso zagotovljena dodatna orodja za razvoj razumevanja, niti niso spodbujeni, da bi uporabili različne računske strategije na fleksibilen način. Raziskava kaže tudi na to, da hrvaški učbeniki prikažejo matematiko kot prakso, ki je zaprta in vnaprej dana, omejena na en sam in edino pravilen način obravnave.

Ključne besede: konstrukcija konceptov, množenje, analiza učbenika

Introduction

In 2014, Croatia started a process of curriculum reform for primary and secondary education. New curricula for all school subjects were finished in 2016 and were sent for an expert review. In the new curriculum for mathematics, learning of multiplication and multiplication tables should start in Grade 2, but full mastery was expected in Grade 3, since it requires time and a great deal of practice. In their commentaries on the first draft of the curriculum, many primary teachers called for the expectation of mastery of the multiplication table by the end of Grade 2 from the old curriculum (MZOŠ, 2007) to be retained. This proposition was adopted in the revised version of the new mathematics curriculum (MZO, 2019).

The many recommendations concerning the mastery of the multiplication table lead us to consider the initial teaching and learning of multiplication as an example of what Stiegler and Hiebert (1999) described when they characterise teaching as a cultural activity. These kinds of activities are rooted in the cultural script: they reflect attitudes and beliefs of a culture and are manifested in the patterns of behaviour. These attitudes, beliefs, and behaviours are learned implicitly through observation and participation inside the culture. Since they are widely shared, they are difficult to see. This makes them highly stable over time and not easily changed.

The old Croatian mathematics curriculum reflects this: 'In the existing elementary school, mathematics is a subject with a long tradition and well-defined content, and no major interventions are required in current programmes. Therefore, the existing programmes formed the starting basis in the design of the mathematics curriculum' (MZOŠ, 2006, p. 238).

To investigate these kinds of practices, one needs to step out of one's cultural circle. In this way, one can become aware of the scripts they are using and, by comparing different cultural scripts, one can see possibilities, not just what is there but also what is not (Stiegler & Hiebert, 1999). One of the ways of doing this for teaching practices is through the comparative international textbook analysis (Haggarty & Pepin, 2002).

Multiplicative reasoning

That 'multiplicative thinking appears early and develops slowly' was the conclusion of Clark and Kamii (1996) based on their analysis involving children in Grades 1 to 5. There are multiple reasons for that.

In studying children's solutions to multiplication problems, Bell et al. (1984) found that the solutions varied according to the numbers involved, the structures involved, and the context of the problem. They saw this as another manifestation of 'meaning blindness'. Although traditionally, the teaching of multiplication and division was the teaching of procedures, this showed the necessity of also teaching different kinds of multiplicative structures. Or, in the words of Kaput (1985, as cited in Greer, 1992), mathematics and its 'applications' should be taught as being of a piece from the very beginning.

In their search for the origins of children's mistakes in multiplication problems, Fischbein et al. (1985) hypothesised that children link every arithmetic operation to an implicit, unconscious, primitive model that mediates the meaning of operation when solving problems. The model, in turn, imposes its own constraints on the understanding of operations. For multiplication, they argue, children use repeated addition as an implicit model. This hypothesis of repeated addition as the origin of multiplicative thinking is reflected in curriculum documents of different countries (Park & Nunes, 2001; MZOŠ, 2006).

When considering multiplicative situations or structures from the point of the underlying process or procedure, different kinds of these situations can be identified: equal grouping, rate, array, measure conversion, multiplicative comparison, Cartesian product (Anghileri, 1989; Bell et al., 1984). Greer (1992) synthesises classifications of multiplicative situations into four major classes: equal groups (with equal measure and rate; *Ann, Ivy, and Kathy have 5 apples each. How many apples do they have altogether?*), multiplicative comparison (*Mary has 3 times as many chocolates as Steve. If Steve has 5 chocolates, how many chocolates does Mary have?*), rectangular array (and area; *Chocolates are sorted in an array of 3 rows and 5 columns. How many chocolates are there?*) and Cartesian product (*John has 3 pants and 5 shirts. In how many ways can John dress?*).

Vergnaud (1988), in contrast, considers the mathematical relations between quantities in these situations. In this way, he finds two main types of situations. The first he calls 'an isomorphism of measures', which involves two pairs of different quantities. There are two types of relationships in this situation: one between the quantities of a different kind (ratio) and one between values of the same kind of quantities (scalar factor). Another type of situation concerns a product of measures, which involves a third quantity, a factor quantity, connecting two other kinds of quantities.

Different researchers considered children's solution strategies for multiplication problems. From the point of the degree of abstraction, these strategies were classified as modelling (and counting), counting (by 1, rhythmic, skip), repeated addition, and multiplicative operations (Anghileri, 1989; Kouba, 1989;

Mulligan & Mitchellmore, 1997). Kouba (1989) also noted that when children used physical objects for solving problems, they either represented individual elements in each set, or they represented only the tallies. For the same multiplication expression, children's solution strategies can change if the meaning of multiplication or its representation changes (Fosnot & Dolk, 2001; Kouba, 1989). This phenomenon is not exclusive to multiplication but has been observed in other topics (Kolar et al., 2018; Tirosh et al., 2018). Fosnot and Dolk (2001) elaborated multiplicative operations strategies by linking them to the properties of multiplication.

The research into children's solution strategies for multiplication problems put forward a theory of repeated addition as an origin of multiplication into question (Kouba, 1989). Although counting and adding are used to obtain the numerical answer, there is something that goes on before the counting (Nunes & Bryant, 1996; Steffe, 1994).

These considerations lead to an alternative hypothesis that views the concept of multiplication as defined by an invariant relation between two quantities, called 'ratio' or 'rate' (Piaget, 1965; Vergnaud, 1988). Thus, the development of a one-to-many correspondence scheme lies at the origin of the understanding of multiplicative situations (Nunes & Bryant, 1996). The works of Piaget (1965) and others (see Nunes & Bryant, 1996) show that children as early as five years of age are able to represent and use one-to-many correspondences. Furthermore, arguments for the repeated addition theory can result from children's previous school experience (Nunes et al., 2010). An intervention study by Park and Nunes (2001) showed that pupils who were taught multiplication through correspondence made significantly more progress than pupils who were taught multiplication through repeated addition.

This shows that multiplication is not simply a new arithmetic operation to be learned after addition but that there are also new meanings and new situations to be learned. Moreover, although there is a procedural connection between multiplication and addition, there is also a conceptual discontinuity between the two (Nunes & Bryant, 1996).

Research on Mathematical Textbooks

Mathematics textbook research is a rich and growing area of investigation. In their overview of the literature, Fan et al. (2013) conceptualised four categories of mathematics textbook research: the role of textbooks in mathematics teaching and learning, textbook analysis and comparison, textbook use, and other areas of textbook research.

Studies from different countries have shown that textbooks have a considerable impact on the teaching and learning taking place in the classrooms (Fan & Kaeley, 2000; Haggarty & Pepin, 2002; Pepin & Haggarty, 2001). Haggarty and Pepin (2001, 2002) analysed textbook structure and content, their usage by the teachers, and the access to the textbooks for the pupils. They found that textbooks from different countries differ both in the structure and in the complexities and learning opportunities they offer for the pupils. Furthermore, the way teachers used textbooks varied between different countries, but also between different educational strands in one country. Authors argued that some of the factors that influenced these issues come from the cultural tradition, consisting of the organisation of the educational system, educational traditions, values and epistemic beliefs, and the socio-economic conditions of both the pupils and teachers. Studies in Croatia also confirmed the impact that textbooks have on teaching and learning (Glasnović Gracin & Domović, 2009; Glasnović Gracin & Jukić Matić, 2016; Jukić Matić, 2019). These studies also revealed social issues that influence textbook utilisation.

A comprehensive TIMSS study of mathematics and science textbooks from 48 countries (Valverde et al., 2002) assumes a four-part model of the curriculum; besides the intended, implemented, and attained curriculum, they add the potentially implemented curriculum as the fourth part, and view textbooks as parts of it, as mediators between intended and implemented curriculum. Their description of textbooks considers their macrostructure (size, length, etc.) but also their microstructure for which textbooks are sequenced into blocks of analysis, which are then coded by block type, mathematical content, presentation expectations, and a wider perspective on the subject. This analysis shows variations in textbook structures across different school systems.

The study by Jones and Fujita (2013), which uses the adapted TIMSS framework, showed that mathematical textbooks in England and Japan clearly reflect the geometrical component of a national curriculum. Charalambous et al. (2010) developed a framework for investigating learning opportunities and pupils' expectations in the textbooks, particularly with respect to the presentation of specific content. Their study considered the treatment of the addition and subtraction of fractions in textbooks from Cyprus, Ireland, and Taiwan and found greater variations between textbooks across cultures than within one country, which is in accordance with the findings of the TIMSS video study (Stigler & Hiebert, 1999). This framework was later used in similar comparative studies in different countries and for different mathematical topics (Hong & Choi, 2014; Yang & Sianturi, 2017). Boonlerts and Inprasitha (2013) analysed the presentation of multiplication in elementary textbooks from Japan, Singapore, and Thailand through content analysis (sequencing of topics and meaning of multiplication). The meaning given to multiplication was conceptualised using Greer's (1992) categorisation of classes of situations involving multiplication, and the analysis showed differences in the textbooks from the three countries from this perspective. Analysis of the sequencing of topics showed that textbooks from Singapore and Japan offer their pupils increasingly sophisticated strategies to solve multiplication problems.

This study aims to describe important features of the treatment of initial multiplication in mathematics textbooks in Croatia. This kind of description could yield important information for the future development of the national curriculum and curricular materials. Also, the findings of this study can be used for informing and changing the teaching practice in primary schools in Croatia.

To describe the treatment of initial multiplication in Croatian textbooks, we compare them with textbooks from Singapore, Japan, and England. The primary reasons for choosing these countries are the good results that students from these countries achieve in international comparative studies and the availability of textbooks in the English language. Furthermore, since these countries are geographically distant, and their tradition is historically and culturally different from that of Croatia, one can expect that insight into culturally hidden assumptions and practices of the Croatian teaching tradition could be obtained by comparison with these countries.

We attempt to describe features of Croatian textbooks by answering the following research questions:

1. What is the structure of a typical lesson in Croatian textbooks?
2. What possible multiplication constructs and representations are used?
3. What multiplication strategies are promoted?
4. Do Croatian textbooks promote the usage of tools and manipulatives for learning?
5. What features of a mathematics classroom are implicated by Croatian textbooks?
6. To what extent do Croatian textbooks reflect the national curriculum?

General assumptions and initial multiplication in the Mathematics Curricula of Croatia, Singapore, Japan, and England

Before the textbook analysis, an overview of national mathematics curricula was made so that it could be used for the interpretation of the results.

The Croatian 2006 curriculum is divided into two parts: the introductory part and the list of themes for each year of study, together with pupils' learning achievements. Themes of beginning multiplication are placed in Grade 2.

The Singapore Mathematics Syllabus (MOE, 2012) is organised through a list of mathematical contents and appropriate learning experiences for each year of study are given at the end of the syllabus. The learning of multiplication and multiplication tables takes place in the first three years.

The 2009 Japanese curriculum standards (Isoda, 2010) operationalises its objectives for each grade through a list of learning outcomes, a list of relevant terms and symbols, a list of mathematical activities to be used, and some specific remarks on the ways of handling the content. The learning of multiplication and multiplication tables takes place in Grade 2.

The 2013 English Mathematics Curriculum (DfE, 2013) for each year gives statutory requirements but also some notes and guidance. Pupils learn multiplication during the first four years.

Table 1 summarises different aspects of the mathematics curriculum of these four countries. One can observe that curricula have much more in common regarding their main emphasis. However, when it comes to the elaboration of the content of initial multiplication and description of the ways of working, the Croatian curriculum is rather concise on that matter. It will be interesting to see whether this will also be reflected in the textbooks.

Table 1
Comparison of mathematics curricula of Croatia, Singapore, Japan, and England regarding multiplication

	Croatia	Singapore	Japan	England
Emphasis of the Curriculum	basic mathematical knowledge, development of skills, problem-solving	problem-solving, conceptual understanding, skills proficiency, mathematical process, attitudes, metacognition	basic knowledge and skills, ability to think and express, attitudes	fluency, mathematical reasoning, problem-solving
Grades	2	1, 2, 3	2	1, 2, 3, 4
Multiplication constructs	addition	addition, equal groups, scaling	equal groups	addition, equal groups, array, scaling, Cartesian product
Multiplication strategies	commutativity	counting, concretisation, patterns, heuristics	counting, properties of multiplication, commutativity, adding next, patterns	counting, concretisation, doubling, patterns, connections, commutativity, associativity, distributivity
Representations and manipulatives		pictorial, concrete objects	pictorial, concrete objects	pictorial, concrete objects
Forms of work		group work, sharing ideas		

Method

Based on the data of the Croatian Ministry of Science and Education (MZOS, 2014), the three most commonly used Croatian textbook series were chosen for the analysis: *Matematika*, *Moj sretni broj*, *Nove matematičke priče*. In accordance with the Croatian curriculum, all these textbooks are used in the second year of study. Together they comprised a market share of 78.33% of mathematics textbooks in the second grade in Croatia. In addition, the following textbook series were analysed: *Primary Mathematics* and *My Pals are Here* from Singapore, *Sansu Math* from Japan, and *Power Maths* from England, used from Grades 1–3, 2–3, and 1–4, respectively. Throughout this paper, the names of these textbooks will be abbreviated as CT1, CT2, CT3, ST1, ST2, JT, and ET, respectively.

An analytical framework was based on the framework of Charalambous, Delaney, Hsu, and Mesa (2010) for textbook analysis, which views a textbook as an environment for the construction of knowledge of a single mathematical concept. All lessons of a textbook concerning initial multiplication learning upon the multiplication table were analysed. Each lesson was divided into smaller blocks that were analysed from several aspects (categories). Based on the results of the pilot study (Baković et al., 2019.), the analytical framework was adapted to include the following categories: block type, social form of work, context, use of concrete materials, images, characters, representations, construct and multiplication strategies. The coding list for each category was created by using the grounded theory approach (Glaser & Strauss, 1967); the initial list was created on the basis of the literature review and was later revised on the basis of the observed data. The full list of codes is given in Table 2.

The category *block type* concerns its function in the text with the way that the block is communicated to pupils.

The category *context* deals with the context of the mathematical problem posed in a block. Authentic context refers to a pupil's personal experience while realistic to a possible experience.

The category *construct* refers to the meaning of multiplication. For this purpose, Greer's (1992) classification was revised so that multiplicative comparison, rate, and equal measure problems were included in a joint class that was named 'scaling'. This change was motivated by children's different ways of representing grouping and matching problems in Kouba's study (1989).

Three categories were concerned with the visual features of blocks. The first one, the *use of pictures* category registers the presence of the picture and whether it is mathematically relevant or not.

The *representations* category refers to the way that multiplication is represented. The representation usually corresponds to the construct, but sometimes the two can differ. For instance, the text can be about four groups of three, while the picture shows 12 objects sorted in a 4×3 array.

Finally, the category *other graphical objects* refers to the function of graphical objects that were used. These were mostly images of pupils or some other characters that communicate some information to pupils, but there were also diagrams, drawings, and images of objects to use.

The category *use of manipulatives* assesses if pupils are instructed to use concrete, manipulative materials and, if so, which ones. Physical manipulatives are divided into standardised, which are used throughout the textbook, and sporadic, which are used only once.

The category *multiplication strategy* assesses both which strategies for carrying out multiplication are promoted in the block, as well as problem-solving and general learning strategies.

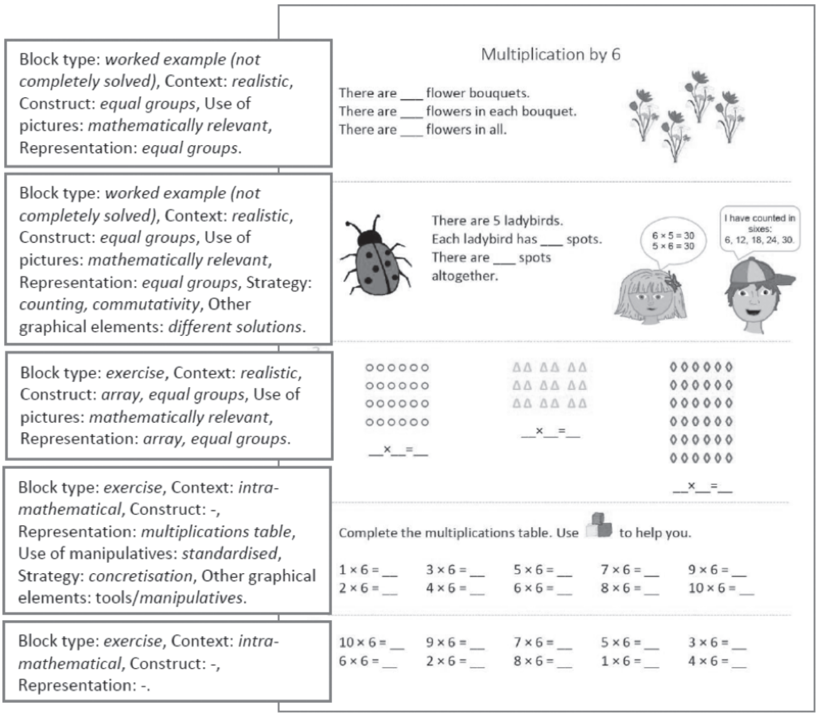
The category *social form of work* characterises organisational ways pupils are supposed to work. Only blocks for which this is explicitly stated were categorised, other blocks were categorised as 'unclear'.

Table 2
Coding list

Block Type	**Use of manipulatives**
Situation	Standardised manipulative
Recap	Sporadic manipulative
Worked example	10x10 table
Definition/rule	Games
Exercise	Multiplication cards
Activity/game	Hands
Self-assessment	
	Strategy
Context	Commutativity
Authentic	Multiplications table
Realistic	Multiplication by 1 and 0
Intra-mathematical	Counting
	Adding/Subt next
Construct	Distributivity
Addition	Doubling
Equal groups	Bar-model
Array	Modelling/Concretisation/Drawing
Scaling/Comparison	Metacognition
Counting	Estimation
Cartesian product	Mnemonic
Without construct	Associativity
Use of pictures	**Other graphical elements**
Absent	Instruct/describe
Mathematically relevant	Solution
Context related	Explanation
	Guidance
Representation	Different solutions
Equal groups	Additional questions
Linear	Terminology
Array	Rule
Number line	Drawing/diagram
Multiplications table	Tools/manipulatives
Counting	
Bar/Scaling	**Social form of work**
10×10 table	Pair
Cartesian product	Group
Without Representation	Family
	Unclear

We demonstrate the coding procedure on the sample page given in Figure 1.

Figure 1
Example of coding a textbook page



Results and discussion

The total number of blocks that were identified and analysed in the textbooks is given in Table 3.

Table 3
Number of analysed blocks in the textbooks

CT1	CT2	CT3	ST1	ST2	ET	JT
139	117	171	197	136	420	101

The results of the analysis of blocks are given below, for each category, in terms of percentages of the blocks in the textbooks.

Block type

The distribution of codes inside this category is shown in Table 4. Recaps appear more frequently in Croatian textbooks than in non-Croatian textbooks. In non-Croatian textbooks, they appear only at the beginning of chapters, while in Croatian textbooks each lesson begins with a recap. This can be considered an indicator of a compartmentalised and unconnected view of mathematical knowledge. Also, Croatian textbooks had neither activities/games nor self-assessment blocks.

Table 4
Distribution of block type in the textbooks

Block Type	CT1	CT2	CT3	ST1	ST2	ET	JT
Situation	1%	1%		1%			2%
Recap	10%	18%	15%	1%	1%		
Worked example	24%	9%	26%	24%	29%	8%	30%
Definition/rule	6%	15%	6%	8%	4%		9%
Exercise	59%	57%	53%	57%	48%	89%	53%
Activity/game			1%	9%	15%		6%
Self-assessment					3%	2%	

Context and Construct

Table 5
Distribution of Context in the textbooks

Context	CT1	CT2	CT3	ST1	ST2	ET	JT
Authentic				13%	16%	6%	3%
Realistic	32%	24%	30%	63%	71%	62%	47%
Intra-mathematical	68%	76%	70%	23%	13%	33%	50%

The context of the majority of the blocks in Croatian textbooks is intra-mathematical, and it is never authentic (see Table 5). This is different from non-Croatian textbooks in which there is a prevalence of realistic content, and all of them include blocks with authentic context.

Table 6

Distribution of Construct in the textbooks

Construct	CT1	CT2	CT3	ST1	ST2	ET	JT
Addition	17%	27%	19%	8%	7%	14%	4%
Equal groups	23%	15%	11%	65%	60%	41%	23%
Array	2%	9%	3%	21%	13%	17%	17%
Scaling/Comparison	17%	5%	24%	6%	7%	20%	33%
Counting				20%	7%	2%	1%
Cartesian product						2%	
Without construct	55%	64%	54%	17%	19%	28%	49%

The majority of blocks of Croatian textbooks do not have a construct specified (see Table 6). In non-Croatian textbooks, a construct is evident in almost all blocks. The high rate of blocks without evident constructs in JT textbooks is a consequence of the fact that the focus of the last 20% of lessons is to encourage pupils to use different multiplication strategies and to investigate multiplication properties and patterns; thus, a construct in these lessons is not evident.

Some non-Croatian textbooks systematically introduce different constructs and sometimes explicitly ask pupils to make connections between different constructs. This is not the case in Croatian textbooks in which new constructs usually appear 'out of nowhere', with the occasional observation that 'this is also multiplication'.

It can be noted that in Croatian textbooks repeated addition is promoted at a higher rate of blocks than in non-Croatian textbooks. Also, non-Croatian textbooks promote counting as a construct, which is not the case in Croatian textbooks.

An explanation for these findings could be rooted in differences between curricula. Unlike other curricula, the Croatian curriculum does not point out different aspects of multiplication and defines multiplication only as repeated addition.

Representations, use of pictures and manipulatives

Table 7
Use of pictures in the textbooks

Use of pictures	CT1	CT2	CT3	ST1	ST2	ET	JT
Absent	76%	75%	74%	37%	17%	28%	38%
Mathematically relevant	24%	25%	24%	60%	79%	66%	55%
Context related			2%	3%	1%	5%	7%

Almost 75% of blocks in Croatian textbooks are without representation, which is significantly lower than in non-Croatian textbooks (see Table 7). They are usually present in worked examples at the beginning of a lesson, while in exercises a representation is usually absent.

Table 8
Distribution of Representation in the textbooks

Representation	CT1	CT2	CT3	ST1	ST2	ET	JT
Equal groups	15%	11%	8%	46%	32%	36%	28%
Linear	5%	2%	4%	3%	15%	3%	7%
Array	1%	10%	8%	20%	29%	20%	27%
Number line	3%	12%	8%			14%	
Multiplications table	7%	11%	6%	6%	13%	3%	20%
Counting	1%	7%		20%	10%	2%	2%
Bar/Scaling				4%	2%	11%	10%
10×10 table				2%	2%	1%	
Cartesian product						2%	
Without Representation	70%	58%	75%	21%	18%	35%	39%

Besides pictorial representation, some non-pictorial representations also appear; in Croatian textbooks, these are a multiplication table and counting, while in non-Croatian textbooks, we additionally registered a 100-table (see Table 8). There is also a difference in the usage of multiplication tables; in Croatian textbooks, they are simply given without any further instructions. In non-Croatian textbooks, pupils are encouraged to use these tables to observe patterns and properties of multiplication.

Table 9

Use of manipulatives in the textbooks

Use of manipulatives	CT1	CT2	CT3	ST1	ST2	ET	JT
Standardised manipulative				15%	10%	5%	3%
Sporadic manipulative				2%	1%		
10×10 table				2%	1%	1%	8%
Games					1%	1%	3%
Multiplication cards				2%	2%		2%
Hands				8%	1%		

As shown in Table 9, Croatian textbooks do not promote the use of manipulatives. In non-Croatian textbooks, one can observe the intention to standardise the use of manipulatives: the same manipulatives are used throughout different lessons, giving pupils the opportunity to make connections between mathematical content and its representation with those manipulatives.

Strategies

Table 10

Distribution of Strategy in the textbooks

Strategy	CT1	CT2	CT3	ST1	ST2	ET	JT
Commutativity	6%	15%	13%	23%	8%	10%	16%
Multiplications table	7%	10%	12%	5%	13%	8%	24%
Multiplication by 1 and 0	7%	6%	8%	2%	1%	3%	4%
Counting	1%	7%		21%	10%	6%	1%
Adding/Subt next	1%			3%	3%	6%	24%
Distributivity	2%			9%	5%	4%	14%
Doubling				2%		9%	3%
Bar-model				5%	2%	4%	8%
Modelling/Concretisation/Drawing			3%	12%	4%	12%	15%
Metacognition	1%			5%	6%	15%	22%
Estimation						3%	
Mnemonic		1%		4%	1%		1%
Associativity							1%

Croatian textbooks promote only multiplication tables, multiplication by 1 and 0 and commutativity as strategies for carrying out multiplication (see Table 10). Multiplication by 0 and 1 are mostly present inside corresponding lessons listed in the curriculum. In non-Croatian textbooks, this strategy is far less present. Commutativity is also differently treated: in non-Croatian textbooks it is a strategy that pupils can use in their calculations, or when looking for patterns or global properties of multiplication. Unlike this, Croatian textbooks usually present commutativity as a property that pupils are just supposed to check.

Modelling, concretisation, and drawing as strategies that enable pupils to organise their own thinking around multiplication problems are almost never present in Croatian textbooks. This is in accordance with the Croatian curriculum which, unlike non-Croatian curricula, pays no attention to these strategies. The same holds for counting as a strategy.

Strategies that enable pupils to calculate flexibly, such as doubling, adding/subtracting one and distributivity, are not present at all in Croatian textbooks, while non-Croatian textbooks systematically encourage pupils to use these strategies.

Non-Croatian textbooks systematically introduce the bar model as a pre-algebraic strategy for solving word problems. This is not present in Croatian textbooks where pupils are not given additional tools to help them with modelling or solving problems.

Unlike Croatian textbooks, non-Croatian textbooks pay attention to the development of different metacognitive strategies: recognition of boundaries of certain concepts, self-assessment, scaffolding in problem-solving, making connections between different parts of knowledge, and similar.

Other graphical elements

Table 11
Use of other graphical elements in the textbooks

Other graphical elements	CT1	CT2	CT3	ST1	ST2	ET	JT
Instruct/describe		2%	14%	2%	5%	1%	5%
Solution		5%	1%		4%	2%	3%
Explaining			2%	16%			2%
Guidance				11%	8%	11%	14%
Different solutions				2%	1%	5%	11%
Additional questions				2%	7%	6%	9%
Terminology		1%		1%	7%		
Rule		1%			4%	1%	2%
Drawing/diagram				2%			5%
Tools/manipulatives				13%	15%		

The majority of *other graphical elements* refer to certain characters in textbooks. These are mostly pupils, but some abstract characters also appear (e.g., a magician in CT3). It is interesting to note the difference between the functions of these characters in textbooks (see Table 11). Characters in Croatian textbooks give instructions, solutions, and explanations. The magician in CT3 leads pupils through the story. Besides these functions, characters in non-Croatian textbooks give guidance for pupils' solutions, they ask additional questions and give and seek more than one solution. We believe that the functions of these characters give insight into the way teaching is supposed to unfold in the classroom. Croatian textbooks present a classroom in which the teacher is the central figure, 'the sage on the stage'. Pupils in this classroom are expected to give solutions to problems in only one, correct way. The mathematics that this classroom presents is closed and given, restricted to the one and only right way through it.

Unlike this, non-Croatian textbooks present classrooms that are pupil-centred. Pupils are the ones that solve problems, guide each other, discuss different solutions, and open new questions. Mathematics in this case is an open and connected domain of flexible thinking and discussions among practitioners.

Social form of work

Table 12
Social forms of work in the textbooks

Social form of work	CT1	CT2	CT3	ST1	ST2	ET	JT
Pair				3%	3%	2%	4%
Group				7%	2%		4%
Family				1%			

Unlike non-Croatian textbooks, Croatian textbooks do not call for work in pairs or in a group (see Table 12). This additionally supports the idea of a mathematical activity as a solitary endeavour. Furthermore, it gives fewer educational opportunities for pupils (Boaler, 2016).

Conclusion

This analysis provides evidence that there are only minor variations of teaching multiplication within different Croatian textbooks and at the same time all Croatian textbooks are noticeably different from the observed textbooks from England, Japan, and Singapore. The initial learning of multiplication in Croatia puts practice and automation in focus. The majority of content is intra-mathematical, not connected to the real world or authentic problems, and pupils are not encouraged to use manipulatives to help them model realistic problems or represent abstract mathematical problems. The meaning of multiplication is usually not clear nor is it visually represented. It is expected that the pupils solve exercises, but they are not encouraged to use different strategies and approaches.

The study also indicates some general features of teaching and learning mathematics in Croatia. Unlike non-Croatian textbooks, Croatian textbooks do not promote the development of metacognitive skills and strategies. Pupils are not encouraged to find different solutions and to make connections between them, to discuss their solutions, or to pose problems.

These findings suggest that the teaching practice should strive to compensate for these deficits. This means helping pupils in making sense of multiplication by situating problems in a familiar context, giving them opportunities to explore multiplication by using pictures and manipulatives, and giving them opportunities to develop different strategies for calculating.

Focusing solely on the textbooks is a limitation of this study. To obtain a fuller picture, one should see what is going on inside the actual classrooms. In fact, to get insight into this, an observational study had been started in March 2020 but was closed soon because of a lockdown due to the Covid-19 pandemic. Since the implementation of the new curriculum in Grade 2 started in September 2020, it was no longer possible to continue it.

All these features are in accordance with the Croatian curriculum from 2006. The new curriculum, which puts more stress on the use of manipulatives and pictorial representations, on problem-solving and formative assessment, will be fully implemented by 2022. It would be interesting to conduct a follow-up study in the future and see to what extent the changes made in the curriculum will be reflected in the textbooks and the teaching practice.

References

Anghileri, J. (1989). An investigation of young children's understanding of multiplication. *Educational Studies in Mathematics*, *20*(4), 367–385. https://doi.org/10.1007/BF00315607

Baković, T., Trupčević, G., & Valent, A. (2019). Treatment of initial multiplication in textbooks from Croatia and Singapore. In Z. Kolar-Begović, R. Kolar-Šuper, & Lj. Jukić Matić (Eds.), *Towards new perspectives on mathematics education* (pp. 216–228). Element.

Bell, A., Fischbein, E., & Greer, B. (1984). Choice of operation in verbal arithmetic problems: The effects of number size, problem structure and context. *Educational Studies in Mathematics*, *15*(2), 129–147. https://doi.org/10.1007/BF00305893

Boaler, J. (2016). *Mathematical mindsets: unleashing students potential through creative math, inspiring messages. and innovative teaching*. Jossey-Bass.

Boonlerts, S., & Inprasitha, M. (2013). The Textbook Analysis on Multiplication: The Case of Japan, Singapore and Thailand. *Creative Education*, *4*(4), 259–262. https://doi.org/10.4236/ce.2013.44038

Charalambous, C.Y., Delaney, S., Hsu, H.-Y., & Mesa, V. (2010). A comparative analysis of the addition and subtraction of fractions in textbooks from three countries. *Mathematical Thinking and Learning, 12*(2), 117–151. https://doi.org/10.1080/10986060903460070

Clark, F. B., & Kamii, C. (1996). Identification of multiplicative thinking in children in grades 1-5. *Journal for Research in Mathematics Education*, *27*(1), 41–51. https://doi.org/10.2307/749196

Department for Education [DfE]. (2013). *Mathematics programmes of study: Key stages 1 and 2. National curriculum in England*. DfE. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/335158/PRIMARY_national_curriculum_-_Mathematics_220714.pdf

Fan, L., & Kaeley, G. S. (2000). The influence of textbooks on teaching strategies: An empirical study. *Mid-Western Educational Researcher*, *13*(4), 2–9.

Fan, L., Zhu, Y., & Miao, Z. (2013). Textbook research in mathematics education: Development status and directions. *ZDM: The International Journal on Mathematics Education*, *45*(5), 633–646. https://doi.org/10.1007/s11858-013-0539-x

Fischbein, E., Deri, M., Nello, M. S., & Marino, M. S. (1985). The role of implicit models in solving verbal problems in multiplication and division. *Journal for Research in Mathematics Education*, *16*(1), 3–17. https://doi.org/10.2307/748969

Fosnot, C. T., & Dolk, M. (2001). *Young mathematicians at work. Constructing multiplication and division.* Heinemann.

Foong, P. Y., Chang S.H., Lim L. G. P., & Wong O.H. (2014). *Primary mathematics 1B*. Shinglee.

Foong, P. Y., Chang S.H., Lim L. G. P., & Wong O.H. (2015). *Primary mathematics 3A*. Shinglee.

Foong, P. Y., Lim L. G. P., & Wong O. H. (2014). *Primary mathematics 2A*. Shinglee.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Aldine.

Glasnović Gracin, D. & Domović, V. (2009). Upotreba matematičkih udžbenika u nastavi viših razreda osnovne škole [The use of mathematics textbooks in teaching of upper primary grades]. *Odgojne znanosti*, *18*(2), 45–65. https://hrcak.srce.hr/48441

Glasnović Gracin, D., & Jukić Matić, Lj. (2016). The role of mathematics textbooks in lower secondary education in Croatia: An empirical study. *The Mathematics Educator 16*(2), 31–58.

Greer, B. (1992). Multiplication and division as models of situations. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 276–295). Macmillan.

Haggarty, L., & Pepin, B. (2002). An investigation of mathematics textbooks and their use in English, French and German classrooms: Who gets an opportunity to learn what? *British Educational Research Journal*, *28*(4), 567–590. https://doi.org/10.1080/0141192022000005832

Hong, D. S., & Choi, K. M. (2014). A comparison of Korean and American secondary school textbooks: the case of quadratic equations. *Educational Studies in Mathematics, 85*(2), 241–263. https://doi.org/10.1007/s10649-013-9512-4

Isoda, M. (Ed.) (2010). *Elementary school teaching guide for the Japanese course of study mathematics (Grade 1-6)*. CRICED, University of Tsukuba. https://www.criced.tsukuba.ac.jp/math/apec/ICME12/Lesson_Study_set/Elementary_School_Teaching_Guide-Mathematics-EN.pdf

Jones, K., & Fujita, T. (2013). Interpretations of national curricula the case of geometry in textbooks from England and Japan. *ZDM Mathematics Education*, *45*(5), 671–683. https://doi.org/10.1007/s11858-013-0515-5

Jukić Matić, Lj. (2019). The teacher as a lesson designer. *Center for Educational Policy Studies Journal*, *9*(2), 139–160. Ttps://doi.org/10.26529/cepsj.722

Kheong, F. H., Ramakrishnan, C., & Choo, M. (2014). *My pals are here maths 2A*. MC Education.

Kheong, F. H., Ramakrishnan, C., & Choo, M. (2015). *My pals are here maths 3A*. MC Education

Kheong, F. H., Ramakrishnan, C., & Wah, B. L. P. (2013). *My pals are here maths 1B*. MC Education.

Kolar, V. M., Hodnik Čadež, T., & Vula, E. (2018). Primary teacher students' understanding of fraction representational knowledge in Slovenia and Kosovo. *Center for Educational Policy Studies Journal*, *8*(2), 71–96. https://doi.org/10.26529/cepsj.342

Kouba, V. L. (1989). Children's solution strategies for equivalent set multiplication and division word problems. *Journal for Research in Mathematics Education*, *20*(2), 147–158. https://doi.org/10.2307/749279

Markovac, J. (2015). *Matematika 2* [Mathematics 2]. Alfa.

Miklec, D., Prtajin, G., & Jakovljević Rogić, S. (2015). *Moj sretni broj 2 [My lucky number 2*]. Školska knjiga.

Ministarstvo znanosti obrazovanja i športa [MZOŠ]. (2006). *Nastavni plan i program za osnovnu školu. [Course of study for primary school].* MZOŠ.

Ministry of Education Singapore [MOE]. (2012). *Mathematics syllabus: Primary one to Six*. MOE. https://www.moe.gov.sg/-/media/files/primary/mathematics_syllabus_primary_1_to_6.pdf

Mulligan, J. T., & Mitchelmore, M. C. (1997). Young children's intuitive models of multiplication and division. *Journal for Research in Mathematics Education*, *28*(3), 309. https://doi.org/10.2307/749783

Ministarstvo znanosti i obrazovanja [MZO]. (2019). Odluka o donošenju kurikuluma za nastavni predmet Matematike za osnovne škole i gimnazije u Republici Hrvatskoj [Decision of declaration of the curriculum of Mathematics for primary and secondary school]. *Narodne novine* 7/2019. https://narodne-novine.nn.hr/clanci/sluzbeni/2019_01_7_146.html

Ministarstvo znanosti obrazovanja i sporta [MZOS]. (2014). Konačne liste odabranih udžbenika i pripadajućih dopunskih nastavnih sredstava prema odabiru stručnih aktiva u osnovnim i srednjim školama [Final lists of selected textbooks and related supplementary teaching aids according to the selection of professional boards in primary and secondary schools]. https://mzo.gov.hr/vijesti/konacne-liste-odabranih-udzbenika-i-pripadajucih-dopunskih-nastavnih-sredstava-prema-odabiru-strucnih-aktiva-u-osnovnim-i-srednjim-skolama/1031

Nunes, T., & Bryant, P. (1996). *Children doing mathematics*. Blackwell.

Nunes, T., Bryant, P., Evans, D., & Bell, D. (2010). The scheme of correspondence and its role in children's mathematics. *British Journal of Educational Psychology*, *2*(7), 83–99. https://doi.org/10.1348/97818543370009X12583699332537

Park, J., & Nunes, T. (2001). The development of the concept of multiplication. *Cognitive Development*, *16*(3), 1–11. https://doi.org/10.1016/S0885-2014(01)00058-2

Pepin, B., & Haggarty, L. (2001). Mathematics textbooks and their use in English, French and German classrooms: A way to understand teaching and learning cultures. *ZDM: The International Journal on Mathematics Education*, 33(5), 158–175. https://doi.org/10.1007/BF02656616

Piaget, J. (1965). *The child's conception of number*. Norton.

Polak, S., Cindrić, D., & Duvnjak, S. (2014). *Nove matematičke price 2 [New mathematical tales 2].* Profil.

Sansu Math 2. (2015). Koyo Publishing.

Sansu Math 3. (2015). Koyo Publishing.

Staneff, T. (Ed.) (2017). *Power maths 1C*. Pearson.

Staneff, T. (Ed.) (2017). *Power maths 2A*. Pearson.

Staneff, T. (Ed.) (2018). *Power maths 3A*. Pearson.

Staneff, T. (Ed.) (2018). *Power maths 3B*. Pearson.

Staneff, T. (Ed.) (2018). *Power maths 4A*. Pearson.

Steffe, L. (1994). Children's multiplying schemes. In G. Harel & J. Confrey (Eds.), *The development of multiplicative reasoning in the learning of mathematics* (pp. 3–40). SUNY Press.

Stiegler, J., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. The Free Press.

Stručna radna skupina [SRS]. (2016). *Nacionalni kurikulum nastavnog predmeta Matematika, Prijedlog* [The national curriculum of Mathematics, A proposal]. http://www.kurikulum.hr/wp-content/uploads/2016/03/Matematika.pdf

Stručna radna skupina [SRS]. (2016). *Odgovori na pristigle priloge stručnoj raspravi Nacionalni kurikulum nastavnoga predmeta Matematika* [Answers to the received contributions to the expert discussion of the national curriculum of Mathematics]. http://www.kurikulum.hr/wp-content/uploads/2016/06/komentari-stručna-rasprava-matematika.pdf

Tirosh, D., Tsamir, P., Barkai, R., & Levenson, E. (2018). Engaging young children with mathematical activities involving different representations: Triangles, patterns, and counting objects. *Center for Educational Policy Studies Journal*, *8*(2), 9–30. https://doi.org/10.26529/cepsj.271

Valverde, G. A., Bianchi, L. J., Wolfe, R. G., Schmidt, W. H., & Houang, R. T. (2002). *According to the book: Using TIMSS to investigate the translation of policy into practice through the world of textbooks*. Kluwer.

Vergnaud, G. (1988). Multiplicative structures. In J. Hiebert & M. Behr (Eds.), *Number concepts and operations in middle grades* (pp.141–161). NCTM.

Yang, D. C., & Sianturi, I. A. (2017). An analysis of Singaporean versus Indonesian textbooks based on trigonometry content. *EURASIA Journal of Mathematics Science and Technology Education*, *13*(7), 3829–3848. https://doi.org/10.12973/eurasia.2017.00760a

Biographical note

Goran Trupčević, PhD, is an assistant professor in the field of mathematics at the Faculty of teacher education, University of Zagreb, Croatia. His research interests include representation theory, mathematical textbooks and mathematics teacher noticing.

Anđa Valent, PhD, is a master lecturer in the field of mathematics at Zagreb University of Applied Sciences, Croatia. Her research interests include representation theory, history of mathematical education and mathematical textbooks.

doi: 10.26529/cepsj.1309

Nature of Science in Greek Secondary School Biology Textbooks

Nausica Kapsala*[1], Apostolia Galani[2] and Evangelia Mavrikaki[2]

The nature of science describes what science is, how it works, and its interactions with society under the perspectives of philosophy, history, sociology, and psychology of science. Understanding it is an essential aspect of scientific literacy. Given the critical role that school textbooks hold, considering what is taught and how it is taught in schools, we find the presence of the nature of science in school science textbooks to be significant. In this research paper, all Greek biology textbooks of lower secondary education are analysed to evaluate whether principal elements of the nature of science can be found in them. The whole array of educational resources available (textbooks, workbooks, lab guides, teachers' books) was analysed as well as the corresponding official biology curricula. Content analysis was the method of choice, and the 'meaning unit' was the unit of analysis. We found that most of the nature of science references in the material that students were taught in 2021/22 was implicit and not especially designed by the curriculum. Some nature of science aspects were more commonly found (e.g., evidence is vital in science) than others (e.g., science has limits). The most opportunities for the nature of science to be introduced were found in history of science vignettes, laboratory activities, and some optional inquiry activities. However, without a structured design from the curriculum, it is the teachers' responsibility to design and facilitate nature of science instruction (or not). We conclude that lacking explicit references, the nature of science falls into the hidden curriculum and becomes falsely depicted, enforcing a positivist image of science.

Keywords: biology textbooks, Greece, nature of science, secondary education

1 *Corresponding Author. Second Lyceum of Elefsina and Faculty of Primary Education of the National and Kapodistrian University of Athens, Greece; nkapsala@gmail.com.

2 National and Kapodistrian University of Athens, Greece.

Narava znanosti v grških srednješolskih bioloških učbenikih

Nausica Kapsala, Apostolia Galani in Evangelia Mavrikaki

Narava znanosti opisuje, kaj je znanost, kako deluje in kakšni so njeni stiki z družbo z vidikov filozofije, zgodovine, sociologije in psihologije znanosti. Razumevanje je temeljni vidik znanstvene pismenosti. Upoštevajoč pomembno vlogo, ki jo imajo šolski učbeniki glede tega, kaj in kako je poučevano v šolah, ugotavljamo, da je prisotnost narave znanosti v šolskih naravoslovnih učbenikih bistvena. V tem prispevku so analizirani vsi učbeniki za biologijo v osnovni šoli v Grčiji, in sicer z namenom ovrednotenja, ali je v njih mogoče najti osnovne elemente, povezane z naravo znanosti. Preiskana je bila celotna zbirka razpoložljivih učnih gradiv (npr. učbeniki, delovni zvezki, laboratorijski vodniki, priročniki za učitelje) pa tudi pripadajoč uradni učni načrt za biologijo. Izbrana metoda je bila vsebinska analiza, pri čemer je kot enota analize služila »pomenska enota«. Ugotovili smo, da je večina sklicev na naravo znanosti v gradivih, s pomočjo katerih so poučevali v letih 2021–2022, implicitna. Nekateri vidiki narave znanosti so bili pogosteje najdeni (npr. dokazi so nujni v znanosti) kot drugi (npr. znanost ima omejitve). Največ priložnosti, ob katerih je bila narava znanosti vpeljana, je bilo vezanih na vinjete zgodovine znanosti, laboratorijske naloge in na nekatere izbirne poizvedovalne aktivnosti. Brez strukturiranega načrta, izhajajočega iz učnega načrta, je odgovornost učiteljev, da načrtujejo in vpeljejo v pouk tudi vsebine narave znanosti (ali ne vpeljejo). Sklepamo, da zaradi primanjkljaja neposrednih sklicev narava znanosti pristane znotraj skritega kurikuluma in je napačno prikazana, vsiljujoča pozitivno sliko znanosti.

Ključne besede: učbeniki za biologijo, Grčija, narava znanosti, sekundarno izobraževanje

Introduction

Textbooks hold an influential role in education. They constitute an essential source for students to obtain knowledge; therefore, inadequate and inconsistent science knowledge presented in the textbooks can affect students' conceptions. According to Dimopoulos and Karamanidou (2013, in Slapničar, 2014), science is unfortunately mispresented as static and absolute knowledge in school textbooks. This provides students with a false view of what science is and how it works.

The Nature of Science (NOS) is the field that describes what science is, how it works, and its interactions with society under the perspectives of the philosophy, history, sociology, and psychology of science (McComas, 2020). Understanding NOS affects how people conceive scientific information and how much they trust science and its exhortations (Song et al., 2021). Citizens' knowledge of the NOS positively impacts the acceptance of evolution, Global Climate Change (GCC), and vaccination, regardless of political ideology and religiosity (Weisberg et al., 2020).

Understanding NOS constitutes 'a critical component of scientific literacy that enhances students' understanding of science concepts and enables them to make informed decisions about scientifically based personal and societal issues' (NSTA Board of Directors, 2020, p. 1), making the importance of its inclusion in the school curriculum indubitable. However, the incorporation of NOS into school practice is challenging, as textbooks mostly focus on traditional science content (McDonald & Abd-El-Khalick, 2017), and as a result, teachers do not value the NOS content the same as the 'traditional' science content and tend to ignore it (Haagen-Schützenhöfer & Joham, 2018). Even when teachers attempt to teach NOS aspects, they encounter difficulties, mainly due to lack of time, lack of proper instructional material, as well as due to their own misunderstandings (Höttecke & Silva, 2011). It is indicative that NOS instruction is seen as a 'progressive' part of science education that becomes fragile in crucial times, such as during the Covid-19 pandemic (Taber, 2021).

NOS, its definition, and its components may still be a matter of controversy at a philosophical level, but at the level of school instruction, and about what NOS aspects should be taught in school, in other words, what NOS aspects concern a scientifically literate citizen, there is a great deal of consensus among researchers and education stakeholders (Clough, 2011; Lederman, 2007; McComas, 2020; Osborne et al., 2003; Scharmann & Smith, 1999). Different researchers (Lederman & Lederman, 2012; McComas, 2020; Osborne et al., 2003) provide sets of NOS aspects that generally overlap, and some may include aspects that others omit. All agree, though, that any of the proposed lists are not constraining; they are not supposed to be taught and memorised; rather, the aspects should be introduced in the form of questions and raise conversations that will broaden students' views about science. Research has shown that to accomplish effective NOS instruction, NOS aspects should be introduced explicitly, reflectively, and contextualised via inquiry learning experiences or through genuine historical and contemporary episodes of science in action (McComas et al., 2020).

There is criticism against this 'general aspects' conceptualisation of NOS, claiming that it does not adequately describe science (Allchin, 2011; and others). Kampourakis (2016) answers such criticism by stating that NOS school instruction needs to be pragmatic and that the 'general aspects' conceptualisation of NOS has proved useful and effective in introducing students to thinking about NOS and in addressing their preconceptions about science.

For the needs of this study, we adopted McComas' (2020) proposal that is in accordance with most of the relevant existing studies (Lederman & Lederman, 2012; Osborne et al., 2003) and suggests three main domains of NOS with their relevant aspects: the *tools and products of science*, the *human elements of science*, and *science knowledge and its limits*. These, along with explanations and possible misconceptions, are presented in Table 1.

Table 1

NOS aspects proposed to be included in the science curriculum

NOS domain	NOS aspect	Short explanation	Argument against possible relevant misconceptions
A. Tools and products of science	A1. Evidence is vital in science.	Evidence (direct or inferential) must exist both to inspire scientific investigation initially and to support scientific conclusions.	Scientific evidence is not a matter of opinion and cannot be discounted.
	A2. Laws and theories are related but distinct.	Theories and Laws are different kinds of knowledge: Laws are generalisations, principles or patterns in nature that are discovered. Theories are the explanations of those generalisations that are invented.	Theories are not a sort of guess, and they do not mature into laws. Hypothesis may mean a prediction or if it is generalising a 'baby law' or if it is explanatory of a 'baby theory'.
	A3. There are many shared methods in science, but no single scientific method.	There are many shared scientific methods, such as induction, deduction, inference, use of multiple data sources, making testable assertions, etc.	There is no standard stepwise scientific method that all scientists use to explore nature.

NOS domain	NOS aspect	Short explanation	Argument against possible relevant misconceptions
B. Human elements of science	B1. Creativity is everywhere in science.	Two individuals with access to the same facts may reach different conclusions based on their prior knowledge and creativity of one.	Science is not a linear mechanical process. In contrast, it is like making art, creativity and imagination are crucial in seeing problems, recognising patterns, and intuiting solutions.
	B2. Subjectivity and bias are present in science.	Scientists are more knowledgeable about what they study, but they also hold preconceptions and biases about how the world operates, which are usually based on their previous knowledge, theory-laden observation, and the paradigm, which provides direction to the research but may also limit investigation.	Scientists are not more objective than other people. Science as an enterprise, though, makes use of intersubjectivity that cancels the biases of the individuals.
	B3. Society and culture interact with science and *vice versa*.	Scientific work is a human endeavour. Humans interact with each other in various ways. These interactions affect the scientific work. Firstly, most science relies on external funding, which is controlled by governments and private foundations. Then, science is a community affair; new ideas are validated by peers.	Funding influences what scientists will investigate. They cannot work on any problems of interest. Validation by the scientific community also may limit them.
C. Science knowledge and its limits	C1. Science has limits.	There are things that science can never know. Firstly, it is impossible to make all the possible observations and to secure all relevant facts for all time. Second, some areas like religion, ethical decision-making ethics cannot be explored with the tools of science.	Science does not offer absolute proof and cannot potentially address all questions.
	C2. Science is tentative, durable, and self-correcting.	Scientific conclusions are long-lasting but might change when compelling new evidence becomes available. Science is constantly undergoing fine-tuning with the occasional radical changes.	The results of science are not final. Scientific interpretation can change through the self-correcting mechanism built into science.
	C3. Science is distinct from engineering and technology.	Science is not necessarily practical. The pursuit of knowledge for the sake of knowledge is called 'pure science' (knowledge-gaining agenda), while its exploitation in the production of a commercial product is applied science or technology facilitated by engineers (profit-gaining agenda).	Science and engineering are not parts of the same pursuit. They interact, but they are not synonymous.

Note. Adapted from McComas, 2020.

School textbooks hold a dominant role in science education as science teachers worldwide rely almost totally on them for both classroom teaching and students' homework; therefore, they significantly affect both the chosen teaching strategies and students' learning experiences (Galili, 2015; Klassen, 2006; McDonald & Abd El-Khalick, 2017 and references therein). According to McDonald and Abd-El-Khalick (2017), science textbooks become the curriculum and determine what is taught and learned in the classrooms. Therefore, NOS should be appropriately incorporated into school science textbooks.

Unfortunately, school textbooks tend globally to emphasise science as a body of knowledge; they refer to science as a way of investigating; they pay little attention to the interaction among science, technology, and society; they rarely present science as a way of thinking (McDonald & Abd-El-Khalick, 2017 and references within). In most, if not all, science textbooks, there is an introductory chapter with references to NOS but mostly focuses on a stereotypical view of the 'Scientific Method' and without any connection of the scientific knowledge to the community that produces it. In general, the authors of science textbooks tend to pay little attention to NOS content, and when they do, they portray naïve views of NOS (Abd-El-Khalick et al., 2007; Campanile et al., 2013; Irez, 2009, McDonald & Abd-El-Khalick, 2017 and references within; Niaz, 2014).

In Greece, there is only one textbook for each subject in each grade published by the Ministry of Education and distributed to students and teachers, which teachers compulsorily use, so they cannot choose among various textbooks as is the case in other countries. The biology textbooks were written by a team consisting of one university teacher of didactics of biology and two biology teachers in secondary school; they were evaluated by a university teacher, a school counsellor, and a biology teacher in secondary school. Biology textbooks have not been widely researched regarding the representations of NOS. Kampourakis (2017) analysed the Greek biology curriculum and some selective chapters of biology textbooks (the introductory chapter of textbooks for Grades 7, 9, 11, and 12 and the chapters on evolution and Mendelian genetics for Grades 9 and 12). He did not use a fixed list of NOS aspects but attempted to identify references to the nature of scientific inquiry, the nature of scientific knowledge, and the nature of scientific explanation aspects; he concluded that several learning goals related to NOS are included in the Greek biology curriculum, but they are only mentioned as general goals, they are not defined, nor explained, nor is any way of achieving them suggested. The biology textbooks under study approached those goals very superficially. Apart from an introductory chapter that presents a distorted perception of the scientific method, which is also usually not taught at all, there are no direct and explicit references to the

NOS besides some references to the History of Science (HOS) (Kampourakis, 2017). Koumara and Plakitsi (2020) have also conducted research regarding the degree that nature of scientific knowledge aspects is included in the science classes of Greek high schools, in the context of which they analysed the biology textbooks concerning the presence of seven Nature of Scientific Knowledge (NOSK) aspects (Lederman & Lederman, 2012), reaching similar conclusions with Kampourakis (2017).

We aim to analyse all the Greek biology textbooks for lower secondary education regarding NOS references based on a predefined list, such as the one in Table 1. We strongly believe that such research would be important to teachers, who will be provided with a list of NOS references that they can highlight in their teaching, as well as to textbook and curriculum developers who might recognise cases that need to be emphasised. It could also function as a guide for the analysis of other countries' biology school textbooks.

Method

Content analysis (Bryman, 2016) was the method of choice and the 'meaning unit' was the unit of analysis.

We analysed all the biology educational material (student's textbooks, accompanying workbooks, lab books, and the teachers' books) (Mavrikaki et al., 2007a, 2007b, 2007c, 2007d, 2007e, 2007f, 2007g) for Grades 7–9 (secondary school: grades 7–9, ages 12-14, which (along with the primary education) is mandatory education in Greece) as well as the corresponding curriculum (Government Gazette, 2003; Ministry of Education and Religious Affairs, 2021). We must note that the existing curriculum was created in 2003, and the books were developed according to it. Biology was at that time taught in the 7th and 9th grades for two hours per week in each grade. Later, the schools' timetable was changed according to the ministry's guidelines (Government Gazette, 2016), and biology is now taught for one hour per week in all three grades (out of the total school hours, which are 33 hours per week for the 7th grade, 33 for the 8th, and 34 for the 9th grades; the school year consists of 36 weeks); therefore, the total biology teaching hours in the lower secondary education were decreased. However, the books remained the same, and teachers were asked to teach specific chapters in each grade. Every year teaching guidelines are provided to help teachers with the year's syllabus as there are changes every other year.

Two of the authors (coders) performed the analysis independently. The nine NOS aspects (McComas 2020), as presented in Table 1 along with a tenth aspect (the epistemology and the object of biology), guided our analysis. We also evaluated the presence of each NOS aspect according to the scoring rubric that has been adapted from the work of Abd-El-Khalick, Waters, and Le (2008) (Table 2). We evaluated whether each aspect is presented explicitly or implicitly, whether the representation of each NOS aspect is informed or naïve, and whether the representation of the NOS aspect is consistent across each document. With the 10 NOS aspects targeted in the analysis, the overall score for a textbook could range from -30 to +30 points.

To ensure reliability, the two authors independently analysed the educational materials. The process resulted in a high level of inter-coder agreement of 93% (O'Connor & Joffe, 2020).

Table 2

Scoring rubric for NOS references distinguishing between explicit and implicit NOS references

Points	Description
+3	explicit, informed, consistent
+2	explicit, partially informed - incomplete representation, consistent
+1	implicit, informed, consistent
0	absent
-1	implicit, naïve
-2	implicit informed combined with explicit naïve or explicit with conflicting messages
-3	explicit naïve

Note. Adapted from Abd-El-Khalick et al., 2008.

Results

In total, 221 meaning units concerning NOS were identified in the educational material under study, but not all of them are meant to be taught as many have been excluded from teaching according to the ministry's teaching guidelines for the 2021/22 school year. Table 3 presents the distribution of NOS in each document.

Table 3

Number of meaning units concerning NOS found in each of the analysed documents and the number of meaning units concerning NOS that correspond to the material to be taught according to the 2021-2022 ministry's teaching guidelines.

Document	NOS references in the documents (N)	NOS references that are meant to be taught (N)
Curriculum	28	2
2021/22 teaching guidelines	1	1
Biology textbook for Grades 7,8	41	22
Biology workbook for Grades 7,8	5	2
Biology lab guide for Grades 7,8	7	2
Biology teachers' book for Grades 7,8	15	2
Biology textbook for Grades 8, 9	72	40
Biology workbook for Grades 8, 9	12	4
Biology lab guide for Grades 8, 9	11	6
Biology teachers' book for Grades 8, 9	29	7
Total	**221**	**88**

The curriculum includes 28 NOS references, most of which are included in a general part that describes the course's purpose and some of its general goals. The same is true for the teachers' books. Many NOS references can be found in the introductory chapters of the students' textbooks (named 'The science of biology'), the corresponding parts of the curriculum, and the teachers' books, which are excluded from teaching.

The nine NOS aspects, as well as the tenth aspect about the objective of biology, were detected in all the analysed documents, though some were quite scarce and absent from the material taught to the students. Table 4 presents their distribution and score according to the scoring rubric for NOS references that grades the holistic presence of each NOS aspect, regarding them as being explicit or implicit, informed, or naïve, and consistent or with conflicting messages (see Table 2) with the higher the score, the better the way that the NOS aspect is included in the document. It must be noted that we considered it important to show in Table 4 any differences that exist between the educational material that was originally developed and the material that was supposed to be taught in 2021/22 according to the ministry's teaching guidelines for that school year (Ministry of Education and Religious Affairs 109009/Δ2/07-09-2021).

Table 4

Distribution of references to NOS aspects and scores on the target NOS aspects, considering: A) the total analysed material and B) (in parentheses) the material to be taught according to the ministry's 2021/22 teaching guidelines.

NOS aspects	Number of References	Score
Objective of Biology	14 (0)	+3 (0)
A1. evidence	28 (22)	+1 (+1)
A2. laws & theories	11 (8)	-1 (-1)
A3. shared methods	50 (16)	-2 (+1)
B1. creativity	24 (18)	+2 (+1)
B2. subjectivity	4 (2)	-1 (-1)
B3. social & cultural influence	45 (8)	+3 (+2)
C1. Limits	1 (0)	0 (0)
C2. tentative	23 (8)	+2 (+1)
C3. relationship with technology	21 (6)	+1 (+1)
Total	**221 (88)**	**+8 (+5)**

As shown in Table 4, most NOS aspects are more informed in the original material than in the taught one, except for A3: shared methods. The NOS aspect about shared methods in the scientific endeavour scores -2 in all the educational material, as there are explicit naïve references like the following activity: '*Place the following concepts in the appropriate order so that the scientific method can be applied: conclusions, hypothesis formulation, experiment design, observation*' combined with informed implicit references like the description of historical experiments: '*In 1956, a team of scientists conducted an experiment to determine the effect of fluoride on tooth decay in children*'. In contrast, in the taught material, the same NOS aspect scores +1, as the chapters with the explicit naïve references have been excluded from the taught material, and only implicit informed references are included.

Table 5 presents the distribution of the NOS references to the three grades, along with the score on each NOS aspect, considering only the material meant to be taught in 2021/22 according to the ministry's teaching guidelines. Aspects about evidence, shared methods, creativity, and relationships with technology are found in all grades. In Grade 9, there are also aspects about laws and theories, and social and cultural influences. Table 6 contains examples of NOS references found in the textbooks and an evaluation of whether they are explicit/implicit and informed/naïve. The scores in other tables regard the total presence of each NOS aspect in the whole document and cannot be estimated for each example, as it includes one more dimension: consistency throughout the document.

Table 5

Number of references to NOS aspects and scores on the target NOS aspects in each grade, considering the material to be taught according to the ministry's 2021/22 teaching guidelines

	Number of References / Score		
NOS Aspect	**7th grade**	**8th grade**	**9th grade**
A1. evidence	7/+1	6/+1	11/+1
A2. laws & theories	1/-1	0/0	7/-1
A3. shared methods	6/+1	2/+1	9/+1
B1. creativity	4/+1	5/+1	10/+1
B2. subjectivity	1/-1	0/0	1/+1
B3. social & cultural influence	2/+1	3/+1	3/+2
C1. limits	0/0	0/0	0/0
C2. tentative	1/+1	4/+1	4/+1
C3. relationship with technology	3/+1	3/+1	1/+1

Table 6

Examples of references to NOS aspects included in the biology textbooks

NOS Aspect	Grade	Example	Evaluation
A1. evidence	7th	The use of the optical microscope and the observation of cells enabled scientists to reveal another unique feature of organisms that is not found in inanimate objects: cellular organisation.	implicit, informed
A2. laws & theories	9th	There are two theories about the evolution of the human brain and the human's agile fingers.	implicit, naïve
A3. shared methods	7th	In 1956, a team of scientists conducted an experiment to determine the effect of fluoride on the appearance of tooth decay in children. For this purpose, the children of one area drank water to which fluoride had been added, while the children of another area drank water without fluoride. The researchers then compared the percentage of caries-free children in the two areas.	implicit, informed
B1. creativity	8th	Abu Bakr Mohamed Alrazi, better known as Razi, was a famous 9th-century Arab physician. When he had to move to Baghdad to set up a hospital, he chose the area by hanging pieces of meat in various parts of the city and selecting the place where the piece of meat was slow to rot.	implicit, informed

NOS Aspect	Grade	Example	Evaluation
B2. subjectivity	9th	Mendel first published the results of his research in 1865, thus establishing the science of genetics. His work was recognised thirty years after his death.	implicit, informed
B3. social & cultural influence	9th	Thanks to vaccination, many diseases that once plagued humanity have disappeared. A typical example is smallpox.	implicit, informed
C2. tentative	8th	People in the past did not know that many of the diseases they suffered from were caused by micro-organisms. It took many years for the microscope to be discovered, and at the end of the 19th century, this situation changed. Thanks to the work of two important researchers, Louis Pasteur and Robert Koch, it turned out that the cause of many diseases was microorganisms. This discovery initially led to the opposite of the older belief. In other words, it was considered that all microorganisms are pathogenic. Today we know that this is not true.	implicit, informed
C3. relationship with technology	7th, 8th	With the help of optical and electron microscopy, we have now investigated the cells of monocytes and multicellular organisms.	implicit, informed

Discussion

The Greek secondary school biology textbooks, like most science textbooks that have been analysed in the literature (McDonald & Abd-El-Khalick, 2017 and references within), begin with an introductory chapter that discusses biology as a discipline, its objectives, and the process of the scientific method. These chapters, although part of the books, have been excluded from the official 2021/22 syllabus, so they are not taught.

The Objective of Biology. All the analysed documents begin with a reference to biology as a science and its objective, followed by examples (for the textbooks) or educational goals (for the reform documents, the curriculum and the teachers' books). The curriculum begins with the phrase '*Biology is the science that deals with the study of the phenomena and processes of life*'. In the teachers' books, there are corresponding texts; for instance, for the 7th grade, it is stated that '*[students should] acquire the ability to recognise the unity and continuity of scientific knowledge in matters concerning organisms, as well as the ability to recognise the relationship of biology with other sciences*'. This was the extra tenth aspect that we decided to include in our analysis, and it was found to be explicit, well informed (especially in the textbooks) and consistent as it was presented in all the analysed documents in more or less the same way. Unfortunately, the corresponding units are not taught.

A. Tools and Products of Science

A1 – *Science relies on empirical evidence*. This is a critical NOS aspect that can relatively easily be approached compared to others (Yacoubian, 2020). Indeed, it was one of the most commonly found aspects. There are few explicit references, in the chapter on evolution, in the unit 'Evolution and its testimonies', in which it is explicitly claimed that to draw a scientific conclusion, scientists must rely on evidence. Other than that, A1 is present in the textbooks mostly implicitly (e.g., '*In 1665 R. Hook, observing thin sections of cork under his microscope, spoke for the first time about cells*'). However, if the teacher does not grasp the opportunity to discuss it explicitly, it is doubtful that students will understand.

A2 – *Laws and theories are related but distinct*. This NOS aspect is one of the most difficult ones to deeply understand and teach (Mesci & Schwartz, 2016). In the analysed material, there is no reference to the nature of laws or theories as forms of scientific knowledge neither their inter-relationships. In the curriculum, as well as the teachers' books there is a goal that mentions laws and theories '*to acquire knowledge related to concepts, theories, laws and principles related to biology […]*'. Accordingly, in the textbooks, the only times that laws and theories are mentioned are to introduce the cell theory and Mendel's laws. An inconsistency comes up in the chapter of evolution, where the word 'theory' is not only used to convey a system of ideas based on general principles that explains phenomena but also the meaning of supposition: for example, '*There are two theories about the evolution of the human brain and the agile fingers*'. There is a common misconception that theories are a sort of guess and that they are not secure and credible enough, so they can easily be dismissed (McComas, 2020). This kind of thinking may lead to sayings like 'evolution is just a theory'. The way that the word 'theory' is used in the chapter of evolution may strengthen that misconception. To prevent that, it is essential that the biology teacher makes an intervention to explicitly address this issue.

A3 – *Thoughts about methods that scientists use*. This is the most common NOS aspect in the analysed documents. It is the only one that is explicitly mentioned and thoroughly explained, unfortunately supporting the naive conception about the one, linear, steplike, scientific method that all scientists use. In the reform documents, there are several relevant educational goals, such as '*Describe the scientific method and apply it to solving a simple problem*'. In these documents, the scientific method is parallelised with a scientific way of thinking. In the students' textbooks, in the introductory chapter, which is not currently taught, '*the scientific method*' is thoroughly described, discussed, and presented in diagrams as a stepwise linear process, like most science textbooks do throughout the world (Reiff-Cox, 2020). In the taught chapters and the corresponding activities and laboratory exercises, there is no explicit reference to the scientific method. Nevertheless, different methods that scientists use are implicitly presented via the short description of some scientists' work. There is one exception in the 9th -grade teachers' book with an indicative teaching action that explicitly mentions the scientific method: *'Next, we can refer to the discovery of the structure of the DNA molecule by J. Watson and Fr. Crick. [...] We urge them to look at the steps of the scientific method and to combine the content, the history and the significance of this discovery with the scientific way of thinking'*. With appropriate coordination, this could be a suitable activity, although it requires much preparation from the teacher. For the students to understand shared methods in science, they should experience a broad range of the paths that scientists follow (thought experiments, correlational, descriptive, exploratory studies, and serendipitous moments) (Reiff-Cox, 2020). Laboratory exercises and activities about historical scientific episodes give the opportunities to integrally address the aspect.

B. Human Elements of Science

B1 – *Creativity is part of the scientific process*. This NOS aspect is explicitly present twice: in the introductory note for the student at the laboratory guide for the 7th grade (p. 5) *'Designing and executing an experiment requires imagination and hard work'* and in the introductory chapter of the 9th-grade textbook *'The questions that are asked must be explained and this often requires the use of imagination'*. Moreover, creativity can be implicitly detected in the textbooks and the lab guides. There are a few references to HOS so that the biology teacher could present more about the scientists and the circumstances under which they were led to their discoveries and to explicitly discuss creativity as part of the process with the students . However, the provided material is not enough. For instance, in the 9th-grade textbook there is a picture of Robert Koch, accompanied by a text mentioning he is an important scientist who contributed to proving that the cause of many diseases are microorganisms. This is a very typical example of the presence of HOS in science textbooks internationally (Lin et al., 2010), which is by no means sufficient to illuminate the human nature of scientists. Such are the most HOS references in the textbooks and the lab guides (nine cases). In five more cases, the experimental procedure that led to the discovery is also described, some elements are provided to illuminate the fact that for someone to design such an experiment, creativity and imagination are needed.

B2 – *Subjectivity and bias are present in science.* This is a subtle hue of NOS that is difficult to be understood and taught, and in some cases, it is hardly accepted even by the science teachers as it may challenge their broader epistemological stance (Mesci & Schwartz, 2016). Thus, it should not come as a surprise that it is almost absent from secondary school biology. In all the material, we detected only two implicit references that could be made about scientists' subjectivity. The first one is in the 7th-grade textbook, which refers to scientists as authorities: *'Scientists have formed a food pyramid to help us choose the right food'*. In the certain context this reference is not problematic, but when the only way that scientists are pictured is as authorities, which is the case in the secondary education curriculum, then a distorted image of scientists is communicated to the students: superhuman scientists being always objective. The second detected reference is in the 9th-grade textbook: *'Mendel first published the results of his research in 1865 […]. His work was recognized thirty years after his death.'* If this is properly utilised by the teacher, then a discussion about how a scientific discovery is approved and recognised by the scientific community may occur. Mendel's case is an example of how occasionally some innovative ideas are rejected because they fall outside the accepted paradigm. (McComas, 2020).

B3 – *Society and culture interact with science and vice versa.* This aspect has a dual nature: on the one hand, it sheds light on the socio-cultural impacts of science and, on the other hand, on the impact of science on society. The Greek biology curriculum mostly focuses on the latter. It is explicitly stated multiple times in the reform documents that students must *'[…]ascertain the contribution of Biology in the improvement of the quality of human life but also reflect on the effects (positive or negative) of the applications of Biology'*, as well as in the introductory chapters of the students' textbooks *'Thanks to advances in all fields of biology, and especially in molecular biology and genetic engineering, great strides have been made in studies directly or indirectly involving humans'*.

Later the socio-cultural impacts of science are also approached. For example, there is an activity in the 7th-grade textbook that asks students to *'gather historical and other data on the political, social and cultural situation that prevailed and write a text in which you will document the rapid development of science during the 18th century in Europe'*. Another one in the 9th-grade textbooks asks students to reflect on the rapid development of biology, and the ethical, legal, and humanitarian issues that arise. Unfortunately, these chapters are not taught. Neither is the chapter of biotechnology that illuminates the applications of genetic engineering and the related issues that arise. Thus, in the taught chapters there are very few implications about the contribution of biology to society that are (A) about the many ways that humans use germs in various applications, such as the production of alcoholic beverages, in wastewater treatment, in the food industry, and in the pharmaceutical industry, and (B) that *'Thanks to vaccination, many diseases that once plagued humanity, like smallpox, have disappeared'*. The biology teacher must seize this opportunity to introduce all the above concepts that have been excluded from the syllabus. Finally, the 9th-grade teachers' book proposes activities that refer to the impact of the historical context, society, and culture on science: for example,*'We refer briefly to Charles Darwin, his journey and his time. We can assign students to work on Darwin's views and their impact within the specific historical, political, economic and social climate of Victorian England'*.

C. Science Knowledge and its Limits

C1 – *Science has limits*. This aspect is totally absent from the Greek biology curriculum. The only relevant reference is in the 9th-grade's textbook, in the biotechnology chapter, which is not taught, where the concept of bioethics is introduced. This could provide the opportunity for discussions about the questions that science cannot answer, and which are ascribed to ethics, religion, or aesthetics. The second notion of this aspect, concerning how science cannot offer absolute proof, since induction underlies the law of falsifiability (Popper, 2005), is also absent. This omission is not the case in other countries. For example, in the AAAS, the NGSS documents, the Swedish context (Hansson, 2020) and the German curriculum (Marniok & Reiners, 2017), it is clearly stated that the students should learn about the limits of science. It is important to explicitly discuss this notion to avoid creating the feeling that science is omnipotent and that we need no other explanatory or investigative tools (McComas, 2020). Moreover, encouraging the understanding that science is compatible with many different worldviews (including religious ones) may make science meaningful for more students (Hansson, 2020).

C2 – *Science is tentative, durable, and self-correcting*. This is another NOS aspect that is hard to be accepted and understood (Mesci & Schwartz, 2016). Indeed, this NOS aspect is present in the Greek biology curriculum: *'the great development of the science of Biology and the constant updating of data'*, and it is mentioned in all the reform documents; biology is characterised as *'the science whose modern achievements surprise with the pace of revisions and upheavals, they impose on our knowledge and perceptions of them'*. In the textbooks, there are historical examples about cases in which parts of scientific knowledge were reviewed and if needed corrected. For instance, in the 7th-grade textbook it is mentioned that Schleiden and Schwann formulated the cell theory, which was later revised and completed by Virchow. In the 9th-grade textbook, what was considered to be scientifically correct about the cause of diseases is presented *'People did not know in the past that many of the diseases they suffered from were caused by microorganisms. [...] it turned out that the cause of many diseases were some microorganisms. This discovery initially led to the opposite of the older belief. In other words, it was considered that all microorganisms are pathogenic. Today we know that this is not true*'. Such descriptions are important to reveal the tentative, yet durable character of science. According to Niaz (2014, p. 44) 'various topics of the science curriculum provide an opportunity to illustrate the tentative nature of scientific knowledge, and still very few textbooks refer (explicitly) to this important aspect of nature of science.'

C3 – *Science is distinct from engineering and technology.* This NOS aspect has three dimensions; firstly, it concerns the impact of technology on the advancement of science; secondly, the impact of science on the advancement of technology; last, the fact that science is distinct from technology and engineering: they have different goals and different processes; this is contrary to the false image that tends to be created by the Science, Technology, Engineering, and Mathematics (STEM) curricula that are a trend in science teaching nowadays (McComas, 2020). In the analysed documents, the third dimension is absent. There is one explicit learning goal in the 9th-grade teachers' book concerning the biotechnology chapter which is not taught: '*Understand the importance of science and technology interactions in the advancement of biotechnology.*'. Accordingly, in the corresponding textbook's chapter there are several mentions about the ways science and technology affect each other. As far as the taught material is concerned, there are a few mentions about the advancement of the microscope and its vital contribution to the development of biology; for example (7th-grade textbook), '*The use of the optical microscope and the observation of cells enabled scientists to reveal another unique feature of organisms: cellular organization*'. There is also one reference about how science affects technology in the 9th-grade textbook: '*With the advancement of biology, vaccines are constantly evolving*'.

Based on the above, NOS is poorly presented in the secondary school biology educational material. Most NOS aspects (except for the one about shared methods in science) are slightly more informed in the total material than in the taught one. It is a fact that biology teaching in lower secondary education in Greece has been degraded the recent years with the conversion of the course to a one-hour weekly lesson, which is damaging to meaningful biology teaching. Some chapters had to be excluded and among those are the introductory chapters about Biology as a Science, and the chapter of Biotechnology; both are rich in NOS references; by excluding them from the taught material, NOS introduction is impoverished.

Overall, more NOS concepts are included in the 9th grade, and fewer in the 7th and 8th grades. This is not due to pedagogical reasons, as NOS can, and some researchers believe that it should, be introduced as soon as children begin to learn about science in the kindergarten (McComas et al., 2020). Of course, some of the NOS aspects are more difficult for younger students (e.g., 'the theory/law distinction', or 'subjectivity' than others (e.g., 'need for evidence' and 'creativity') due to cognitive development, prior knowledge, and previous experiences (McComas et al., 2020). However, lower secondary education grades that concern us in our study are appropriate for the introduction of all NOS aspects. The introduction of NOS should be included in the curriculum, and it should be designed carefully to have a sequence throughout the curriculum, revisiting the same ideas in more depth at every level, in a developmentally appropriate way at all levels from the elementary to the college level (Yacoubian, 2020).

Conclusions

Secondary school biology textbooks follow the global trend (McDonald & Abd-El-Khalick, 2017) dedicating almost all their content to declarative, or else conceptual knowledge, little content to procedural knowledge (through laboratory experiments and activities), and neglect epistemic knowledge, including NOS aspects. As it has been previously reported (Kampourakis, 2017; Koumara & Plakitsi, 2020), NOS is not sufficiently depicted in the secondary school biology textbooks.

NOS references are made implicitly, with the exceptions of a few misleading explicit references to the steplike scientific method. Other than that, there are many opportunities to introduce informed views about NOS aspects, starting from the implicit textbook references that concern the importance of empirical evidence, shared methods in the scientific endeavour, the role of creativity in science, social and cultural influences, and the tentative nature of the scientific knowledge. Less frequent are the opportunities to discuss the nature of laws and theories, and science's relationship with technology. Subjectivity in science is scarce, whereas the fact that science has limits is practically absent.

Similar are the findings of global research. Unfortunately, the myth of the stepwise, single, procedural, and/or universal 'Scientific Method' seems to immortal and present in high school chemistry textbooks in the USA

(Abd-El-Khalick, et al., 2007) and in secondary school biology textbooks in Turkey (Irez, 2007) as well as in other countries (Reiff-Cox, 2020). Most references in NOS aspects, such as the social and cultural embeddedness of science, creativity, and subjectivity in science, tend to be discussed only implicitly, while theories, laws, and the tentative nature of scientific knowledge, when approached, are presented naively. Subjectivity in science and the fact that science has limits are the rarest ones (Abd-El-Khalick, et al., 2007; Campanile, et al., 2013; Irez, 2007).

In the Greek secondary school biology textbooks that we analysed, the introductory chapters, discussing biology as a science that contain explicit NOS aspects, provide a distorted image of NOS (e.g., the step-like scientific method). However, if they were taught, they could be useful to introduce discussion about NOS and, with careful guidance from the instructor, become the spark for the introduction of accurate NOS aspects. Unfortunately, according to the ministry's guidance, they are excluded from the taught material, enforcing teachers' and students' ideas that NOS aspects are not worthy to be given teaching time especially when compared to scientific content.

The remaining chapters of the textbooks contain implicit references to NOS aspects, mostly through referral to the HOS, which has been repeatedly acknowledged as a valuable way to teach NOS (Abd-El-Khalick & Lederman, 2000; Kapsala & Mavrikaki, 2020; McComas & Kampourakis, 2015). Nevertheless, for the NOS instruction to be successful, it has to be contextualised, explicit, and reflective (McComas et al., 2020). HOS references cover the first part of providing the context. It is left to the teacher to make the hidden NOS message explicit and provide their students with the opportunity to explore it, discuss it, and reflect on it. Internationally, few science textbooks provide meaningful historical discussions about the development of science ideas (McComas, et al., 2020). References to the HOS usually focus on individual achievements that give students the false impression that science moves straightforwardly and that it always has a direct impact and wide acceptance therefore creating misconceptions about how science is integrated in our lives (Abd El-Khalick & Lederman, 2000). Thus, opportunities to discuss the dynamic nature of science, the role of creativity of the human mind, the interaction between scientists, and the influence of different contexts on the development of science are lost (Lin, et al., 2010).

Moreover, implicit NOS references can be found in the laboratory activities, some of the workbook activities and the indicative activities that are proposed in the teachers' books; it is doubtful though that teachers use this material in the context of the one-hour per week lesson and the pressure to complete the syllabus in a limited time. As for the laboratory activities, it has been reported that the experiments have been designed to accept or reject hypotheses and not as a method to structure theoretical and conceptual meanings, enforcing a distorted positivist image of science (Stasinakis & Koliopoulos, 2009).

When NOS is removed from science teaching, myths about it are reproduced in science classes. Scientific knowledge is presented as unchangeable, objective facts, and the scientific procedures as 'a strict scientific method always being applicable'. Such images of science are claimed to be part of the hidden curriculum in many science classes (Hansson, 2018). Research shows that such are the students' views about science (Hansson, 2018 and references therein). It seems that although empiricism and positivism are 'dead to philosophical circles', policy makers and curriculum writers consider them very much alive (Kokkotas, 2004 as referenced in Stasinakis & Koliopoulos, 2009).

Since textbooks have so many deficiencies in the NOS content, measures regarding the teachers' training should be taken (Solaz-Portolés, 2010). To achieve effective NOS instruction, science teachers should hold several competencies including general NOS knowledge, subject matter knowledge, knowledge of learners with respect to NOS, knowledge about NOS instructional strategies, knowledge about NOS assessment, general pedagogical knowledge, and motivations and positive beliefs about teaching NOS (Nouri, et al., 2021).

In the given circumstances, for NOS to be taught, teachers are called to put in action their 'pedagogical design capacity', and like the charismatic teacher depicted in Matić's research (2019), manoeuvre through the syllabus, utilise textbooks' sources, improvise, and create productive instructional episodes to achieve their educational goals. This requires high levels of the abovementioned competencies. Unfortunately, most science teachers in Greece have not been trained either in NOS content or in NOS instruction; they are not familiar with it, and when they approach NOS, they do so intuitively, not explicitly nor reflectively, and they do not assess their students' understandings (Koumara & Plakitsi, 2020). Such an expectation from them would be unrealistic and unfair. Only when science teachers are trained can we expect them to be able to utilise the minimum available NOS content of the textbooks, to effectively teach NOS aspects.

In conclusion, when there are no direct explicit NOS references in the taught material, NOS slips to the hidden curriculum. It is taught indirectly through the implicit messages of the textbooks. Thus, too much is left on the shoulders of the biology teacher who is called to correctly interpret the implicit NOS references, to create educational material in order to make them explicit and to employ their students in a reflective dialogue about them. All this of this is with the time pressures of the syllabus that does not include NOS in its educational goals, and with almost no official training on NOS instruction. The NOS instruction ends up depending on the teacher's worldview about science, which rarely goes along with the precise NOS. This cultivates misunderstandings and feeds myths about NOS enforcing the distorted positivist image of science and leaves no room for a different, more human, and closer-to-the-truth image of science.

Limitations of the research: in this study we analysed the Greek secondary school biology textbooks along with the curriculum and the ministry's 2021/22 teaching guidelines. It would be interesting to investigate what actually is taught in the classrooms. Perhaps some teachers could highlight ideas about NOS that are absent from the textbooks.

References

Abd-El-Khalick, F., Waters, M., & Le, A. P. (2008). Representations of nature of science in high school chemistry textbooks over the past four decades. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, *45*(7), 835–855. https://doi.org/10.1002/tea.20226

Abd-El-Khalick, F., & Lederman, N. G. (2000a). The influence of history of science courses on students' views of nature of science. *Journal of Research in Science Teaching*, *37*(10), 1057–1095. https://doi.org/10.1002/1098-2736(200012)37:10<1057::AID-TEA3>3.0.CO;2-C

Allchin, D. (2011). Evaluating knowledge of the nature of (whole) science. *Science Education, 95*(3), 518–542. https://doi.org/10.1002/sce.20432

Bryman, A. (2016). *Social research methods*. Oxford University Press.

Campanile, M. F., Lederman, N. G., & Kampourakis, K. (2015). Mendelian genetics as a platform for teaching about nature of science and scientific inquiry: The value of textbooks. *Science & Education*, *24*(1), 205–225. https://doi.org/10.1007/s11191-013-9607-4

Clough, M. P. (2011). Teaching and assessing the nature of science. *The Science Teacher, 78*(6), 56–60.

Government Gazette. (2003). 304/13-03-2003. Διαθεματικό ενιαίο πλαίσιο προγραμμάτων σπουδών (Δ.Ε.Π.Π.Σ.) και αναλυτικά προγράμματα σπουδών (Α.Π.Σ.) Δημοτικού Γυμνασίου [Interdisciplinary unified framework of study programs and curricula of primary school and high school].

Government Gazette. (2016). 1640/Β΄/9-6-2016. Ωρολόγιο πρόγραμμα των μαθημάτων των Α›, Β›, Γ› τάξεων του ημερησίου γυμνασίου [Schedule of the lessons of the A›, B›, C› classes of high school].

Haagen-Schützenhöfer, C., & Joham, B. (2018). Professionalising physics teachers in doing experimental work. *Center for Educational Policy Studies Journal, 8*(1), 9–34. https://doi.org/10.26529/cepsj.333

Hansson, L. (2020). Teaching the limits of science with card-Sorting activities. In W. McComas (Ed), *Nature of science in science instruction: Rationales and strategies* (pp. 627–639). Springer. https://doi.org/10.1007/978-3-030-57239-6_34

Hansson, L. (2018). Science education, indoctrination, and the hidden curriculum. In M. Matthews (Eds.), *History, philosophy and science teaching. Science: philosophy, history and education* (pp. 283–306). Springer. https://doi.org/10.1007/978-3-319-62616-1_11

Höttecke, D., & Silva, C. C. (2011). Why implementing history and philosophy in school science education is a challenge: An analysis of obstacles. *Science and Education, 20*(3), 293–316. https://doi.org/10.1007/s11191-010-9285-4

Kampourakis, K. (2017). Nature of science representations in Greek secondary school biology textbooks. In C. V. McDonald & F. Abd-El-Khalick (Eds.), *Representations of nature of science in school science textbooks. A global perspective* (pp. 130–146). Routledge. https://doi.org/10.4324/9781315650524-6

Kampourakis, K. (2016). The "general aspects" conceptualisation as a pragmatic and effective means to introducing students to nature of science. *Journal of Research in Science Teaching*, *53*(5), 667–682. https://doi.org/10.1002/tea.21305

Koumara, A., & Plakitsi, K. (2020). The degree that nature of scientific knowledge aspects are included in the science classes of Greek high schools. *World Journal of Education*, *10*(5), 1–17. https://doi.org/10.5430/wje.v10n5p1

Irez, S. (2009). Nature of science as depicted in Turkish biology textbooks. *Science Education*, *93*(3), 422–447. https://doi.org/10.1007/978-94-007-7654-8_44

Kapsala, N., & Mavrikaki E. (2020). Storytelling as a pedagogical tool in nature of science instruction. In W. F. McComas (Ed.), *Nature of science in science instruction: rationales and strategies* (pp. 485–512). Springer. https://doi.org/10.1007/978-3-030-57239-6_27

Lederman, N. G., & Lederman, J. S. (2012). Nature of scientific knowledge and scientific inquiry: Building instructional capacity through professional development. In B. J. Fraser et al. (Eds.), *Second international handbook of science education* (pp. 335–359). Springer.

Lederman, N. G. (2013). Nature of science: Past, present, and future. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 831–880). Routledge.

Lin, C.-Y., Cheng, J.-H., & Chang, W.-H. (2010). Making science vivid: Using a historical episodes map. *International Journal of Science Education, 32*(18), 2521–2531. https://doi.org/10.1080/09500691003746015

Marniok, K., & Reiners, C. S. (2017). Representations of nature of science in German school chemistry textbooks. In C. V. McDonald & F. Abd-El-Khalick (Eds.), *Representations of nature of science in school science textbooks* (pp. 201–214). Routledge. https://doi.org/10.4324/9781315650524-11

Matić, L. J. (2019). The teacher as a lesson designer. *Center for Educational Policy Studies Journal, 9*(2), 139–160. https://doi.org/10.26529/cepsj.722

Mavrikaki, E., Gouvra, M., & Kampouri, A. (2007a). *Βιολογία Α΄ Γυμνασίου. Αθήνα: Οργανισμός Εκδόσεως Διδακτικών Βιβλίων* [7th grade biology students' textbook]. Textbook Publishing Organization.

http://ebooks.edu.gr/ebooks/v2/books-pdf.jsp?handle=8547/106

Mavrikaki, E., Gouvra, M., & Kampouri, A. (2007b). *Βιολογία Α΄ Γυμνασίου Τετράδιο Εργασιών. Αθήνα: Οργανισμός Εκδόσεως Διδακτικών Βιβλίων* [7th grade biology students' workbook]. Textbook Publishing Organization.

Mavrikaki, E., Gouvra, M., & Kampouri, A. (2007c). *Βιολογία Α΄ Γυμνασίου Εργαστηριακός Οδηγός. Αθήνα: Οργανισμός Εκδόσεως Διδακτικών Βιβλίων* [7th grade biology students' lab-book]. Textbook Publishing Organization.

Mavrikaki, E., Gouvra, M., & Kampouri, A. (2007d). *Βιολογία Α΄ Γυμνασίου Βιβλίο Εκπαιδευτικού. Αθήνα: Οργανισμός Εκδόσεως Διδακτικών Βιβλίων* [7th grade biology teachers' book]. Textbook Publishing Organization. http://ebooks.edu.gr/ebooks/handle/8547/4927

Mavrikaki, E., Gouvra, M., & Kampouri, A. (2007e). *Βιολογία Γ΄ Γυμνασίου. Αθήνα: Οργανισμός Εκδόσεως Διδακτικών Βιβλίων* [9th grade biology students' textbook]. Textbook Publishing Organization. http://ebooks.edu.gr/ebooks/v2/books-pdf.jsp?handle=8547/108

Mavrikaki, E., Gouvra, M., & Kampouri, A. (2007f). *Βιολογία Γ΄ Γυμνασίου Τετράδιο Εργασιών. Αθήνα: Οργανισμός Εκδόσεως Διδακτικών Βιβλίων* [9th grade biology students' workbook]. Textbook Publishing Organization.

Mavrikaki, E., Gouvra, M., & Kampouri, A. (2007g). *Βιολογία Γ΄ Γυμνασίου Εργαστηριακός Οδηγός. Αθήνα: Οργανισμός Εκδόσεως Διδακτικών Βιβλίων* [9th grade biology students' lab-book]. Textbook Publishing Organization.

Mavrikaki, E., Gouvra, M., & Kampouri, A. (2007h). *Βιολογία Γ΄ Γυμνασίου Βιβλίο Εκπαιδευτικού. Αθήνα: Οργανισμός Εκδόσεως Διδακτικών Βιβλίων.* [9th grade biology students' textbook]. Textbook Publishing Organization. http://www.pi-schools.gr/books/gymnasio/biologia_a/kath/kath_1_150.pdf

McComas, W. F., Clough, M. P., & Nouri, N. (2020). Nature of science and classroom practice: A review of the literature with implications for effective NOS instruction. In W. F. McCommas (Ed.), *Nature of science in science instruction: Rationales and strategies* (pp. 67–111). Springer. https://doi.org/10.1007/978-3-030-57239-6_4

McComas, W. F. (2020). Principal elements of nature of science: Informing science teaching while dispelling the myths. In W. F. McCommas (Ed.), *Nature of science in science instruction: Rationales and strategies* (pp. 35–65). Springer https://doi.org/10.1007/978-3-030-57239-6_3

McComas, W. F., & Kampourakis, K. (2015). Using the history of biology, chemistry, geology, and physics to illustrate general aspects of nature of science. *Review of Science, Mathematics and ICT Education*, *9*(1), 47–76. https://doi.org/10.26220/rev.2240

McDonald, C. V., & Abd-El-Khalick, F. (2017). Representations of nature of science in school science textbooks. In C. V. McDonald & F. Abd-El-Khalick (Eds.), *Representations of nature of science in school science textbooks. A global perspective* (pp. 1–19). Routledge. https://doi.org/10.4324/9781315650524-1

Mesci, G., & Schwartz, R. (2016). Changing preservice science teachers' views of nature of science: Why some conceptions may be more easily altered than others. *Research in Science Education, 47*(2), 329–351. https://doi.org/10.1007/s11165-015-9503-9.

Ministry of Education and Religious Affairs (2021). *109009/Δ2/07-09-2021 Οδηγίες για τη διδασκαλία του μαθήματος της Βιολογίας του Ημερησίου και του Εσπερινού Γυμνασίου για το σχολικό έτος 2021-2022* [Instructions for teaching the course of biology of the day and evening high school for the school year 2021-2022].

Niaz, M. (2014). Science textbooks: The role of history and philosophy of science. In M. R. Matthews (ed.) *International handbook of research in history, philosophy and science teaching* (pp. 1411–1441). Springer.

Nouri, N., Saberi, M., Mccomas, W., & Mohammadi, M. (2021). Proposed teacher competencies to support effective nature of science instruction: A meta-synthesis of the literature. *Journal of Science Teacher Education, 32*(6), 1–24. https://doi.org/10.1080/1046560X.2020.1871206.

NSTA Board of Directors. (2020). *Nature of science/NSTA*. https://www.nsta.org/nstas-official-positions/nature-science

O'Connor, C., & Joffe, H. (2020). Intercoder Reliability in Qualitative Research: Debates and Practical Guidelines. *International Journal of Qualitative Methods, 19*, 1–13. https://doi.org/10.1177/1609406919899220

Osborne, J., Collins, S., Ratcliffe, M., Millar, R., & Duschl, R. (2003). What "ideas-about-science" should be taught in school science? A Delphi study of the expert community. *Journal of research in science teaching, 40*(7), 692–720. https://doi.org/10.1002/tea.10105

Popper, K. (2005). *The logic of scientific discovery*. Routledge.

Reiff-Cox, R. (2020). Exchanging the myth of a step-by-step scientific method for a more authentic description of inquiry in practice. In W. F. McCommas (Ed.), *Nature of Science in Science Instruction* (pp. 127–139). Springer. https://doi.org/10.1007/978-3-030-57239-6_6

Smith, M. U., & Scharmann, L. C. (1999). Defining versus describing the nature of science: A pragmatic analysis for classroom teachers and science educators. *Science Education*, *83*(4), 493–509. https://doi.org/10.1002/(SICI)1098-237X(199907)83:4<493::AID-SCE6>3.0.CO;2-U

Slapničar, M. (2014). Khine Swe, M. (Ed.) (2013) Critical analysis of science textbooks: Evaluating instructional effectiveness. *Center for Educational Policy Studies Journal, 4*(1), 137–141.

Solaz-Portolés, J. J. (2010). La naturaleza de la ciencia y libros de texto de ciencias: una revision. *Educación* [The nature of science and science textbooks: A review. *Education] XX1, 13*(1), 65–80. https://doi.org/10.5944/educxx1.13.1.277

Song, J., Chun, J., & Na, J. (2021). Why people trust something other than science. *Science & Education, 30*(6), 1387–1419. https://doi.org/10.1007/s11191-021-00243-w

Stasinakis, P. K., & Koliopoulos, D. (2009). *Ανάλυση εγχειριδίων βιολογίας στη δευτεροβάθμια εκπαίδευση: Η περίπτωση της έννοιας της θρέψης φυτών και ζώων. Θέματα Επιστημών και Τεχνολογίας στην Εκπαίδευση* [Analysis of biology textbooks in secondary education: The case of the concept of plant and animal nutrition]. *Issues of Science and Technology in Education, 2*(1–2), 103–125. http://earthlab.uoi.gr/ojs/thete/index.php/thete/article/view/25

Taber, K. S. (2021). The challenge to educational reforms during a global emergency: The case of progressive science education. *Center for Educational Policy Studies Journal, 11*(Sp. Issue), 67–87.

https://doi.org/10.26529/cepsj.1109

Weisberg, D. S., Landrum, A. R., Hamilton, J., & Weisberg, M. (2020). Knowledge about the nature of science increases public acceptance of science regardless of identity factors. *Public Understanding of Science, 30*(2), 120–138. https://doi.org/10.1177/0963662520977700

Yacoubian, H. A. (2020). Teaching nature of science through a critical thinking approach. In W. F. McComas (Ed.), *Nature of science in science instruction: rationales and strategies* (pp. 199–212). Springer. https://doi.org/10.1007/978-3-030-57239-6_10

Biographical note

Nausica Kapsala is a biologist and a storyteller. She is currently a secondary education biology teacher (2nd Lyceum of Elefsina) and an external teaching and research fellow of the Laboratory of Science Teaching & Epistemology and Educational Technology of the Faculty of Primary Education of the National and Kapodistrian University of Athens, Greece. She has authored and coauthored many articles or chapters in Greek and international journals, conference proceedings, and edited volumes. She holds a master's degree in Modern trends in the teaching of biology with the aid of new technologies and a PhD in Biology teaching. Her research interests lie in storytelling as an educational tool in biology teaching, nature of science, history of science and evolution understanding.

Apostolia (Lia) Galani is Associate Professor in Didactics of Geography at the National and Kapodistrian University of Athens, Greece. She has authored or coauthored many articles or chapters in Greek and international journals or books, and she is the author-coauthor of 15 books, including the Greek primary and lower secondary education geography textbooks. Her research interests focus on socio-scientific issues, on the educational approach of environmental issues with innovative methods (e.g., remote sensing), on the use of ICT in school Geography.

Evangelia Mavrikaki, holds a PhD in biology and she is a Professor and a member of the Laboratory of Science Teaching & Epistemology and Educational Technology of the Faculty of Primary Education of the National and Kapodistrian University of Athens, Greece. She has worked as an academic (lecturer, Assistant Professor) in the Aristotle University of Thessaloniki and the University of Western Macedonia before her current position. She has been a member of the board of reviewers/editorial board of several journals and she has authored or coauthored many articles or chapters in Greek and International Journals or books, and she is the author-coauthor of 15 books including the official Greek Gymnasium (lower secondary education) biology textbooks. She has served for four years as an elected member of the executive board - in charge for Educational Issues – of the Pan-Hellenic Association of BioScientists and has served as President of EvoKE (Evolutionary Knowledge for Everyone).

doi: 10.26529/cepsj.1287

Theorising Textbook Adaptation in English Language Teaching

Stefan Rathert*[1] and Neşe Cabaroğlu[2]

Even though textbooks are a central component of the daily instructional practice of English language teachers, relatively little research has been conducted on how teachers actually use (i.e., adapt) textbooks in the classroom. This gap is aggravated by the fact that the terminology proposed in the literature to analyse teachers' textbook use is characterised by inconsistencies because different terms denote the same adaptation techniques, identical terms refer to different techniques and suggested frameworks differ in the fact that comparable techniques are allocated to different categories. This inconsistency mirrors the difficulty of operationally defining adaptation techniques, as the terms used may be unambiguous but vague and therefore of reduced explanatory power or more specific but potentially unreliable because an adaptation may be matched to different terms given the complexity of a particular textbook adaptation. Discussing these aspects, this paper proposes a research-informed framework to contribute to a systematic description of textbook adaptation in foreign and second language teaching. Examining adaptation as a process, it is argued that teachers, driven by an identified or felt mismatch between the textbook and other factors (e.g., school facilities, the learners, teacher cognition, course requirements, or outdatedness of the materials), engage in adaptation based on principles (i.e., ideas about best practices, by making changes to the content, the language and/or the sequence of activities offered by the textbook authors). Even though related to English language teaching, this paper does not exclusively inform this context as it offers implications for research on textbook use in other disciplines.

Keywords: adaptation, adaptation techniques, coursebook, foreign language teaching, textbook

1 *Corresponding Author. Faculty of Education, Kahramanmaraş Sütçü İmam University, Turkey; strathert@gmail.com.
2 Faculty of Education, Çukurova University, Adana, Turkey.

Razmislek o prilagoditvi učbenika pri poučevanje angleščine

Stefan Rathert in Neşe Cabaroğlu

Čeprav so učbeniki osnovna sestavina vsakodnevne učne prakse učiteljev angleškega jezika, je bilo opravljeno zelo malo raziskav glede tega, kako učitelji dejansko uporabljajo (tj. prilagodijo) učbenike v učilnici. Ta vrzel je še poslabšana z vidika predlagane terminologije, ki se nanaša na učiteljevo uporabo učbenika, saj je ta prepredena z nedoslednostmi, pri katerih se na primer različni izrazi sklicujejo na isto prilagoditev tehnik, identične oznake označujejo različne tehnike in predlagana ogrodja se razlikujejo v tem, da primerljive tehnike dodeljujejo različnim kategorijam. Ta nedoslednost zrcali težavnost operativnega poimenovanja prilagoditvenih tehnik, saj so nekateri izrazi sicer rabljeni nedvoumno, a vseeno nedoločno, kar posledično niža razlagalno moč, pri čemer pa tudi bolj specifični izrazi nazadnje postanejo potencialno nezanesljivi, saj je prilagoditev lahko vezana na številne termine glede na težavnost posamezne adaptacije učbenika. Pri razpravi o teh vidikih prispevek predlaga raziskovalno-informirano ogrodje kot prispevek k sistematičnemu opisu prilagoditev učbenikov pri poučevanju tujega oz. drugega jezika. Pri analizi adaptacij kot procesa argumentiramo, da se učitelji, gnani s strani zaznanega ali občutenega neujemanja med učbenikom in drugimi faktorji (npr. šolske infrastrukture, učečih se, učiteljevega védenja, zahtev predmeta, zastaranosti gradiva), lotijo prilagoditev na osnovi načel (tj. zamisli o najboljši praksi, prek sprememb vsebine, jezika in/ali sosledja aktivnosti, ki jih ponujajo avtorji učbenika). Čeprav je povezano s poučevanjem angleščine, članek ne zadeva ekskluzivno le tega konteksta, ampak ponuja raziskovalne možnosti glede rabe učbenikov tudi znotraj drugih področij.

Ključne besede: prilagoditev, prilagoditvene tehnike, poučevanje tujega jezika, učbenik

Textbooks in language teaching: the need for adaptation

The significance of textbooks as the main medium in English language teaching (ELT) (Mishan, 2021; Tomlinson & Masuhara, 2018) as well as in other subject areas (Smart et al., 2020) has been widely recognised. As a convenient tool, textbooks allow teachers to plan and administer their lessons by using them as a guideline and resource; they expose learners to samples of the target language and help them preview and review lesson content that is presented in a visually appealing way; moreover, they encourage administrators to base courses on the sequence given in textbooks so that they function as coursebooks (Gray, 2016; McGrath, 2013). Based on these advantages, textbooks appear to be an indispensable (and for publishers highly lucrative) component of institutional language teaching as their utilisation fulfils learners' and teachers' expectations, thus they provide courses with 'face validity' (Mishan, 2021, p. 2; Vitta, 2021), and serve as the actual curriculum in numerous contexts (Garton & Graves, 2014): without a coursebook, 'a program may have no central core and learners may not receive a syllabus that has been systematically planned and developed' (Richards, 2001, p. 1).

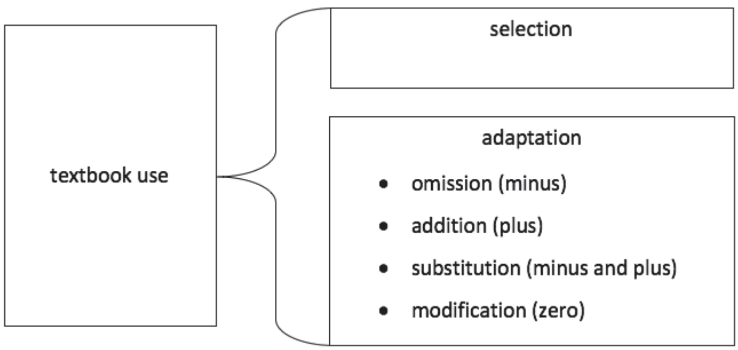
For language teaching, an important distinction is made between global, localised, and local textbooks (López-Barrios & Villanueva de Debat, 2014). Global textbooks are produced for learners worldwide by publishers usually located in countries where the target language is spoken, while localised (i.e., modified global textbooks) and local textbooks are designed for learners situated in a specific location where the target language is learned but not the environmental language. They include materials that engage learners in the comparison of the target language and culture with their own language and culture and may address specific incentives to learn the target language. Local textbooks are usually issued by publishers located in the countries where the language is learned. They comply with curricular regulations enacted by state authorities to receive permission to be used in state schools and are comparatively inexpensive for parents and state schools; due to these production conditions, locally produced textbooks usually do not produce innovative methodological approaches or startling content (Kovač & Šebart, 2019). Especially (but not exclusively) global textbooks are likely to cause the need for adaptation, (i.e., modifications applied to the textbook materials) because of the distance between textbook authors and users (Tomlinson & Masuhara, 2018) and the fact that they 'are written for everyone and therefore for no one' (McGrath, 2013, p. 59). Because textbook authors and publishers cannot consider the numerous specific conditions of varying local contexts where the textbooks are used, the textbook materials are potentially incongruent with the teaching and learning environment (Madsen & Bowen, 1978). Mismatches between what the material offers and the conditions of the learning context (e.g., learner expectations or proficiency levels, teacher beliefs about best teaching practices, school culture and infrastructure, institutional regulations as evidenced in curricula, examinations, or expected methods) force teachers to adapt textbooks (Ariyan & Pavlova, 2019; Macalister, 2016a; McDonough et al., 2013). Furthermore, adaptation is considered a sign of professionalism because it lessens the danger that teachers are patronised by textbooks given their potential 'to exercise a tyrannical function as the arbiter of course content and teaching methods' (Cunningsworth, 1995, p. 7), to deskill teachers (Akbari, 2008; Littlejohn, 2012; Pouromid & Amerian, 2018; Rathert & Cabaroğlu, 2021) and to reinforce transmission-based teaching (Smart et al., 2020).

Research has recently started to pay increasing attention to the use and adaptation of textbooks and other learning materials in language teaching. In relation to this, materials as a field of inquiry currently appear under-theorised both in terms of what learning materials actually are and how materials function as both objects used by teachers and learners and subjects shaping classroom interaction (Guerrettaz et al., 2021; Harwood, 2021). Specifically related to textbook use, there is inconsistency in operationalised terms to analyse and describe why and how teachers and learners engage in adaptation (McGrath, 2013, p. 63; Tomlinson & Masuhara, 2018, p. 105). Recognising the wide spectrum of learning materials and the role of teachers, learners and administrators as users and adapters of materials (Guerrettaz et al., 2021; Harwood, 2021), this paper focuses on textbook adaptation carried out by teachers and identifies techniques employed, principles, reasons and purposes driving teachers to engage in textbook adaptation. Additionally, procedures are reported that teachers can follow to arrive at reflective adaptation approaches. In other words, this contribution aims to theorise textbook adaptation by defining key issues surrounding textbook adaptation. Explicating this issue, our contribution aims at encouraging researchers who are interested in examining this field of study and practitioners to raise their awareness of the complexity of adaptation and to help them reflect on the utilisation of this central medium in instructional practice. Based on scholarly monographs and research conducted in ELT, we hope that the considerations and suggestions are informative for other educational domains.

Techniques in textbook adaptation

Teachers can select textbook materials as they are or adapt them to varying extents ranging from slight modifications without deviating from the guidance offered to use the textbook as a resource book when developing procedures not foreseen by textbook authors (Ur, 2015). The potential utilisations of textbooks by teachers are shown in Figure 1.

Figure 1
Components of textbook use



For adaptation purposes, teachers can benefit from the techniques of omission, addition, substitution, and modification to engage in adaptation. These adaptations may be directed toward the language of texts, the content that is conveyed through the language, or the activities and procedures in which the textbook aims to engage learners. Tomlinson and Masuhara (2004) reasonably point to the reduction or increase of textbook materials, and replacement or modification without change of material amount (indicated through *plus*, *minus*, and *zero*) as criteria to classify adaptation techniques. Table 1, extending an overview by McGrath (2013, p. 64), allocates the terminology used by different authors in relation to the four basic adaptation techniques.

Table 1

Adaptation techniques suggested in the literature

	Omission	Addition	Substitution	Modification
Cunningsworth (1995, p. 136)	leaving out some parts of the material	adding material	replacing material with something more suitable	changing the published material
Harmer (2007, pp. 182–183)	omit		replace (the book completely)	add, rewrite, replace activities, reorder, reduce
Islam and Mares (2003)	deleting; subtracting and abridging	adding including extending and expanding	replacing materials	simplifying, reordering
Maley (2011)	omission	addition	replacement	rewriting, modification, reduction, extension, reordering, branching
McDonough et al. (2013, pp. 69–78)	deleting or omitting including subtracting and abridging	adding including expanding and extending		modifying including rewriting and restructuring, simplifying, reordering
McGrath (2013, pp. 139–147; 2016)	omission	addition including extemporisation, extension, exploitation	change including reordering, replacement, simplification, localisation, complexification, conversion (converting a text into a play)	
Richards (2017, p. 251)	deleting content	adding content addressing omissions extending tasks		modifying or reorganising content, modifying tasks
Tomlinson and Masuhara (2004, as cited in McGrath 2013, p. 64)	minus: delete, subtract, reduce	plus: add expand	zero: modify, replace, reorganise, resequence, convert	

The overview in Table 1 indicates a lack of standardised terms and inconsistencies. Some of the terms are apparently synonymous (e.g., *leaving out, omit, deleting*), but other terms and classifications appear to be problematic. For example, McGrath (2016) defines *exploitation* as 'creative use of what is already there (e.g., text, visual, activity) to serve a purpose which is *additional* to that foreseen by the textbook writer' (p. 71; emphasis in the original); this corresponds to *branching* denoting a technique 'to add options to the existing activity or to suggest alternative pathways through the activities' proposed by

Maley as a form of modification (2011, p. 382). According to Harmer (2007), *reducing* is a kind of modification, which is reasonable because the appearance of the material is changed, while McDonough et al. (2013), equally plausibly, subsume the corresponding techniques of *subtracting* and *abridging* under *omission* emphasising the reduction in the amount of material. Another example of inconsistent and confusing terminology, given by McDonough et al. (2013), is the definition of *rewriting* as an attempt to make materials more communicative and learner-centred, which they exemplify with the instance of a teacher who rewrites a reading text and delivers it orally to generate an extra listening practice for the learners. However, there is some overlap to *restructuring* referring 'essentially to a "modality change"' (p. 74), and it is not completely convincing to introduce *simplifying* as another subcategory denoting language modification: Simplification may be more appropriately comprehended as a principle guiding adaptation (McGrath, 2013; Tomlinson & Masuhara, 2018). To execute a fair evaluation, it should be noted that the scholarly publications considered in the survey in Table 1 differ in that some of them are specifically related to textbook research (e.g., McDonough et al., 2013; McGrath, 2013, 2016; Tomlinson & Masuhara, 2004), while other books allude to the topic as they are introductions to ELT (Harmer, 2007) or deal with curriculum development in language teaching (Richards, 2017).

The fact that two core contributions to the field (Mishan & Timmis, 2015; Tomlinson & Masuhara, 2018) abstain from presenting their own frameworks may be indicative of the difficulty of systematising adaptation techniques. To address the unsatisfactory inconsistency in terminology, we propose a research-informed framework that was developed and tested in a study examining the textbook utilisations of two English language teachers (Rathert & Cabaroğlu, 2021). The framework is shown in Table 2.

Table 2

Research-informed framework of adaptation techniques

Main techniques	Sub-techniques	Descriptions and examples
omission (reduction in material amount)	subtracting	quantitative reduction without changing procedures (e.g., number of gap-fill sentences in exercise is reduced)
	abridging	qualitative reduction with change of procedures (e.g., prereading questions are skipped)
addition (increase in material amount or enhanced function)	extemporisation	(often ad hoc) explanation to address perceived or anticipated challenges (paraphrasing of instruction)
	exploiting	adding a new purpose to materials in the textbook (e.g., learners use reading comprehension questions to generate a text before reading the text in the textbook)
	extending	addition of materials without changing procedures (e.g., number of gap-fill sentences in exercise is increased)
	expanding	addition of activity or material leading to procedural change (e.g., after answering comprehension questions, the learners create their own questions)
	supplementing	adding a component that leads into the textbook material without changing it (e.g., playing hangman to preview vocabulary in the textbook unit)
substituting		replacing material in the textbook with other material for the same or a similar purpose
modification (change of language, procedures, or content)	restructuring	changing the task procedures or modality (e.g., a listening text is delivered as a reading text due to the lack of technical equipment or pair work changed into group work)
	rewriting	changing the vocabulary, grammar or content in texts or rubrics (e.g., reference to alcoholic drinks is removed from a text because drinking alcohol is not considered acceptable in the cultural context of the teaching/learning environment)
	reordering	textbook or text components are presented in a different order (e.g., order of exercises on a textbook page)

Remaining with the four main adaptation techniques, the framework attempts not only to consider forms of minor and major adaptation but also to find a balance between establishing unambiguous but potentially vague categories and more specific categories that may be reduced in their validity because adaptations may fall into more than category given 'all the combinations and permutations' (McDonough et al., 2013, p. 76) in textbook adaptation. The following remarks will clarify and highlight some aspects of the framework.

The main technique *omission* with its sub-techniques is primarily based on McDonough et al.'s (2013) classification. The category *addition* combines the systematic accounts presented by McGrath (2013, 2016) and McDonough et al. (2013). The sub-techniques *extemporisation*, *extending*, *expanding*, and *exploiting* may be best distinguished by observing that *extending* and *expanding* change the 'appearance' of the material in terms of length while *extemporising* and *exploiting* add clarification and a new purpose to the material. Differently from *expanding*, *supplementation* has no direct impact on the textbook material. McGrath (2013) argues that supplementation does not count as adaptation, claiming that 'supplementation involves introducing something new' (p. 72). His own example of supplementation, a '*presentation* activity (based on a new topic – preparatory; books closed)' (p. 145, emphasis in the original), however, suggests that supplementation counts as adaptation because the supplementary material is connected to a component in the textbook as it prepares learners for the textbook component. To give another example, supplementing an image to pre-teach vocabulary in a reading text in the textbook, does count as adaptation because the learners' cognitive load while doing the reading tasks is lowered so that they face fewer difficulties in comprehending the text or engaging in follow-up tasks based on the reading text. Supplementation can also address the content of a textbook component that has already been dealt with in a lesson. A teacher may supplement, for instance, a text that is thematically unrelated to the content of a unit in the textbook but contains examples of a grammar point introduced in the unit to give the learners extra practice.

A significant criterion to identify a realised adaptation as *substitution* is that the replaced and replacing components serve approximately the same purpose. *Modification* draws on the terminology employed by McDonough et al. (2013), but the term is differently defined in our framework, and the corresponding sub-techniques are related differently to each other. We do not follow McDonough et al.'s (2013) definition of *rewriting* based on the more common understanding of the term according to which this technique aims at the change of the language or content of a material, possibly in combination. *Restructuring* addresses a change in modality (i.e., a text is used to practice a different language skill than in the textbook material), or in classroom management. In particular, we consider *simplification* as a principle in line with McGrath (2013; see the next section).

Principles

From a pedagogical standpoint, textbook adaptation should be informed by overarching considerations and guidelines that are beyond immediate purposes arising from specific reasons (McGrath, 2013, p. 66). To label such considerations and guidelines, the term *principles* (used by McDonough et al., 2013, p. 69; McGrath, 2013, p. 66; Tomlinson & Masuhara, 2018, p. 108) referring to 'research and theory about best practices in language teaching and learning' (Macalister, 2016b, p. 44) has been suggested.

The most original contributions to delineate such principles were given by Islam and Mares (2003), McGrath (2013) and Tomlinson and Masuhara (2018). Table 3 summarises their accounts.

Table 3

Principles guiding adaptation offered in the literature

Islam & Mares (2003, pp. 89–90)	McGrath (2013, pp. 66–70)	Tomlinson & Masuhara (2018, p. 108)
– personalise – individualise – localise – modernise – add a real choice – cater for all sensory learner styles – provide for more learner autonomy – encourage higher-level cognitive skills – make the language input more accessible* – make the language input more engaging	– localisation – modernisation – personalisation – individualisation – humanising – simplification/ complexification/ differentiation – variety	– match the needs of target learners – match the wants of target learners – make relevant connections with learners' lives – stimulate affective engagement – stimulate cognitive engagement – provide achievable challenges – provide exposure to language in use – provide opportunities to communicate in L2 – provide opportunities for learners to notice and make discoveries about language use – provide enough varied recycling

Note. *Listed by the authors, but not explicated in the text.

While each of the three sources has things in common (e.g., *personalise/personalisation/make relevant connections with learners' lives*), Tomlinson and Masuhara (2018) more strongly emphasise the need to integrate validated insights gained in second language acquisition (SLA) research into textbook production (cf. Macalister, 2016b). McGrath's notion of *humanising* is a very broad term entailing connecting materials to learners' lives and serving their intellectual, aesthetic, and emotional needs (McGrath, 2013, p. 69). The need to humanise textbooks is grounded in the presupposition that textbooks with their texts, tasks, and activities suit learner needs and will lead to intended outcomes irrespective of the context they are used in. However, local or individual differences such as learner needs and interests, their previous knowledge, gained abilities or learning styles may guide teachers to engage in some kind of adaptation to enhance the relevance, attractiveness or complexity of textbook material (Maley, 2018). Tomlinson (2015, 2018) developed and collected a variety of adaptation ideas to make materials a better match for the learners. Some examples are presented here:

- closed questions (requiring one correct answer) are turned into open questions by, for instance, asking learners to justify their answer (*Explain your answer.*) or to evaluate the content (*Do you think it is a good idea…?*);
- learners invent interviews with characters from a textbook reading or listening text;
- before reading a textbook text, learners write a text based on the comprehension questions in the textbook and then compare their text with the text in the textbook;
- the teacher reads a text in the textbook aloud in a dramatic manner and the learners act it out;
- learners chant out a drill in different voices, imitating, for instance, a young child, an old man, or an angry person.

Such activities along with the integration of music, dance, art, or drama provide learners with sensory experiences and address their kinaesthetic or aesthetic preferences, intensify engagement and lead to deeper cognitive processing (Tomlinson, 2018; cf. Maley, 2018).

An obvious example of an activity that engages learners neither affectively nor cognitively is in the language textbooks' frequently employed *textually explicit comprehension question* activity (Freeman, 2014). For example, a question may read 'What do they have for breakfast?' and the text 'They have eggs and coffee for breakfast'. Because the wording in the question matches the wording in the text, this activity 'simply involves surface recognition' (Tomlinson, 2018, p. 24) and is, therefore, unlikely to facilitate language learning. In the example given, learners do not even have to understand the meaning of the words 'breakfast', 'eggs', or 'coffee' to answer the question correctly. An adaptation to make the question cognitively more activating would be a change of the wording, for example, 'What do they eat in the morning?' Asking the learners to compare the breakfast habits described in the text with their own habits would enhance the relevance of the reading activity for the learners.

In order to compare and synthesise the principles listed in Table 3, we suggest an identification of the main foci addressed. Our reorganisation distinguishes between principles mainly associated with the learners and their personal dispositions, the learning process as contributing to SLA, and the textbook itself.

Following our categorisation, adaptation is associated with the learners specifically addressing:

- their personal needs and interests by selecting relevant content (personalising);
- their individual learning styles, strengths and weaknesses by integrating activities that suit, for example, kinaesthetic or aesthetic preferences or aim at differentiation through simplification or complexification (individualisation);
- their geographic location/cultural background or experienced (thus expected) forms of instruction (localisation).

Adaptation may consider SLA research by employing texts and activities that

- are cognitively challenging and compelling;
- are affectively engaging;
- increase learner self-efficacy;
- expose learners to authentic language and tasks (i.e., language examples and activities referring to language use outside the classroom);
- build communicative competence;
- are informed by the principles of discovery learning.

Adaptation may be a response to the quality of the textbook when

- its content, language, or methodological approach is outdated (modernisation);
- it is characterised by repetitiveness or a lack of variety in activities or linguistic input.

Admittedly, the principles are overlapping. For instance, the provision of cognitively challenging and compelling materials can be potentially realised via personalisation and individualisation. It should also be noted that a specific instance of an adaptation is likely to be driven by one or two principles but not by all principles shown. Putting it differently, teachers are likely to be informed by some of the principles when they engage in an adaptation that is caused by the perceived incongruence of the textbook material with the learning context. In line with this, Tomlinson and Masuhara (2018, p. 108) remark that a principled approach to adaptation is based on the identification of deficiencies in the textbook. For instance, if content in the textbook is outdated (e.g., a reading text dealing with a technological novelty at the time of textbook production that is no longer a novelty when it is used in class), the teacher may decide to substitute the text based on the principle of modernisation.

Reasons and Purposes

The example given at the end of the previous section shows that principles, by nature, point to specific reasons and purposes to adapt textbook materials. Reasons and purposes are intertwined as they are 'like a flip of a coin. The former focuses on what needs improving, the latter targets improving' (Tomlinson & Masuhara, 2018, p. 102). In the example of the outdated textbook content (the reason), a teacher may substitute the text in the textbook with another text in order to avoid material that is likely to have a demotivating effect on learners because they perceive it as irrelevant (the purpose). Menkabu and Harwood (2014) report on an English teacher who delivers instruction in medical English: The teacher omits textbook components because her content knowledge is insufficient (the reason). One may assume that the teacher does not want to expose their learners to erroneous information or embarrass herself in the class by revealing her limited content knowledge (the purpose), but this conclusion entails some speculation.

To lift these and other examples reported in the literature to a more explanatory level, the distinction between external forces including curriculum, assessment and methodology regulations set by authorities (e.g., ministries, school districts), internal forces such as teacher beliefs about teaching and learning or received teacher education and situated forces encompassing expectations of school authorities, parents, colleagues, and learners is a useful starting point to account for teacher decisions on instructional practices including textbook use (Zheng & Davison, 2008, p. 172). More specifically related to how language teachers adapt textbooks, Tomlinson and Masuhara (2018, pp. 102–104), based on a concise review of case studies on textbook use, identified five factors that impact teachers when they utilise textbooks in their instructional practices:

- the national, regional, institutional and cultural teaching environment (e.g., curricula, examinations, school culture, views about appropriate content);
- the learners (e.g., age, learner biographies, learning styles, interests, incentives for learning);
- the teachers (e.g., personalities, belief systems, teaching styles, levels of teacher autonomy, educational background and professional experience);

- the immediate course and lesson context (e.g., objectives, syllabi, the time/day of a lesson);
- the textbook (e.g., outdatedness or modernity, methodological approach favouring a deductive or inductive approach to language learning, presentation/explanation of language features).

Zheng and Davison (2008) and Tomlinson and Masuhara (2018) emphasise that these factors do not exert an effect in isolation. Indeed, textbook adaptation is realised in a complex interplay of factors, may be realised intuitively or unconsciously (Islam & Mares, 2003), or even in 'a haphazard way' (Loh & Renandya, 2016, p. 107). Studies documenting that textbook adaptation by teachers may undermine validated principles of language teaching (Abdel Latif, 2017; Seferaj, 2014) or may be an outcome of routinised behaviour (Menkabu & Harwood, 2014; Rathert & Cabaroğlu, 2021) underline the need to inform teachers and teacher educators about procedures to identify needs for adaptation and steps to implement adaptation in a more systematic way.

Procedures

We have seen that textbook adaptation aims at removing mismatches between the textbook material and other factors. It is a process in which teachers employ (or at least should employ) specific techniques based on principles as a response to an identified problem. However, textbook adaptation is not necessarily an outcome of a reflective process: teachers may not be aware of their own adaptations (Menkabu & Harwood, 2014) and beliefs about one's own textbook adaptation may not reflect actual adaptation (Tasseron, 2017). Teachers may engage in adaptations that are not based on pedagogic concerns but serve pragmatic purposes: Rathert and Cabaroğlu (2021), for example, show that teachers may simplify cognitively engaging textbook activities in order not to fall behind the institutional schedule. They also provide evidence that a strictly structured sequence of textbook tasks encourages teachers to follow the procedures prescribed and prevents them from reflecting on how to enhance the potential of textbooks as a means to facilitate learning through adaptation. It has been argued that particularly global ELT textbooks with their 'recurring structure, ensuring predictability across the materials as a whole' (Littlejohn, 2012, p. 291) contribute to the standardisation of teaching and the deskilling of teachers. These effects are likely to materialise in overreliance on textbooks along with the avoidance of principled adaptations that reflect the learning context.

To arrive at informed and target-oriented adaptations, Cunningsworth (1995, p. 137) calls for an examination of contents, topics, methods, and unit objectives to evaluate whether they are congruent with the teaching context. A more detailed seven-step procedure was proposed by Tomlinson and Masuhara (2004, as cited in Tomlinson & Masuhara, 2018, p. 104). Starting from the development of the teaching context by evaluating, for instance, learner needs, course objectives or school equipment in order to base adaptations on accurate assumptions, reasons to engage in adaptation are identified and objectives informed by the principles explained above are formulated. The adaptation of the textbook material is then realised, and adaptations may be revised in the light of classroom experiences, for example, when the same content is taught in a parallel class the following day. This cyclical procedure is idealised, and teachers will not have to follow the steps to the letter. Indeed, Tomlinson and Masuhara (2018) find it noteworthy that textbook adaptation is

> an intuitive, organic, dynamic but principled creative process that is stimulated by the teachers' motivation to provide the best teaching input and approaches for specific learners in a specific context with specific learning objectives. We would discourage the prescribed use of techniques or advice as this could be counteractive or even damaging to teachers' creativity. (p. 105)

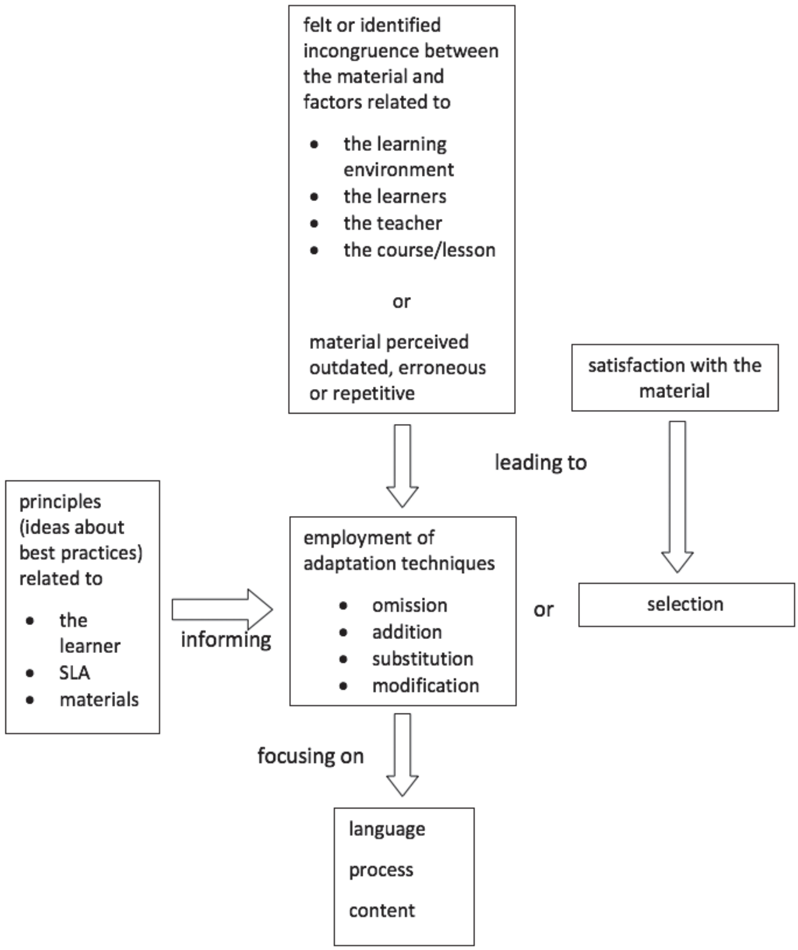
To summarise, adaptation can be a proactive decision as a part of lesson planning as a response to an anticipated challenge, a reactive, ad hoc intervention based on a perceived difficulty while teaching (Islam & Mares, 2003; Li & Harfitt, 2017; McGrath, 2013) or an outcome of experiences of using the material (Amrani, 2011). Even if an unprincipled attempt to adapt material may – by chance – generate learning, the need to tailor adaptation to the learner context in order to generate learning opportunities appears to be crucial irrespective of the procedure followed.

Conclusion

To research textbook adaptation and to inform practitioners (teachers, teacher educators and teacher trainers) about what it means to adapt textbook materials, it is important to possess an array of terminology to describe techniques, underlying principles as well as specific reasons and purposes that contribute to the act of adapting textbook materials. Figure 2 displays a visual summary of the steps and factors involved in textbook adaptation discussed in this paper.

Figure 2

Textbook adaptation: factors and steps in textbook adaptation



Appreciating the need to allow teachers to shape instructional practice creatively, we have focused on principles in textbook adaptation that are calibrated to generate learning opportunities. These principles do not exclusively aim at a textbook utilisation conducive to achieving expected learning outcomes measured in exam scores or desired learner behaviour (Ariyan & Pavlova, 2019; cf. Taggart & Wilson, 2005). They aim to raise the teachers' awareness of the need to critically assess what the textbook offers as 'the engine that drives much current practice' (Thornbury, 2013, p. 217) and to evaluate alternative practices. From this it follows that teachers need support in teacher education and professional development to become critical and informed textbook users with the perspective to develop their own materials (Agba, 2018; Bouckaert, 2019; Matić, 2019; Walterman & Forel, 2015). We hope that our paper contributes to this issue by providing a theoretical underpinning of textbook adaptation in English language teaching.

References

Abdel Latif, M. M. (2017). Teaching grammar using inductive and communicative materials: Exploring Egyptian EFL teachers' practices and beliefs. In H. Masuhara, F. Mishan, & B. Tomlinson (Eds.), *Practice and theory for materials development in L2 learning* (pp. 275–289). Cambridge Scholars Publishing.

Agba, Y. K. M. (2018). EFL materials – from adoption to adaptation. Definitional, practical and operational aspects on textbook development by teachers. *International Journal of English and Education, 7*(4), 260–274. http://ijee.org/yahoo_site_admin/assets/docs/21YOBOUE.28065606.pdf

Akbari, R. (2008). Postmethod discourse and practice. *TESOL Quarterly, 42*(4), 641–652. https://www.jstor.org/stable/40264492

Amrani, F. (2011). The process of evaluation: a publisher's view. In B. Tomlinson (Ed.), *Materials development in language teaching* (2nd ed.; pp. 267–295). Cambridge University Press.

Ariyan M. A., Pavlova L. V. (2019). An English language textbook: A friend or a foe. In Z. Anikina (Ed.), *Going global through social sciences and humanities: A systems and ICT perspective*. Advances in Intelligent Systems and Computing, vol 907 (pp. 83–90) Springer. https://doi.org/10.1007/978-3-030-11473-2_10

Bouckaert, M. (2019). Current perspectives on teachers as materials developers: Why, what, and how? *RELC Journal, 50*(3), 439–456. https://doi.org/10.1177/0033688218810549

Cunningsworth, A. (1995). *Choosing your coursebook.* Macmillan Heinemann.

Freeman, D. (2014). Reading comprehension questions: The distribution of different types in global EFL textbooks. In N. Harwood (Ed.), *English language teaching textbooks: Content, consumption, production* (pp. 72–110). Palgrave Macmillan.

Gray, J. (2016). ELT materials. Claims, critiques and controversies. In G. Hall (Ed.), *The Routledge handbook of English language teaching* (pp. 95–108). Routledge.

Guerrettaz, A. M., Engman, M. M., & Matsumoto, A. Y. (2021). Empirically defining language learning and teaching materials in use through sociomaterial perspectives. *Modern Language Journal, 105*(S1), 3–20. https://doi.org/10.1111/modl.12691.

Harmer, J. (2007). *The practice of English language teaching* (4th ed.). Pearson.

Harwood, N. (2021). Coda: An expanding research agenda for the use of instructional materials. *Modern Language Journal, 105*(1), 175–184. https://doi.org/10.1111/modl.12683

Islam, C., & Mares, C. (2003). Adapting classroom materials. In B. Tomlinson (Ed.), *Developing materials for language teachers* (pp. 86–99). Continuum.

Kovač, M., & Šebart, M. (2019). Educational publishing: How it works: Primary and secondary education publishing. In A. Phillips & M. Bhaskar (Eds.), *The Oxford handbook of publishing* (pp. 1–18). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198794202.013.14

Li, Z., & Harfitt, G. J. (2107). An examination of language teachers' enactment of curriculum materials in the context of a centralised curriculum. *Pedagogy, Culture & Society, 25*(3), 403–416. https://doi.org/10.1080/14681366.2016.1270987

Littlejohn, A. (2012). Language teaching materials and the (very) big picture. *Electronic Journal of Foreign Language Teaching, 9*(1), 283–297. http://www.andrewlittlejohn.net/website/docs/Littlejohn%20Language%20Teaching%20Materials%20and%20the%20(Very)%20Big%20Picture.pdf

Loh, J., & Renandya, W. A. (2016). Exploring adaptations of materials and methods. *The European Journal of Applied Linguistics and TEFL, 4*(2), 93–11.

López-Barrios, M., & Villanueva de Debat, E. (2014). Global vs. local: Does it matter? In S. Garton & K. Graves (Eds.), *International perspectives on materials in ELT* (pp. 37–52). Palgrave Macmillan.

Macalister, J. (2016a). Adapting and adopting materials. In M. Azarnoosh, M. Zeraatpishe, A. Faravani, & H. R. Kargozari (Eds.), *Issues in material development* (pp. 57–64). Sense Publishers.

Macalister, J. (2016b). Applying language learning principles to coursebooks. In W. A. Renandya & H. P. Widodo (Eds.), *English language teaching today* (pp. 41–51). Springer.

Madsen, H., & Bowen, J. (1978). *Adaptation in language teaching*. Newbury House.

Maley, A. (2011). Squaring the circle – reconciling materials as constraint with materials as empowerment. In B. Tomlinson (Ed.), *Materials development in language teaching* (2nd ed.; pp. 357–378). Cambridge University Press.

Maley, A. (2018). Creative materials: An oxymoron? In D. Bao (Ed.), *Creativity and innovations in ELT material development* (pp. 35–50). Multilingual Matters.

Matić, L. J. (2019). The teacher as a lesson designer. *CEPS Journal, 9*(2), 139–160. https://doi.org/10.26529/cepsj.722

McDonough, J., Shaw, C., & Masuhara, H. (2013). *Materials and methods in ELT* (3rd ed.). Wiley-Blackwell.

McGrath, I. (2013). *Teaching materials and the role of EFL/ESL teachers*. Bloomsbury.

McGrath, I. (2016). *Materials evaluation and design for language teaching* (2nd ed.). Edinburgh University Press.

Menkabu, A., & N. Harwood. (2014). Teachers' conceptualization and use of the textbook on a medical English course. In N. Harwood (Ed.), *English language teaching textbooks: Content, consumption, production* (pp. 145–177). Palgrave Macmillan.

Mishan, F. (2021). The global ELT coursebook: A case of Cinderella's slipper? *Language Teaching,* 1–16. https://doi.org/10.1017/S0261444820000646

Mishan, F., & Timmis, I. (2015). *Materials Development in TESOL*. Edinburgh University Press.

Pouromid, S., & Amerian, M. (2018). Iranian and Japanese language teachers' operationalization of materials mode and their pursuit of immunity. *Electronic Journal of Foreign Language Teaching, 15*(1), 5–22. https://e-flt.nus.edu.sg/wp-content/uploads/2020/09/pouromid.pdf

Rathert, S., & Cabaroğlu, N. (2021). Teachers as slaves or masters to their coursebooks: An in-depth study on two English language teachers' coursebook utilization. *Language Teaching Research.* https://doi.org/10.1177/13621688211036239

Richards, J. C. (2001). *The role of textbooks in a language program.* http://www.professorjackrichards.com/wp-content/uploads/role-of-textbooks.pdf

Richards, J. C. (2017) *Curriculum development in language teaching* (2nd ed.). Cambridge University Press.

Seferaj, K. (2014). Coping with new teaching approaches and materials: An East-European EFL teacher's interpretation of communicative teaching activities. In S. Garton & K. Graves (Eds.), *International perspectives on materials in ELT* (pp. 89–103). Palgrave Macmillan.

Smart, A., Sinclair, M., Benavot, A., Bernard, J., Chabbott, C., Russell, S. G., & Williams, J. et al. (2020). *Learning for uncertain futures: the role of textbooks, curriculum, and pedagogy. Background paper for the futures of education initiative*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000374078

Taggart, G.L., & Wilson, A.P. (2005). *Promoting reflective thinking in teachers. 50 action strategies.* Corwin Press.

Tasseron, M. (2017). How teachers use the global ELT coursebook. In H. Masuhara, F. Mishan, & B. Tomlinson (Eds.), *Practice and theory for materials development in L2 learning* (pp. 290–311). Cambridge Scholars Publishing.

Thornbury, S. (2013). Resisting coursebooks. In J. Gray (Ed.), *Critical perspectives on language teaching materials* (pp. 204–223). Palgrave Macmillan.

Tomlinson, B. (2015). Challenging teachers to use their coursebook creatively. In A. Maley & N. Peachey (Eds.), *Creativity in the English language classroom* (pp. 24–28). British Council.

Tomlinson, B. (2018). Making typical coursebook activities more beneficial for the learner. In D. Bao (Ed.), *Creativity and innovations in ELT material development* (pp. 21–34). Multilingual Matters.

Tomlinson, B., & Masuhara, H. (2004). *Developing language course materials.* SAEMO Regional Language Centre.

Tomlinson, B., & Masuhara, H. (2018). *The complete guide to the theory and practice of materials development for language learning.* Wiley Blackwell.

Ur, P. (2015). Using the coursebook: A teacher's perspective. *The European Journal of Applied Linguistics and TEFL, 4*(2), 1–17.

Vitta, J. P. (2021). The functions and features of ELT textbooks and textbook analysis: A concise review. *RELC Journal.* https://doi.org/10.1177%2F00336882211035826

Walterman, E., & Forel, C. (2015). Why and how to include textbook analysis in language teacher education programs. *English Language Education and Development Journal, 18*, 43–48. http://www.elted.net/uploads/7/3/1/6/7316005/elted_vol_18_paper_6_waltermann___forel.pdf

Zheng, X.-M., & Davison, C. (2008). *Changing pedagogy: Analysing ELT teachers in China*. Continuum.

Biographical note

Stefan Rathert, PhD, is an assistant professor in the field of English language teaching at the Faculty of Education, Kahramanmaraş Sütçü İmam University, Turkey. His research interests include language teaching materials development, teaching methods, translanguaging in language teaching and professional development for teachers.

Neşe Cabaroğlu, PhD, is a full professor in the field of English language teaching at the Faculty of Education, Çukurova University, Adana, Turkey. She received her Master of Arts Degree in 1993 (Çukurova University) and Doctor of Philosophy at the University of Reading, UK in 1999. Her research interests and expertise are in the areas of foreign language teaching and learning, language teaching methodology, pre-service teacher education and teachers' professional development.

doi: 10.26529/cepsj.1238

Interview with Richard E. Mayer about Multimedia Materials and Textbooks

Gregor Torkar[1]

This interview is part of the focus issue entitled "The Role of Textbooks in Teaching and Learning Processes", which aims to investigate the advantages and disadvantages of using a textbook, as well as ways of adapting textbooks that provide teachers with an opportunity to personalise teaching material and enable students to be more actively involved in the learning process. The interview was conducted online (exchanging emails back and forth between Ljubljana, Slovenia and Santa Barbara, California, USA) in February 2021.

GREGOR TORKAR: I am aware that it is very difficult and responsible task to summarise all of the important achievements of the distinguished professor Richard Mayer, who has contributed so significantly to the field of education. Working at the University of California, Santa Barbara, Richard Mayer has devoted his entire professional life to educational psychology, making significant contributions to theories of cognition and learning, particularly problem solving and multimedia learning. His best known contribution to the field of educational psychology is multimedia learning theory, which postulates that optimal learning occurs when visual and verbal learning materials are presented simultaneously. Professor Mayer has received numerous prestigious awards for his outstanding scholarly contribution. Most notably, he was awarded the E. L. Thorndike Award for professional achievement in educational psychology and he is the winner of the 2008 Distinguished Contribution of Applications of Psychology to Education and Training Award from the American Psychological Association. He was ranked as the most prolific educational psychologist in the world for the period 1997–2001, and he is the author of hundreds of publications, including more than twenty books on education and multimedia.

Multimedia learning theory is a fundamental theory for the design and use of textbooks and other types of multimedia. Despite its modest length, this interview aims to point out some important directions related to ensuring the present and future quality of textbooks and other educational materials.

1 University of Ljubljana, Faculty of Education, Slovenia; gregor.torkar@pef.uni-lj.si.

GREGOR TORKAR: Distinguished Professor Richard Mayer, I am very privileged to have the opportunity to talk to you about your multimedia learning theory, which is a world-renowned and, above all, very effective educational theory. However, I would like to ask you first, if you don't mind, to begin with a few words about yourself and the professional path that led you to educational psychology.

RICHARD E. MAYER: Thank you, Professor Torkar, for your kind words. My professional path involves receiving a BA degree in psychology from Miami University (in Oxford, Ohio) and a PhD degree in psychology from the University of Michigan (in Ann Arbor, Michigan). After graduating, my path took me to a teaching position in the Psychology Department at Indiana University (in Bloomington, Indiana) for two years, followed by my move to the University of California, Santa Barbara, where I have served on the faculty for more than 40 years. My research focus has always been on the issue of how to help people learn in ways so that they can take what they have learned and apply it to new situations. My curiosity about this issue of how to promote transfer is what led me to the field of educational psychology.

GREGOR TORKAR: Let me start now with some basic questions about the multimedia learning theory that you break down in such detail in your books, especially in Multimedia Learning (2020) and The Cambridge Handbook of Multimedia Learning (2014). In the introduction to your book Multimedia Learning, you wrote that you explore ways that go beyond purely verbal learning. What exactly is multimedia learning and what do all the principles of instructional design (for example, coherence, signalling, redundancy, pre-training, segmenting, modality, personalisation, etc.) actually contribute to learning and teaching?

RICHARD E. MAYER: Multimedia learning is learning from words and graphics. The words can be spoken or printed; the graphics can be static (e.g., photos, drawings, charts, etc.) or dynamic (e.g., video or narration). I became interested in multimedia learning when my lab repeatedly found that people performed better on a transfer posttest when they had studied a lesson that included words and graphics rather than words alone. For example, we found better transfer test performance for students who saw an animation while they heard a narration describe how a bicycle tire pump works, rather than for students who only listened to the narration. In 13 experiments, we consistently found this pattern, with a median effect size greater than 1. We call this the multimedia principle: people learn better from words and graphics than from words alone. In trying to better understand how to optimise multimedia learning, we found that not all multimedia instructional messages are equally effective. For example, in series of experiments, we found that people learn better from multimedia lessons when extraneous words and graphical elements are eliminated (i.e., coherence principle), when key words or aspects of graphics are highlighted during instruction (i.e., signalling principle), and when printed words are placed next to the part of the graphic they refer to (i.e., spatial contiguity principle). These techniques seek to reduce extraneous processing, which is cognitive processing that does not support the instructional objective and wastes precious processing capacity that could have been used for deeper learning. In another series of experiments, we found that people learn better from multimedia lessons when the lesson is presented in bite-size segments paced by the learner (i.e., segmenting principle), when students receive training in the names and characteristics of the key concepts before the lesson (i.e., pre-training principle), and when the words are spoken rather printed on the screen (i.e., modality principle). These techniques seek to manage essential processing, which is cognitive processing aimed at representing the core material in working memory. Finally, another series of experiments showed that people learned better when the words in a multimedia lesson are presented in conversational style rather than formal style (i.e., personalisation principle) and when the instructor engages in appropriate gesture, facial expression, body stance and eye gaze during instruction (i.e., embodiment principle). These techniques seek to foster generative processing, which is cognitive processing aimed at making sense of the material.

GREGOR TORKAR: Which principles of instruction design do you think should be consistently considered when presenting educational material in textbooks? Could you somehow rank them? Is it even possible to rank the importance of such principles?

RICHARD E. MAYER: In our analyses of textbooks used in California schools, the most grossly violated principle was the coherence principle. Books contained beautiful colour graphics that were not related to the essential lesson, and included interesting but irrelevant stories that can be called seductive details. If I had to choose one principle for revising textbooks, I would start by choosing the coherence principle and seek to remove irrelevant and distracting elements so students can focus on learning the essential material in the lesson. Next, I would add the spatial contiguity principle, which calls for removing the captions on figures and moving the essential text (in segments) next to the corresponding part of the graphic. When a textbook has graphics with long captions or legends, that is an indication of poor design.

GREGOR TORKAR: If I may ask one more sub-question about the presentation of multimedia material. I am aware that you also work extensively on e-learning, presented in your books E-Learning and the Science of Instruction: Proven Guidelines for Consumers and Designers of Multimedia Learning (2016) and Computer Games for Learning: An Evidence-Based Approach (2014). In your opinion, what are the most important advantages and perhaps also disadvantages of digital learning materials?

RICHARD E. MAYER: A consistent theme in research on technology-supported instruction is that instructional media do not cause learning, but instead, instructional methods cause learning. It is not computers per se that cause learning, but rather how we use computers to guide instruction in line with theories of how people learn. Thus, digital learning materials can be successful when they employ effective instructional methods and can be unsuccessful when they employ ineffective instructional methods. There may be some instructional methods that are better afforded by computer-based platforms than by textbooks, such as using well-designed interactive simulations, videos and animations. However, learning is caused by using appropriate instructional design with media such as simulations or videos or animations.

GREGOR TORKAR: The digital environment now allows much greater flexibility and personalisation of multimedia materials compared to traditional printed textbooks and other multimedia materials. We can incorporate animations, simulations, digital games, augmented and virtual reality, etc. into the materials, just to name a few. The question arises as to when more and more is too much (i.e., the redundancy principle) in the learning process? How do you see this development in education?

RICHARD E. MAYER: You raise an excellent point. We have found that learning in virtual reality is emotionally arousing (as measured by heartrate and skin conductance), which can lead to distraction. A solution can be to stop the VR lesson at various points and ask the student to summarise what has been learned so far. We have also found that games have many attention-grabbing elements that can be distracting and result in poorer learning than with conventional media. A solution to this problem is to add instructional elements, such as a worksheet that the player uses throughout a game. In short, new media such as games, simulations and immersive virtual reality – although motivating – can create extraneous processing in learners, so I recommend incorporating instructional features (such as self-explaining or worksheets) that prompt learners to reflect on what they are learning.

GREGOR TORKAR: Your multimedia learning theory is a "living" theory, if I may say, and subject to the process of new developments and (re) interpretations. In scholarly works based on your theory, I perceive that some authors draw a line between multimedia materials consisting of pictures and written words, and multimedia materials consisting of pictures and spoken words in defining multimedia learning. I point out this distinction being aware of the visual/pictorial and auditory/verbal communication channels (Paivio's dual channels) and the limited capacity of information processing. You deal with this issue, for example, when explaining narrated animations and of course in your famous figure of the cognitive theory of multimedia learning. Do you think that the notion of multimedia material should only apply to multimedia material using separate communication channels (visual and auditory), as some scholars try to interpret the theory, or not?

RICHARD E. MAYER: I agree that we can make a distinction between book-based multimedia involving printed text and illustrations versus computer-based multimedia involving spoken text and video or animation. However, in general, our research shows that the same basic principles – such as multimedia, coherence, signalling, segmenting, pre-training and personalisation – apply to both venues.

GREGOR TORKAR: As one of the many researchers referencing your work, I am very curious to know how you actually developed multimedia learning theory? What were the initial pieces of the puzzle that later led you to multimedia learning theory? Did you start from existing theories and assumptions (e.g., Paivio, 1986; Baddeley, 1992) in formulating the empirical studies and theory, or were you initially guided by the possibilities that arose from advances in educational technology, which gradually led you to formulate the theory?

RICHARD E. MAYER: You are right in suggesting that the cognitive theory of multimedia learning was influenced by pre-existing theoretical ideas, including Paivio's dual code theory, Baddeley's working memory theory, Wittrock's generative learning theory, and, of course, Sweller's cognitive load theory.

GREGOR TORKAR: Reading your recent scientific articles, I came across a study (Lawson et al., 2021) on the emotional role of animated pedagogical agents in educational material. I would like to use this as a starting point for forming a question about the role of teachers in the process of designing and using multimedia materials. What is the primary role of teachers in the process of multimedia learning and how is their role changing with the introduction of

new technologies, such as the aforementioned animated pedagogical agents or virtual reality, which is also the subject of your research interests?

RICHARD E. MAYER: Teachers are central to the effective use of instructional materials, mainly in their roles in selecting, implementing and, in some cases, creating multimedia instructional materials. An understudied aspect of multimedia learning concerns contextual studies that examine how teachers effectively use multimedia materials in their classrooms.

GREGOR TORKAR: I would like to move away from the main topic slightly and try to include the book A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives (2001) in the textbook discussion. You are a co-author of this book, as well. What is your view of the relevance of a revised Bloom's taxonomy to education today? How is it relevant to designing or using multimedia materials?

RICHARD E. MAYER: I was honoured to be asked by David Krathwohl – the last surviving member of the team that wrote the original document in 1956 – to be part of a team to write a revision of Bloom's taxonomy that was published in 2001. I learned a lot about the thinking that went into Bloom's taxonomy and I came away with a greater respect for the science of assessment – the scientific study of determining what students have learned. In my opinion, the science of assessment is an essential component in educational psychology, along with the science of learning and the science of instruction. What I appreciate most about Bloom's taxonomy is the focus on measuring transfer (i.e., the ability to use the learned material) in addition to retention (i.e., the ability to remember the learned material).

GREGOR TORKAR: In the last three or four decades, the way we access information and the way we learn has changed a great deal. Information and communication technologies have changed the learning environment and will continue to do so. What are the most important changes over this period that you think have benefited education the most? And secondly, which current technological innovations hold the most promise for education and will continue to do so in the future?

RICHARD E. MAYER: We now have easy access to a vast amount of information as well as exciting new information formats (such as interactive simulations and virtual reality), which has important implications for education. Students need to learn an expanded form of literacy that includes what I call multimedia literacy – the ability to understand multimedia materials and to create multimedia materials that others can understand. Another aspect of multimedia literacy is that students need to be able to work with multiple sources of information, make judgements of credibility and relevance, and integrate the information. In short, we need to equip students with the skills they need for the world of multimedia information.

In terms of the promise of technology for education, I am interested in how we can design effective instruction with interactive simulations, games, animated pedagogical agents and immersive virtual reality. Although these media have promise, research is needed to determine how best to take advantage of that promise.

GREGOR TORKAR: The last question is also directed to the future. I try to follow your work and that of your younger colleagues, and am impressed by your drive and fresh ideas. What are your current and future research goals and what are the educational challenges that we should all address in the near future?

RICHARD E. MAYER: We are currently examining how the emotional stance of instructors (both animated and human) affects learning, how games can be used to train cognitive skills, how to design effective academic learning in immersive virtual reality, and how to incorporate prompts for generative learning activities in online multimedia lessons. Much of our work involves international collaborators, as the search for multimedia design principles is clearly a global effort. This global effort is indicated, for example, by the author list of the forthcoming third edition of The Cambridge Handbook of Multimedia Learning. Thanks again for your thought-provoking questions.

GREGOR TORKAR: Thank you professor for your valuable contribution to the focus issue of the CEPS Journal.

doi: 10.26529/cepsj.976

Interdisciplinary Interaction between Music Education and History: Shaping the Musical Preferences in Classical Music of the 20th Century

Jerneja Žnidaršič[1]

The purpose of the current study was to investigate whether an experimental programme, based on interdisciplinary interactions between music education and history and the implementation of arts and cultural education objectives, could influence pupils' interest in Western classical music of the 20th century. The programme was designed on the basis of collaborating with music education and history teachers at two Slovenian primary schools and a Slovenian composer. Classes of pupils, aged fourteen and fifteen, were divided into an experimental and a control group. According to the outcome, the pupils in the experimental group showed a higher level of interest in contemporary classical music after the experiment than their peers in the control group. Furthermore, the pupils in the experimental group reported having listened on their initiative, to more classical compositions after the experiment than the pupils in the control group had.

Keywords: arts and cultural education, history, interdisciplinary interactions, music education, musical preferences

1 Faculty of Education, University of Maribor, Slovenia; jerneja.znidarsic@um.si.

Medpredmetno povezovanje glasbene umetnosti in zgodovine: oblikovanje glasbenih preferenc klasične glasbe 20. stoletja

Jerneja Žnidaršič

Namen raziskave je bil preučiti, ali eksperimentalni program, ki temelji na medpredmetnem povezovanju med glasbeno umetnostjo in zgodovino ter na implementaciji ciljev kulturno-umetnostne vzgoje, vpliva na zanimanje učencev za zahodno klasično glasbo 20. stoletja. Program je bil zasnovan na osnovi sodelovanja z učitelji glasbene umetnosti in zgodovine iz dveh slovenskih osnovnih šol ter s slovensko skladateljico. Razredi z učenci, starimi od štirinajst do petnajst let, so bili razdeljeni v eksperimentalno in kontrolno skupino. Rezultati raziskave so pokazali, da so učenci eksperimentalne skupine po izvedenem eksperimentalnem programu pokazali višjo stopnjo zanimanja za sodobno klasično glasbo kot njihovi vrstniki v kontrolni skupini. Poleg tega so učenci iz eksperimentalne skupine po izvedenem eksperimentalnem programu poročali, da so na lastno pobudo poslušali več klasičnih skladb kot učenci iz kontrolne skupine.

Ključne besede: glasbena umetnost, glasbene preference, kulturno-umetnostna vzgoja, medpredmetno povezovanje, zgodovina

Introduction

Contemporary society confronts us with the challenges of finding the best and most effective and sustainable teaching and learning methods used in the education process. Modern approaches in education emphasise the need to move from a model of transmission to a model of transformation. Such approaches are characterised by a growing role of the pupils within the learning process. With the help of the teacher, pupils acquire knowledge through their activity, discovery, and experience. Within this framework, an interdisciplinary interaction[2] in the learning environment enables pupils to recognise meaningful connections across different curricular disciplines and discuss various topics from different angles.

Several research studies discuss positive correlations of interdisciplinary interactions, such as encouragement of pupils to discover the connections between seemingly unrelated domains, facilitation of a personalised process of organising knowledge, development of generic skills, which are independent of the content and can be transferred and used in different situations (Hodnik Čadež, 2013; Rich, 2009; Sicherl-Kafol, 2013); improvement of motivation and interest for the learning process (Cheung, 2008; Michelsen & Sriraman, 2009; Serrano Pastor, 2013), openness to different points of view (LTTA, 2010) and improvement of pupil-teacher relationships (Drake & Burns, 2004). Furthermore, interdisciplinary collaborations are recognised as a concept for sustainable learning (learning that is continuous, enduring, proactive) and education (current and viable theories, practices, systems that influence positive change) (Hays & Reinders, 2020).

Given its positive effects in recent years, interdisciplinarity is also often a part of the learning process in Slovenian schools (Sardoč, 2004; Sicherl-Kafol, 2013). However, the studies show that among the teachers, the understanding of interdisciplinarity is questionable (Hodnik Čadež, 2013), most studies are carried out only at the level of the first and the second triennium of primary school[3]

2 The terms 'interdisciplinary interactions/ interdiciplinary(ity)' is used generally, to describe various interactions between two or more disciplines.

3 The primary school applies to primary and lower secondary education, which is organized in a single-structure nine-year basic school, attended by students aged 6 to 15 years. The basic school programme is divided into three educational cycles, each covers three years. In the first educational cycle, pupils are taught by the general (class) teacher. In the second educational cycle there is primarily a general teacher and at the end of the second and in the third educational cycle the lessons are delivered by specialist teachers (The education system in the Republic of Slovenia 2018/2019). Music education is a part of the curriculum for nine years of primary school, in which pupils develop their musical abilities, skills, and informative knowledge through various musical activities (listening, creating, playing on instruments). Pupils can also attend a non-compulsory basic music education, outside mainstream formal education, through which they can upgrade the knowledge with music-theoretical subjects and individually learn to play their chosen instruments.

(Devjak et al., 2013; Marjanovič Umek et al., 2011) and that most teachers implement interdisciplinary interactions spontaneously when the right occasion arises (Štemberger, 2013); this means that interdisciplinarity is seldom planned in Slovenian schools, which casts doubt on the quality of its implementation. Furthermore, the teachers most frequently actualise the connections between the subjects based on subject-matter and objectives, and not on the process-developmental model that involves the transfer of certain skills, experience and enables the use of teaching methods, which best contribute to the development of pupils' abilities, such as the ability of critical thinking (Pevec Semec, 2007).

Regarding music education in interdisciplinary settings, we can find various studies (Bresler, 2010; Serrano Pastor, 2013) dealing with the positive effects of interdisciplinary interactions between artistic and non-artistic subjects. Nevertheless, music is too often used only to reinforce the content in subjects of a non-artistic area (Barry, 2008; Suraco, 2006), failing to explore the possible implications that an interdisciplinary interaction could have on different aspects of music education.

Arts and cultural education

The level of quality in education can be raised through elevated awareness and inclusion of arts and cultural education into the teaching process. The latter can be introduced as the integration of different art disciplines within school subjects or other curricular activities. Moreover, arts and cultural education enables the possibility for goal-oriented and process-developmental planning. The main objectives are to develop creativity, individual capabilities, aesthetic sensitivity, and critical reflection of culture and art, to promote the expression of cultural diversity, to raise cultural awareness, and to promote national cultural heritage (National Guidelines for Arts and Cultural Education, 2009). When dealing with the practical implications, one of the key principles of arts and cultural education is the implementation of cooperation among schools and artists (Bamford, 2006; De Backer et al., 2012; Hout et al., 2017; National programme for culture, 2018-2025; Road Map for Arts Education, 2006; Sicherl-Kafol & Denac 2011), where pupils' have, for example, the opportunity to engage in the creative process while being mentored by the composer. According to Ivon and Kuščević (2013), perspectives of authentic, interactive, and integrative learning place pupils not only as the consumers of cultural values but also as the creators of culture and its future values.

Listening to Music and Musical Preferences

Listening to, exploring, choosing, and discussing favourite music represent an essential part of a young adult's life. Background and intentional music listening are the most frequent adolescent leisure activities (Lavrič et al., 2011; Marketing Charts, 2015). They affect the formation of their values and relations to others and society in general. What kind of music do young adults choose and, more importantly, why? What are the determinants that influence people's musical preferences? The results of various studies suggest that, in a wider context, musical preferences are influenced by individual, social and situational factors or context (Bonneville-Roussy et al., 2017).

Individual factors are reflected in personal characteristics (Delsing et al., 2008), individual values (Tekman et al., 2012), arousal potential (North & Hargreaves, 2008), cognitive abilities (Getz et al., 2014), social identity (Mans, 2009), age (Bonneville-Roussy & Eerola, 2017), and gender (Brittin, 2014; North & Hargreaves, 2008).

Regarding social factors, musical preferences are conditioned by social class (North & Hargreaves, 2007), peers (Selfhout et al., 2009), parents (Schäfer & Sedlmeier, 2009), mass media (Wingstedt et al., 2008) and the music industry (Evans, 2010; Negus, 2006).

In addition to the above-mentioned factors, researchers have identified some other influences on musical preferences, such as music functions (Miranda & Claes, 2009; Schäfer & Sedlmeier, 2009), musical knowledge (Getz et al., 2014), general knowledge of context and facts, as well as sociological and analytical comments while becoming familiar with a specific composition (Johnson, 2009) and features of music itself (Imbir & Gołąb, 2017; North & Hargreaves, 2008).

The influence of musical features on music preferences is a field of research within experimental aesthetics. One of the leading researchers in that field was Berlyne (1971), who suggested that the main factor influencing musical preferences in terms of artistic stimuli is arousal, which is largely conditioned by 'collative' stimulus properties, specifically, familiarity, complexity, and the degree of surprise while listening to a specific composition. A high degree of novelty and complexity of a musical piece evokes a high degree of arousal potential and a lower degree of preference, which escalates, as the musical piece becomes familiar, to the point of overexposure and by declining attractiveness. Many researchers have considered the U-shaped relationship between arousal and intrinsic attractiveness (Imbir, 2015; Moors et al., 2013; North & Hargreaves, 2008). Following the series of criticisms of Berlyne's theory and experimental aesthetics, some authors pointed out the problem of dealing with the concept of 'complexity of artistic stimuli' (North & Hargreaves, 2008). North and Hargreaves (2008) suggest replacing the term *objective complexity,* which can be defined as the statistical probability of one note following another (suggesting (non) sequence by means of tonality and key functions, particularly in Western music) with the term *subjective complexity* notwithstanding the less accurate methodological verifiability. The latter is not associated with music itself as much as it focuses on the listener and one's perception of complexity. Various researchers (Krumhansl et al., 2000; Orr & Ohlsson, 2001) supported that theory as they have concluded that, through repetition of melody, the listener perceives it as less complex, thus raising the preference level. Repeated listening reduces unpredictability in music – the objective complexity is still the same, but the subjective complexity decreases. Furthermore, repeated listening and music experiences (Oxenham et al., 2003) help listeners develop listening skills. Individuals with more musical training and musical experiences (objectively) prefer more complex music than people without musical background (Getz et al., 2014).

However, mere exposure does not directly lead from previously unknown music to a higher interest in a specific composition or song. The listener has to increase the effort to become familiar with the music (Lamont & Webb, 2009), and experience is less than an emotional response. The majority of popular music is accompanied by lyrics – a tool for regulating the emotional response (Fiveash & Luck, 2016) that provides additional musical information that influences the interpretation of music (Thompson et al., 2008). Regarding instrumental music, Tan and Kelly (2004) argue that pupils (with no prior musical training) focus on the arousal of emotions or sensations and create stories to accompany music.

In adolescence, pupils most frequently choose popular music (Dobrota & Ercegovac, 2017; Vries, 2010), whereas classical music is most likely reserved for formal music education (Vitale, 2011). One of the reasons can be found in the fact that musical genres are connected with social stereotypes. Studies have shown that young adults hold consistent beliefs about others claiming to be fans of a certain genre (Rentfrow & Gosling, 2007). Furthermore, the preferences for classical music are complex because atonal music is less favoured compared to tonal music (Ball, 2011; Meyers, 2012). The music teacher is thus confronted with a challenge: how to raise pupils' level of interest in classical music above school settings and to develop the abilities of pupils' critical attitudes towards influences of different factors such as the music industry and mass media.

The present research aims to establish possible effects of interdisciplinary interactions between music education and history on shaping the musical preferences in classical music of the 20th century while implementing the objectives and principles of arts and cultural education.

Method

Research objectives

The main research objectives are:

- To design an experimental programme for the interdisciplinary teaching of music and history for the ninth grade of primary school based on teaching processes, objectives, methods, activities and subject-matter while implementing the learning objectives of arts and cultural education,
- To evaluate the effectiveness of existing and experimental teaching programmes with regard to pupils' interest in classical music of the 20^{th} century.

Experimental Programme

The experimental programme is based on interdisciplinary interactions, namely intertwined multidisciplinary and 'interdisciplinary' perspectives.[4] Music and history teachers had collaboratively planned didactic units, deriving from developmental and goal-oriented models, while considering general and operative objectives and the subject-matter covered by the curriculums. A team of teachers identified the cohesive elements at the level of learning processes, objectives, teaching methods, subject-matter, and activities. An important cohesive element was the concept of arts and cultural education. Based on the theoretical background of *National Guidelines for Arts and Cultural Education* 2009, we have implemented the objectives of arts and cultural education, such as developing pupils' critical thinking, critical attitudes towards culture and arts, aesthetic sensitivity, creativity, tolerance to other cultures, cultural awareness, attitude towards the preservation of art and cultural heritage, enabling the ability to experience and re-experience cultural creations, understanding the importance of intercultural dialogue, building cultural identity, and learning about one's own culture and the cultures of other nations.

In the context of culture and arts education, the experimental programme included an interdisciplinary project in collaboration with a famous Slovenian composer. The project aimed to enable the pupils to engage in the musical experience of the creative process with the help of an artist. Inspiration for the creative process were historical events of the 20^{th} century.

4 The term 'interdisciplinary' perspective describes a specific form of interdisciplinary interactions. The typology by Klein (2010) is being used: multidisciplinarity (juxtaposing, sequencing and coordinating; absence of or partial integration); 'interdisciplinarity' (integrating, interacting, linking, focusing, blending; integration); transdisciplinarity (transending, transgressing, transforming; full integration).

The experiment comprised:

- 12 music lessons (37.5% of all music lessons in the ninth grade) covering the following topics: Impressionism, Science and art in the 20th century, Music in the 1st half of the 20th century, Music in the 2nd half of the 20th century, Slovenian music in the 1st half of the 20th century, Slovenian music in the 2nd half of the 20th century, jazz;
- 25 history lessons (39.0% of all history lessons in the ninth grade) covering the following topics: World War I, the world between the World Wars, Slovenians between the World Wars, World War II, Slovenians during World War II, the world after World War II;
- 14 lessons (6 music lessons, 2 history lessons, 6 lessons before and after class) – implementation of the interdisciplinary project *Music through History* in collaboration with a composer.

To study the effectiveness of interdisciplinary teaching in music education and history while implementing objectives of arts and cultural education, we used an experimental method of empirical/analytical pedagogic research paradigm.

The survey was designed as a single factor experiment with classes serving as comparison units. The experimental factor had two modalities: (1) teaching music education and history according to standard curriculum and traditional teaching methods; (2) teaching music education and history according to the standard curriculum with multidisciplinary and interdisciplinary perspectives, implementing the objectives of arts and cultural education in the intended and delivered school curriculum and including an interdisciplinary project in collaboration with a composer.

Research Hypotheses

H 1.1: We presume that after the experiment, the pupils from the experimental group will have a higher degree of interest in classical music of the 20th century than the pupils from the control group.

H 1.2: We presume that after the experiment, the pupils from the experimental group will have a higher degree of interest in a musical composition, classified as a classical composition of the 20th century, than the pupils from the control group.

H 1.3: We presume that during the experiment, the pupils from the experimental group will have listened, at their initiative, to more classical compositions from the 20th century than the pupils from the control group.

Study Population

We surveyed four classes of ninth graders from two Slovenian primary schools (aged fourteen to fifteen).

As the participating schools were randomly chosen, there were initial problems with the inability of the schools and teachers to cooperate. Once the schools were acquired, we divided pupils into two groups: the control group (n = 33) and the experimental group (n = 43). Due to the extended work and preparations needed in the experimental group, we have, in assigning the roles of research groups, considered the willingness of music and history teachers at the specific school to implement the experimental programme.

The initial state has indicated that there were no statistically significant differences between the pupils of the experimental and the control group, regarding the gender ($\chi^2 = 0.244$; g = 1; P = 0.622); the attendance of music school ($\chi^2 = 0.045$; g = 1; P = 0.832); the level of musical education of pupils parents (mother: $\chi^2 = 3.661$; g = 3; P = 0.160; father: $\chi^2 = 1.562$; g = 3; P = 0.668); the frequency of listening to music with parents in childhood ($\chi^2 = 1.275$; g = 2; P = 0.529) and the musical genres pupils were most frequently listening to with their parents in childhood ($\chi^2 = 12.846$; g = 11; P = 0.303).

In terms of statistical hypothesis testing, the two selected groups of pupils represent a simple random sample from a hypothetical population.

Data Collection Procedures and Instrument

Prior to conducting the experiment, we selected the participating classes from the chosen primary schools, which provided us with some general information. Using a preliminary questionnaire and music examples, we have defined the music genres (rock, turbo-folk, metal, R'n'B, reggae, rap, house/techno, jazz, punk, classical music, popular folk music, pop, traditional folk music) recognised by pupils. Then we gathered information on the initial and final states of the control group and the experimental group by means of a questionnaire and musical interest test (for the final state). Research instruments had been designed specifically for the survey.

The initial questionnaire comprises five closed types (multiple answers) questions referring to the study population, which are not repeated in the final questionnaire and question regarding the ranking of the level of interest in the specific musical genre. *The final questionnaire* comprises two questions (one closed type: ranking of musical genres; one semi-open question: self-motivated listening).

The validity was ensured with the pilot study on a sample of pupils (n = 30) from the 9th grade of a primary school in Ljubljana. The final list of musical genres was determined based on a pilot study, which included 9th-grade primary school pupils (n = 20) who identified the most familiar musical genres.

The reliability of the questionnaire was achieved with the same unambiguous questions for all respondents. The objectivity of the questionnaire was ensured by closed-ended questions, whereby the answers were not subjected to subjective assessment. Clear and unambiguous instructions were given in solving the questionnaire.

The *musical interest test* comprises 14 consecutive numbers, each number representing the particular piece of music: (1) Queen: The Show Must Go On; 2) System of a Down: Toxicity; 3) Pankrti: Bandiera Rosa; 4) Atomic Harmonic: Turbo Polka; 5) Destiny's Child: Lose my Breath; 6) Richie Spice: Youth them cold; 7) Cypress Hill: How I could kill a man; 8) DJ Umek: Posing As Me; 9) Maceo Parker: Pass the Peas; 10) Johann Sebastian Bach: Bandiniere; 11) Avseniki: Na Roblek; 12) Katy Perry: California Gurls; 13) Vlado Kreslin: Vsi so venci vejli; 14) Krzysztof Penderecki: Threnody to the Victims of Hiroshima). Prior to and after the experiment, pieces of music were evaluated (level of interest: 1 - the lowest; 10 - the highest; familiarity with pieces of music: yes/no) by pupils.

The validity of the test of musical interests was ensured by pieces of music representing the considered musical genres. Cronbach's alpha coefficient (α) was used to determine the reliability of the musical interest test. The coefficient is $\alpha = 0.823$ for the initial test and $\alpha = 0.813$ for the final, which indicates that the test of musical interests is reliable. The objectivity of the musical interest test was ensured by a closed type of question and a given scale for measuring the level of interest. The evaluation criteria for the responses were uniform. Pupils of the control and experimental group did not listen to any of the included pieces of music before the experiment during the school process of the third educational cycle. During the experiment, students in both comparison groups listened to a composition under number fourteen as part of the expressionism theme.

Data processing

The data were processed at descriptive and inferential levels, using the following statistical methods: Frequency distribution of variables (f, f %), Arrays of variables by average ranks (R), χ^2 test (Pearson's χ^2 test) of the hypothesis of independence, Mann-Whitney U-test, Cronbach's alpha (reliability of musical preferences test).

Results

Interests in Classical Music

Pupils in the two comparison groups had to indicate their musical interest towards thirteen music genres. Before engaging in the study, we carried out a survey, in which pupils were asked to write down the genres they know best, together with examples of musical pieces. The analysis showed that pupils recognised the following musical genres: rock, turbo-folk, metal, R'n'B, reggae, rap, house/techno, jazz, punk, classical music, popular folk music, pop, traditional folk music. These were included in the questionnaire used in the study, where pupils ranked the musical genres accordingly. They used number one to indicate their favourite genre and number thirteen for the least favourite one. We shall only present the results regarding classical music of the 20th century, which is the subject of our hypotheses.

Table 1

Differences between the experimental (EG) and the control group (CG), regarding their level of interest in classical music

Music genre	Group	Rank average R	Mann-Whitney test	
			U	P
Classical music	EG	33.90	495.5	.035
	CG	44.50		

Prior to the experiment, the Mann-Whitney test results showed no statistically significant differences between the pupils in the experimental group and those in the control group regarding their interest in classical music (U = 607.0; P = .277). Table 1 indicates a statistically significant post-experiment difference between the two comparison groups regarding their interest in classical music (P = .035). The pupils demonstrated a higher level of interest in the experimental group. The results thus confirm our hypothesis H 1.1.

Interest in a Particular Classical Composition

During the test of musical interest, pupils listened to fourteen musical compositions from various music genres. The fourteenth composition was a classical piece from the 20th century, which is included in the musical education curriculum of the ninth grade of primary school and was a part of the learning process in both comparison groups. Pupils graded the compositions on a scale from 1 to 10, number 1 representing the lowest and number 10 the highest level of interest. In addition, they had to indicate whether they had been familiar with each composition prior to the test. The following results apply to the musical composition that is directly connected with our hypothesis.

Table 2

Differences in the range of musical interest in musical compositions between the experimental (EG) and control (CG) groups

Musical composition	Group	Rank average	Mann-Whitney test	
		R	U	P
Krzysztof Penderecki: 'Threnody to the Victims of Hiroshima'	EG	53.99	43.500	.000
	CG	18.32		

Prior to the experiment, there were no statistically significant differences between the two comparison groups regarding their interest in the composition (U = 656.6, P = .296). However, a statistically significant difference between the experimental and the control group was noted after the experiment (P = 43.500, P = .000), with a higher level of interest in the composition *Threnody to the Victims of Hiroshima* identified by the pupils in the experimental group. The results thus confirmed our hypothesis H 1.2.

Because the composition was a part of the learning process in both groups, we also tested the evaluation of the familiarity with the composition. Before the experiment, there were no statistically significant differences between the comparison groups regarding their familiarity with the respective composition. None of the pupils from the experimental or control groups had known the *Threnody to the Victims of Hiroshima*. After the experiment, a statistically significant difference exists in familiarity with the composition between the two comparison groups (χ^2 = 19.101, g = 1, P = .000), the composition was largely recognised by the pupils of the experimental group (f%= 93.0%) and in a lesser extent by pupils of the control group (f% = 48.5%).

Self-motivated Listening to the Classical Music of the 20^th^ Century

During the experiment, neither the pupils in the experimental group nor those in the control group were given any instructions regarding intentional listening to any classical music. Also, their school tests in music education and history did not include any listening tasks. The results of the χ^2 test of the hypothesis of independence between pupils groups (EG: experimental group, CG: control group) after the experiment have shown that there is a statistically significant difference in the pupils' self-motivation for listening to classical music of the 20^{th} century (χ^2 = 20.535; g = 1; P = .000); 55.8% of pupils of the experimental group reported listening to the classical composition at their initiative, and 6.1% of pupils in the control group did.

The most listened to composition at pupils' (EG) initiative was *Einstein on Beach* (Glass), *Arabesque* (Debussy) and *The Moonfleck* (Schönberg). Other answers of pupils encompassed: *Lan* – (Choir song from the 20th century); *Opera – Wozzeck*; *Threnody to the Victims of Hiroshima*; *Classical Symphony*; *October Cantata; Prelude to the Afternoon of a Faun; I do not recall the title of the composition.* In the control group, two pupils listened to a composition from the 20^{th} century: *Ljubljana Postcards* (Dekleva – Works for a piano from the 20th century), and one of the compositions was not named. Based on the presented results, we can confirm hypothesis H 1.3.

Discussion and conclusions

In the present study, we examined the efficiency of the experimental programme of interdisciplinary interaction, connecting music and history, while implementing the arts and cultural education objectives. In that process, we were focusing on increasing the musical preferences of pupils, especially the preference for music genre and specific composition of classical music of the 20^{th} century, and creating an environment with different musical experiences that would encourage pupils to include listening to classical music in their leisure activities. The study results have confirmed the positive influence of the designed experimental programme of teaching, as opposed to the traditional one, regarding the pupils' musical preferences. The pupils of the experimental group have shown higher levels of interest in classical music of the 20^{th} century as a genre and specific compositions and have listened to more classical composition at their initiative in comparison to pupils from the control group. The interdisciplinary aspect of the experimental programme with the integration of music education and history offered pupils an in-depth perception and an analytical approach to various music-related subject matters, with classical music playing a particularly important role. The pupils, through classical music, explored the events and lives of people living in the 20^{th} century. Unveiling the historical background of the musical compositions contributed to experiencing and understanding music at a more profound level and provided a possibility of an emotional response.

In contrast, musical compositions helped pupils (through their musical expression) to understand the nature of historic events better. A higher degree of interest in *Threnody to the Victims of Hiroshima*, expressed by the pupils of the experimental group, was related to the pupils' associations of musical elements with their perception of the horrors of war and atomic bombings, which is why they experienced the composition more intensely and emotionally. As Kemperl

(2013) stated, when pupils discuss the context of origin, the idea behind it, and the reasons for the creation of the work of art, they discuss the world, broadening their horizons and recognising the problems of modern society.

From the neurological aspect, due to the development of the brain (Galvan et al., 2006; Sturman & Moghaddam, 2011), adolescents are prone to emotional behaviours; therefore, the musical preferences of adolescents are, among other factors, influenced by a strong connection between music and emotional response (Miranda & Gaudreau, 2011). Various researchers have proven that empathic behaviour is connected with the relationship between an emotional expression of music itself and the emotions felt by the listener (Evans & Schubert, 2008; Schubert, 2013). Furthermore, some researchers argue that familiarity with non-musical elements of musical compositions results in a higher level of musical interest (Johnson, 2009; Thompson et al., 2008).

Another factor of musical preference that has been considered when designing the experimental programme was the manipulation with the familiarity of specific compositions. That has been developed on the basis of the repetitive and analytical listening of composition in music and history lessons by pupils of the experimental group. At the end of the experiment, they showed higher levels of interest in the particular piece than pupils from the control group. Pupils of the experimental group perceived classical compositions of the 20th century as less complex and difficult, which positively impacted their musical interest. These results are confirmed by other studies (North & Hargreaves, 2008; Orr & Ohlsson, 2001), as a higher frequency of listening to a composition reduces its complexity as perceived by the listener, which in turn elevates the level of preference and helps the listener to develop listening skills (Oxenham et al., 2003).

An important aspect of the experimental programme was also the implementation of the arts and cultural education. Through developmental and objective-related models of teaching, pupils of the experimental group were able to experience and re-experience cultural creations and develop critical thinking, creativity, and aesthetic sensitivity. The concept of arts and cultural education was also comprised of an interdisciplinary project, which allowed pupils to identify and explore their creative potential, based on the features of classical music of the 20th century, while being inspired by historical events. Pupils' own musical experience through the creative process resulted in a better understanding of the genre and motivated them to listen to classical compositions at their initiative, outside the school environment. Vuk et al. (2015) argue that when the creative process is aimed at creative problem solving, creating original ideas, expressing thoughts and feelings, communicating with the environment, being sensitive to moral and social issues, it encourages student's motivation and identification with student's work.

Furthermore, according to Green (2006), negative experience with certain types of music might arise due to being unfamiliar with a particular musical style, to the point that we find the musical syntax boring or feel that the music is not ours. Making pupils feel closer to a musical style makes it more approachable, understandable, and more their own, which in turn affects their interest in it. Furthermore, some authors discovered that pupils with an opportunity to positively experience music (in our case, collaboration with an artist) are encouraged to extend their musical lives beyond the classroom (Cabedo-Mas & Diaz-Gómez, 2013).

The limitation of the study can be recognised in the size of the research sample and thus in the availability for generalisation. Furthermore, based on the results, it is not possible to conclude that the pupils' musical preferences for classical music had been entrenched, as there is no evidence that they still listen to classical music at their initiative. Further research that would monitor the preferences over a longer period would be welcomed. However, we find the results encouraging, as they offer an example of possible paths towards the positive development of musical preferences. The experimental model offers guidelines for effective planning of interdisciplinary interaction and implementation of arts and cultural education; furthermore, it contributes to greater awareness of the importance of partnerships between schools and artists.

References

Ball, P. (2011). Schoenberg, serialism and cognition: Whose fault if no one listens? *Interdisciplinary Science Reviews*, *36*(1), 24–41.

Bamford, A. (2006). *The wow factor: Global research compendium on the impact of the arts in education*. Waxmann.

Barry, N. H. (2008). The role of integrated curriculum in music teacher education. *Journal of Music Teacher Education*, *18*(19), 28–38.

Berlyne, D. E. (1971). *Aesthetics and psychobiology*. Appleton-Century-Crofts.

Bonneville-Roussy, A., & Eerola, T. (2017). Age trends in musical preferences in adulthood: 3. Perceived musical attributes as intrinsic determinants of preferences. *Musicae Scientae*, *22*(3), 394–414. https://doi.org/10.1177/1029864917718606

Bonneville-Roussy, A., Still Well, D., Kosinski, M., & Rust, J. (2017). Age trends in musical preferences in adulthood: Conceptualization and empirical investigation. *Musicae Scientae*, *21*(4), 369–389. https://doi.org/10.1177/1029864917691571

Bresler, L. (2010). Integracija glasbe v splošnoizobraževalne predmete: prikaz izobraževalnega podjetništva [Integration of music into general academic subjects: Introducing educational entrepreneurship]. *Glasba v šoli in vrtcu, 15*(1), 3–17.

Brittin, R. V. (2014). Young listeners' music style preferences: Patterns related to cultural identification and language use. *Journal of Research in Music Education*, *61*(4), 415–430. https://doi.org/10.1177/0022429413509108

Cabedo-Mas, A., & Diaz-Gómez, M. (2013). Positive musical experiences in education: music as a social praxis. *Music Education Research, 15*(4), 455–470.

Cheung, J. (2008). The arts-in-programme. Towards effective learning through partnership. In L. Joubert (Ed.), *Educating in the arts. Education in the Asia-pacific region: Issues, concerns and prospects* (pp. 65–81). Springer Link.

De backer, F., Lombaerts, K., De Mette, D., Buffel, T., & Elias, W. (2012). Creativity in artistic education: Introducing artists into primary schools. *International Journal of Art & Design Education, 31*(1), 53–66.

Delsing, M. J. M. H., Ter Bogt, T. F. M., Engels, R. C. M. E., & Meeus, W. H. J. (2008). Adolescents' music preferences and personality characteristics. *European Journal of Personality, 22*(2), 109–130.

Devjak, T., Hodnik Čadež, T., Krek, J., Sicherl Kafol, B., Štemberger, V., & Vogrinc, J. (2013). *Učitelj v vlogi raziskovalca* [Teacher as a researcher]. Pedagoška fakulteta Univerze v Ljubljani.

Dobrota, S., & Eregovac, R. (2017). Music preferences with regard to music education, informal influences and familiarity of music amongst young people in Croatia. *British Journal of Music Education*, *34*(1), 41–55.

Drake, S., & Burns, R. (2004). *Meeting standards through integrated curriculum*. ASCD.

Eijnden, J., Hout, F., Kox, R., Poll, J., Tal, M., & Wervers, E. (2017). *Basis for cultural education. guide for the future of school-based and extracurricular cultural education.* LKCA. https://www.lkca.nl/~/media/downloads/lkca/english/basis_for_cultural_education.pdf

Evans, P. & Schubert, E. (2008). Relationships between expressed and felt emotions in music. *Musicae Scientae*, *12*(1), 75–99. https://doi.org/10.1177/102986490801200105

Evans, D. N. (2010). Ryan Moore: Sells like teen spirit: Music youth culture and social crisis. *Youth Adolescence, 39*(4), 434–435.

Fiveash, A., & Luck, G. (2016). Effects of musical valence on the cognitive processing of lyrics. *Psychology of Music, 44*(6). https://doi.org/10.1177/0305735615628057

Galvan, A., Hare, T. A., Parra, C. E., Penn, J., Voss, H., Glover, G., & Casey, B. J. (2006). Earlier development of the accumbens relative to orbitofrontal cortex might underlie risk-taking behaviour in adolescents. *Journal of Neuroscience*, *26*(25), 6885–6892. https://doi.org/10.1523/JNEUROSCI.1062-06.2006

Getz, L. M., Marks, S. & Roy, M. (2014). The influence of stress, optimism, and music raining on music uses and preferences. *Psychology of Music*, *42*(1), 71–85.

Greene, L. (2006). Popular music education in and for itself, and for 'other' music: Current research in the classroom. *International Journal of Music Education, 24*(2), 101–118.

Hays, J., & Reinders, H. (2020). Sustainable learning and education: A curriculum for the future. *International Review of Education*, *66*(1), 29–52. https://doi.org/10.1007/s11159-020-09820-7

Hodnik Čadež, T. (2013). Učitelj kot raziskovalec medpredmetnega povezovanja [Teacher as a researcher of interdisciplinary connections]. In T. Devjak, T. Hodnik Čadež, J. Krek, B. Sicherl Kafol, V. Štemberger, & J. Vogrinc (Eds.), *Učitelj v vlogi raziskovalca* (pp. 131–149). Pedagoška fakulteta Univerze v Ljubljani.

Hui, W. V. (2009). Music listening preferences of Macau pupils. *Music Education Research, 11*(4), 485–500.

Imbir, K. (2015). Affective Norms for 1,586 Polish Words (ANPW): Duality of mind approach. *Behaviour Research Methods*, *47*(3), 860–870. https://doi.org/10.3758/s13428–014–0509–4

Imbir, K., & Golab, M. (2017). Affective reactions to music: Norms for 120 excerpts of modern and classical music. *Psychology of Music*, *45*(3), 432–449. https://doi.org/10.1177/0305735616671587

Ivon, H., & Kuščević, D. (2013). School and the cultural-heritage environment: Pedagogical, creative and artistic aspects. *CEPS Journal*, *3*(2), 29–50.

Johnson, D. A. (2009). *The effect of prior information on musical preference rating* (Master's thesis). University of Jyväskylä.

Kemperl, M. (2013). Contemporary art and citizenship education: The possibilities of cross-curricular links on the level of content. *CEPS Journal, 3*(2), 97–117.

Klein, J. M. (2010). A taxonomy of interdisciplinarity. In R. Frodeman, J. T. Klein, & C. Mitsham (Eds.), The Oxord handbook of interdisciplinarity (pp. 15-30). Oxford: Oxford Universitiy Press.

Krumhansl, C. (2000). Cross-cultural music cognition: cognitive methodology applied to North Sami yoiks. *Cognition*, *76*(1), 13–58. https://doi.org/10.1016/s0010-0277(00)00068-8

Lamont, A., & Webb, R. (2009). Short- and long-term musical preferences: what makes a favourite piece of music? *Psychology of Music*, *38*(2), 222–241. https://doi.org/10.1177/0305735609339471

Lavrič, M., Flere, S., Tavčar Krajnc, M., Klajnšek, R., Musil, B., Natarer, A. … Lešek, P. (2011). *Mladina 2010: Družbeni profil mladih v Sloveniji* [Youth 2010: The Profiles of Youth in Slovenia]. Aristej.

LTTA (2010). Selected research findings for LTTA students. The Royal Conservatory. https://www.rcmusic.com

Mans, M. (2009). *Living in worlds of music. A view of education and values.* Springer Science + Business Media. https://doi.org/ 10.1007/978-90-481-2706-1

Marjanovič Umek, L., Pečjak, S., Krajnc, S., Kordigel Aberšek, M., Saksida, I., Žakelj, A., Kern, M. (2011). *Lili in Bine. Priročnik za Poučevanje in medpredmetno povezovanje v prvem triletju* [Lili and Bine. Manual for interdisciplinary teaching in first triennium]. Rokus.

Marketing Charts. (2015). Teens' most common-and preferred-media activities. https://www.marketingcharts.com

Meyers, C. (2012). Influences on music preference formation. *PURE Insights*, *1*(7). https://digitalcommons.wou.edu/pure/vol1/iss1/7

Moors, A., De Houwer, J., Hermans, D., Wanmaker, S., Van Schie, K., Van Harmelen Brysbaert, M. (2013). Norms of valence, arousal, dominance and age of acquisition for 4,300 Dutch words. *Behavior*

Research Methods, 45(1), 169–77. https://doi.org/10.3758/s13428-012-0243-8

Michelsen, C., & Sriraman, B. (2008). Does interdisciplinary instruction raise students' interest in mathematics and the subjects of the natural sciences? *ZDM*, *41*(1-2), 231–244. https://doi.org/10.1007/s11858-008-0161-5

Miranda, D., & Claes, M. (2009). Music listening, coping, peer affiliation and depression in adolescence. *Psychology of Music, 37*(2), 215–233.

Miranda, D., & Gaudreau, P. (2011). Music listening and emotional well-being in adolescence: A person and variable-oriented study. *European Review of Applied Psychology, 61,* 1–11. https://doi.org/ 10.1016/j.erap.2010.10.002

National guidelines for arts and cultural education. (2009). Ministrstvo Republike Slovenije za šolstvo in šport. http://www.mss.gov.si

National programme for culture (2018-2025). Ministrstvo Republike Slovenije za kulturo. http://www.mk.gov.si/fileadmin/mk.gov.si

Negus, K. (2006). *Music genres and corporate cultures.* Routledge.

North, A. C., & Hargreaves, D. J. (2007). Lifestyle correlates of musical preference: 3. Travel, money, education, employment and health. *Psychology of Music, 35*(3), 473–497.

North, A. C., & Hargreaves, D. J. (2008). *The social and applied psychology of music*. Oxford University Press.

Orr, M. G., & Ohlsson, S. (2001). The relationship between complexity and liking across musical styles. *Psychology of Music*, *29*(2), 108–127.

Oxenham, A. J., Fligor, B. J., Mason, C. R., & Kidd, G. (2003). Informational masking and musical training. *The Journal of the Acoustical Society of America*, *114*(3), 1543–1549. https://doi.org/10.1121/1.1598197.

Pevec Semec, K. (2007). Kurikularna teorija in posodabljanje učnih načrtov [Curricular theory and curriculum updating]. In A. Žakelj, M. Pušnik, M. Turk Škraba, & M. Lesjak Reichenberg (Eds.), *Zbornik prispevkov posveta Postojna. Kurikul kot proces in razvoj* (pp. 132–138). Zavod RS za šolstvo.

Rentfrow, P. J., & Gosling, S. D. (2007). The content and validity of music-genre stereotypes among college students. *Psychology of Music*, *35*(2), 306–326. https://doi.org/10.1177/0305735607070382

Rich, S. J. (2009). Interdisciplinary learning: A stimulant for reflective practice. In R. K. Chem, K. M. Hibbert, & T. V. Deven (Eds.), *Radiology Education* (pp. 51–61). Springer. https://doi.org/10.1007/978-3-540-68989-8_5

Road map for arts education. (2006). UNESCO. *Document from The world conference on arts education: Building creative capacities for the 21st Century*. Lisbon, 6–9 March, 2006. http://www.unesco.org

Sardoč, M. (2004). *Medpredmetno povezovanje vzgojno-izobraževalnega procesa v 9-letni osnovni šoli: evalvacijska študija: zaključno poročilo* [Interdisciplinary connections in educational process in 9-year primary school: Evaluation report]. Pedagoški inštitut.

Schäfer, T., & Sedlmeier, P. (2009). From the functions of music to music preference. *Psychology of Music*, *37*(3), 297–300.

Schubert, E. (2013). Emotion felt by the listener and expressed by the music: Literature review and theoretical perspectives. *Front. Psychology*, *4*, 837. https://doi.org/ 10.3389/fpsyg.2013.00837

Selfhout, M. H. W., Branje, S. J. T., ter Bogt, T. F. M., & Meeus, W. H. J. (2009). The role of music preferences in early adolescents' friendship formation and stability. *Journal of Adolescence*, *32*(1), 95–107. https://doi.org/10.1016/j.adolescence.2007.11.004.

Serrano Pastor, R. M. (2013). Interdisciplinary education through music. *Human and Social Sciences at the Common Conference,* 237–241.

Sicherl-Kafol, B. (2013). Procesni in vsebinski vidiki medpredmetnega povezovanja [Process and content related aspects of interdisciplinary integration]. In T. Devjak, T. Hodnik Čadež, J. Krek, B. Sicherl Kafol, V. Štemberger, & J. Vogrinc (Eds.), *Učitelj v vlogi raziskovalca* (pp. 112–130). Pedagoška fakulteta Univerze v Ljubljana.

Sicherl - Kafol, B., & Denac, O. (2011). Through musical communication to development of competence in culture awareness and expression. *US-China Education Review, 8*(2), 129–136.

Sturman, D. A., & Moghaddam, B. (2011). The neurobiology of adolescence: Changes in brain architecture, functional dynamics, and behavioral tendencies. *Neuroscience & Biobehavioral Reviews*, *35*(8), 1704–1712. https://doi.org/10.1016/j.neubiorev.2011.04.003

Suraco, T. L. (2006). An interdisciplinary approach in the art education curriculum. *Art and Design Thesis*. Georgia State University.

Štemberger, V. (2013). Načrtovanje in izvajanje medpredmetnih povezav [The planning and the execution of interdisciplinary connections]. In T. Devjak, T. Hodnik Čadež, J. Krek, B. Sicherl Kafol, V. Štemberger, & J. Vogrinc (Eds.), *Učitelj v vlogi raziskovalca* (pp. 93–111). Pedagoška fakulteta Univerze v Ljubljani.

Tan, S., & Kelly, M. E. (2004). Graphic representations of short musical compositions. *Psychology of Music,* 32(2), 191–212.

Tekman, H. G., Boer, D., & Fischer, R. (2012). *Values, functions of music, and musical preferences.* Proceedings of the 12th International conference on music perception and the 8th triennial conference of the European Society for the cognitive sciences of music, July 23–28., 2012, Thessanoliki, Greece.

Thompson, W. F., Russo, F. A., & Quinto, L. (2008). Audio-visual integration of emotional cues in song. *Cognition & Emotion*, *22*(8), 1457–1470. https://doi.org/10.1080/02699930701813974

Vitale, J. L. (2011). Formal and informal music learning: Attitudes and perspectives of secondary school non-music teachers. *International Journal of Humanities and Social Sciences*, *1*(5), 1–14.

De Vries, P. (2010). What we want: The music preferences of upper primary school students and the ways they engage with music. *Australian Journal of Music Education,* 1, 3–16.

Vuk, S., Tacol, T., & Vogrinc, J. (2015). Adoption of the creative process according to the immersive method. *CEPS Journal, 5*(3), 51–71. https://cepsj.si/index.php/cepsj/article/view/127

Wingstedt, J., Brändström, S., & Berg, J. (2008). Young adolescents' usage of narrative functions of media music by manipulation of music expression. *Psychology of Music, 36* (2), 193–214.

Biographical note

Jerneja Žnidaršič, PhD, is an assistant professor in the field of didactics of music education at the Faculty of Education, University of Maribor, Slovenia. Her research interest include interdisciplinary connections, arts and cultural education as well as listening to music in educational settings. In the field of her professional work she strives for closer connections between research work and practice.

doi: 10.26529/cepsj.1029

Creating and Developing a Collaborative and Learning-Centred School Culture: Views of Estonian School Leaders

Katrin Poom-Valickis*[1], Eisenschmidt Eve[2] and Ann Leppiman[2]

The present study aims to analyse how school leaders perceive their activities in creating and developing a collaborative school culture that promotes the school learning process. The data were collected in semi-structured interviews with nine school leaders and analysed using thematic content analysis. The results revealed that only three of the school leaders focused on the shared values and shared leadership necessary for creating a systematic and analytic approach to organisational and teacher development. The school leaders understood the importance of leading the development of the learning process, but this did not take place as expected in practice. Organisational and teacher development seemed to be unsystematic or not based on the continuous monitoring of processes. The findings of our study indicate that development programmes for school leaders should concentrate more on shaping the views, knowledge and skills needed to develop a collaborative and learning-centred school culture.

Keywords: school leaders, collaborative school culture, learning-centred leadership

1 *Corresponding Author. Tallinn University, Estonia, katrinpv@tlu.ee.
2 Tallinn University, Estonia.

Vzpostavljanje in razvijanje sodelovalne in na učenje usmerjene šolske kulture: stališča estonskih ravnateljev šol

Katrin Poom-Valickis, Eisenschmidt Eve in Ann Leppiman

V raziskavi skušamo analizirati, kako ravnatelji šol dojemajo aktivnosti vzpostavitve in razvoja sodelovalne šolske kulture, ki spodbuja šolski učni proces. Podatki so bili zbrani s polstrukturiranimi intervjuji devetih ravnateljev šol in analizirani z uporabo tematske vsebinske analize. Izsledki so pokazali, da so se le trije ravnatelji osredinjali na skupne vrednote in deljeno vodenje za ustvarjanje sistematičnega in analitičnega pristopa pri organizacijskem razvoju in razvoju učiteljev. Ravnatelji so razumeli pomen vodenja razvoja učnega procesa, a se to ni odražalo v praksi. Organizacijski razvoj in razvoj učiteljev sta bila nesistematična ali pa nista bila osnovana na stalnem spremljanju procesa. Ugotovitve raziskave kažejo, da se morajo razvojni programi za ravnatelje šol bolj osredinjati na stališča, znanje in na veščine, ki so potrebne za razvoj sodelovalne in na učenje usmerjene šolske kulture.

Ključne besede: ravnatelji šol, sodelovalna šolska kultura, na znanje usmerjeno vodenje

Introduction

Previous research has demonstrated that changes in teaching do not occur if the organisation is not oriented to the learning and development of its members (Opfer et al., 2011). The collaboration of teachers influences the quality of teaching and consequently the learning outcomes of students. However, such collaboration depends on whether the school management has created favourable conditions for it to thrive (Kruse & Louis, 2009; Woodland et al., 2013). Estonian schools provide interesting contexts to study leadership, as they show high academic performance and achievement levels in international assessments, such as PISA (OECD, 2019). Interestingly, however, the Teaching and Learning International Survey (TALIS) (OECD, 2014) points to a poor culture of collaboration among teachers and a lack of substantive feedback on their teaching activities. Although not designated as head teachers, Estonian school leaders have a high level of autonomy, with authority to appoint and dismiss staff, negotiate working conditions, and make decisions about school finances, educational priorities and development plans for the school (Estonian Parliament, 2010). They are the main actors in designing the school curriculum, which is based on the national framework, and in supporting teachers' professional development in order to implement changes. At the same time, based on the TALIS 2013 survey, only 7% of Estonian school leaders regularly visit lessons to observe and monitor learning processes (the TALIS average is 49%), while 41% of school leaders develop measures to support teacher collaboration on new teaching methods (the TALIS average is 64%). In order to gain a clearer understanding of the contradictory results referred to above, we decided to explore school leaders' views on their activities in creating and developing a collaborative school culture within their everyday leadership practice. This is particularly pertinent considering that the Estonian Lifelong Learning Strategy (Estonian Ministry of Education and Research, 2014) and the Estonian school leader's competence model (Innove, 2016) emphasise the school leader's essential role in leading changes and developing a leadership style that focuses on the development of learners, teachers and all school personnel.

The school leader's role in creating a learning-centred collaborative school culture

A collaborative school culture that focuses on the improvement of the learning process has been addressed by researchers for a long time. Such a culture makes an important contribution to both the success of the school improvement process and its effectiveness (Campo, 1993; Fullan & Hargreaves, 1991), creating an environment in which changes are faster, problems are solved together, and teachers support each other (Eisenschmidt et al., 2015). It has also been realised that a collaborative school culture helps teachers to increase their self-esteem and self-confidence, take responsibility for managing various situations (Angelides, 2010; Kohm & Nance, 2013), and find additional meaning in their work (Kruse & Louis, 2009). School leaders play an important role in creating a school culture that values collaborative learning by providing time for teachers to do research, plan and design together. Researchers have identified that the school leader's support for and participation in the professional learning of teachers is the most significant means by which school leadership impacts student learning (Robinson et al., 2008). A collaborative learning-centred school culture provides a climate and structure that encourages teachers to work with each other, fosters staff learning and professional growth, and benefits all members of the school community.

Research into school leaders has highlighted various strategies and opportunities for making a school culture more collaborative (Day & Sammons, 2006; Hallinger, 2011; Leithwood et al., 2008; Liu et al., 2016; Sales et al., 2017). Based on earlier research, we can point out four key activities in developing learning-focused leadership and collaborative school cultures: (a) building a shared vision and setting goals; (b) sharing responsibility and creating an environment for collaborative learning; (c) improving the quality of teaching and learning; and (d) modelling and building trusting relationships within the organisation (see Table 1).

Table 1

School leaders' key activities in creating a learning-focused collaborative school culture.

Leader's key activities	Description of activities
1. Building a shared vision and setting goals	- Formulating a broad, long-term agenda (vision) and explaining it to the entire staff. - Planning actions based on the vision. - Constantly communicating the vision. - Fostering the ownership and acceptance of shared goals. - Clarifying roles and objectives.
2. Sharing responsibility and creating an environment for collaborative learning	- Joint decision-making and shared responsibility. - Sharing expertise. - Team building. - Creating opportunities for working together (working groups, formal and informal meetings, professional networks, etc.). - Creating conditions for learning and sharing experiences (time and space).

Leader's key activities	Description of activities
3. Improving the quality of teaching and learning	- Systematically supporting teachers' professional development. - Monitoring the learning process and collecting evidence (observing lessons, monitoring student achievement, research, measuring impact, etc.). - Providing feedback to teachers. - Mentoring.
4. Modelling and building trusting relationships	- Being a role model (motivating, inspiring). - Open and positive communication. - Recognising and acknowledging progress. - Creating a safe environment for learning and risk-taking.

Building a shared learning vision and goal setting reflects the extent to which school leaders articulate and communicate an inspiring vision that motivates learning in the school (Liu et al., 2016). Creating and communicating a shared vision in dialogue with the members of the organisation is a foundation for effecting changes in an organisation (Kruse & Louis, 2009; Leithwood et al., 2008). Efficient school leaders who share the school vision are always visible in the school building: they go around and talk with teachers, give feedback about the latest developments, set new development goals, and create a feeling of success by giving positive recognition (Barber et al., 2010; Engels et al., 2008), thus fostering ownership of the directions for development. Researchers have stressed the significance of clearly communicating values and directions of development (Day & Sammons, 2006). However, it is important that the vision is also operational, that is, openly formulated as explicit and clear activities in the development plan of the school. In addition, these goals and activities should constantly be kept in mind when making choices and decisions (Youngs & King, 2002).

Convincing research has shown that ***sharing responsibility*** is one of the essential steps for ***creating a collaborative school culture*** (Vangrieken et al., 2015). Shared responsibility and expertise encourage the sustainability of changes, enabling the changes to have a more solid footing (Harris, 2005; Heck & Hallinger, 2009). Implementing shared leadership presupposes shared goals and a favourable working climate, ensuring time for teachers to get together and benefit from the school leader's friendly and supportive attitude (Bush & Glover, 2014). Thus, the leader's role is to create a hospitable environment, devote time and space to collaboration and sharing expertise, provide resources, and support the implementation of teacher learning (Barber et al., 2010; Hallinger, 2011; Kruse & Louis, 2009). A school culture that is based on open communication and flexibility allows teachers to participate in decision making and express their opinions, thus increasing their feelings of control in work engagement (Zahed-Babelan et al., 2019). Empowering teachers' learning communities and working groups (although professional learning communities can include members from outside the school, e.g., parents, staff from other schools and external stakeholders) supports teacher leadership and plays a crucial role in improving the quality of teaching.

In order to ***improve the quality of teaching and learning*** it is essential that the main process of learning is constantly and consistently planned, monitored and analysed in light of the shared goals. Research has shown that student achievement is higher when the school leader focuses on developing and leading teaching and learning at school (Leithwood et al., 2008; Robinson et al., 2009).

Effective leaders pay attention to teaching with a particular focus on student learning. They devote as much time as possible to supporting teachers in their efforts to strengthen teaching and learning in the classroom. They monitor the process and collect evidence to provide feedback to teachers in order to ensure that high quality and alignment between learning goals and classroom instruction is maintained (Murphy et al., 2007). Instructional quality can be strengthened by systematically supporting teachers' collaborative planning, evidence-based practice development, reflection and mentoring (Robinson et al., 2008; Lai et al., 2016). Thus, we can say that school leaders who see themselves as pedagogical leaders focus first and foremost on developing and guiding the learning process (Bush & Glover, 2014; Hallinger, 2011; Hallinger et al., 2017). Researchers have observed that one important aspect through which school leaders influence teacher commitment to change and professional learning is gaining their trust in the school vision and leadership (Li et al., 2016; Tschannen-Moran, 2009).

Modelling and building trusting relationships highlights the role that school leaders play in supporting the values of openness, risk-taking and collaboration in their own behaviour (Hallinger, 2011; Leithwood et al., 2010). Research indicates that trust develops between school leaders and teachers when the school leader's beliefs and actions are consistent with school goals, when school leaders share responsibility and support teachers' work, and when they manage conflicts proactively and effectively (Youngs & King, 2002). In other words, trust favours effecting changes, but well-managed and successful changes also build up trust in the school leader. School leaders cannot build trust simply by talking; it must be seen in their actions (Kruse & Louis, 2009). By creating a climate of psychological safety, leaders can increase learning (by learning from mistakes and failures) and encourage teachers' creativity and their readiness for risk-taking in implementing novel ideas in their teaching process (Yukl, 2012). Trust and a supportive working atmosphere in which school leaders model learning leadership and help other leaders and teachers to grow are considered important traits of a learning organisation (Bruggencate et al., 2012).

In order to understand how Estonian school leaders perceive their role in fostering the learning process at school, we seek the answer to the following question: How do school leaders view their activities in creating and developing a learning-centred collaborative school culture?

Method

Sample

School leaders who had participated in a large-scale survey conducted in Estonian schools in 2017 were invited to participate in a qualitative study. Nine school leaders agreed to take part: five of them were men and four were women, and they were all aged between 30 and 60 years (see Table 2). Their work experience varied from one to thirteen years. The type of school varied from preschool/primary to upper secondary school (grades 10 to 12), and the number of students in the schools varied from 208 to 1520.

Table 2
Data on the school leaders and schools studied.

School	Gender	Age	Work experience as a school leader (years at the present school)	Type of school	Number of students as of 2017/2018
1	Male	50–59	1	Upper secondary school (with basic school)	1520
2	Female	60 +	Over 10	Upper secondary school (with basic school)	876
3	Male	50–59	13	Upper secondary school (with basic school)	900
4	Male	50–59	8	Basic school	438
5	Female	50–59	9	Upper secondary school (with basic school)	381
6	Male	40–49	2	Upper secondary school	479
7	Male	30–39	1	Upper secondary school	208
8	Female	50–59	8	Preschool/primary school	764
9	Female	60 +	5	Upper secondary school	323

Data collection and analysis

The data were collected using semi-structured interviews. The questions were compiled based on interview questions devised by Nevgi and Korhonen (2016), which focus on mapping and analysing the leadership styles of middle managers at universities, and adapted to the context of Estonian schools. In the present study, we focused only on the questions concerning the school leader's perceptions of their role and activities in developing a learning-focused collaborative school culture. The questions focused on: (a) organisational culture (e.g., How would you describe the work culture of your organisation?); (b) leading the learning process of the school (e.g., What is your role and what are your goals as a leader in pedagogical development? What do you view as the changes that need to be implemented in this regard?). A more detailed overview of the questions used in the study is provided in Appendix 1.

The duration of the interviews varied from 53 minutes to 1 hour and 33 minutes. The interviews were transcribed and analysed using thematic content analysis (Braun & Clarke, 2006). The unit of analysis was a complete answer or single sentences, or sets of sentences that expressed a conceptual whole. A deductive approach was used: a list of key activities for supporting learning and a collaborative school culture (see Table 1), compiled based on the texts, was used to structure the analysis. The interviewees' responses were analysed based on this list, and similarities and differences in the school leaders' patterns of behaviour were mapped. In the first phase, the first author undertook the initial data analysis. In the second phase, the themes and codes were reviewed by the second author and the reliability of the coding was checked. During mutual discussions, disagreements were discussed until the researchers reached a common interpretation of the coding.

Excerpts from the interviews with the school leaders are provided with the corresponding school number (see Table 2) in order to distinguish the views of different school leaders.

Results

Building a shared vision and setting goals

An analysis of the school leaders' key activities supporting a learning-centred and collaborative school culture reveals that although five school leaders clearly stated that their schools do have a vision, only in three schools does this vision actually serve as the foundation for daily decision making and activities: "*…what is important for us at the moment and what is not. What do we do, and what don't we do, what is relevant and what is not…*" (7). In this case, the role of the school leader is described by one of the school leaders as: "*Seeing the big picture. Whether something is not right, whether something needs to be changed, or maintained /.../ and for understanding, if it gets too tight somewhere you need to look at those main documents all the time /.../ the vision has already been created, and the main direction has been set, but now it is about how we get all those wobbly things better...*" (9). The need for constant communication is viewed as follows: "*…we know our vision and our direction in the sense that I see it like this /.../ I simply explain, or at least it seems to me that I have to work at it all the time*" (8).

In the case of four of the school leaders, the shared values and long-term goals that they aim to achieve remained unclear: "*...maybe this vision for the future and what happens, but as I said, it is kind of a vague situation*" (3). "*I don't, uhm, create this vision by myself, it has to come through someone else* (4)".

The school leaders claimed that they aimed to make and maintain changes, but what those changes were remained unclear. Two of the school leaders admitted that the school development plan had a life of its own, so to speak, and administrative duties did not leave enough time to take the initiative; consequently, the school leader's essential role in communicating the vision remained largely unfulfilled. Five of the school leaders mentioned that they do not manage to "*get around and about the school*" (2) as much as they would like and be visible to the teachers as much as needed. Nevertheless, when it comes to implementing the vision in practice, the shared understanding of the entire team plays a significant role. If the vision is not shared by everyone, then it may happen that: "*…head teachers move things in a different direction, the goals of the development plan are not actualised, or only some people contribute*" (1).

Thus, the interviews reflected the tendency for school leaders to become dissatisfied with the work of the management team in schools, which either lacked clear direction and strategic goals, or in which the aims were not sufficiently communicated and not used in everyday decision-making and operating processes.

Sharing responsibility and creating an environment for collaborative learning

The second key activity for a learning-centred collaborative school culture is related to joint decision making and shared responsibility. The interviews revealed that in most schools, the responsibility for pedagogical development was delegated to head teachers. School leaders saw their role rather as that of a generator of ideas and a discussion partner: "*...I think that my role is about letting the others know when I hear about innovations /.../ then I create* *the opportunity for the head teacher*" (5). One school leader considered their own role to be mainly about managing finances: "*...I tell them all the time in the teachers' council /…/ that you are top specialists, I am not managing your time, I am managing finances, those things, so that you do the things that we have agreed on together...*" (1). Only two leaders said that they had a close collaboration with the head teacher in substantively developing teaching and learning processes. However, these school leaders also mentioned that their long work experience as a head teacher made it easy for them to have this collaboration.

Based on the interviews, it seems that shared leadership works mostly from top to bottom, meaning that the management assign a task, propose an idea, and lead the formation of working groups. On the one hand, the school leaders expected the teachers themselves to show more initiative and take responsibility: "*...I wish there were fewer questions about how to do this or that and more about what needs to be done; at the moment, the way it works is that once you set up a task the next question is how are we going to do it /.../ Kind of taking responsibility...*" (2). On the other hand, not enough opportunities were given for such initiative, as the processes were lead from top to bottom.

Teachers' lack of initiative may also be due to vague development goals, which influence motivation as well as the feeling of ownership of innovations. Here, the school leaders with a clear vision of strategic goals clearly stood out: "*…things like 'I got this really great idea, let's do it like this now!' just don't work. You've got to talk about it first, in an information meeting, for example, that I've got this idea, and how I came to it and why it came or how we could get it, whether someone sees any obstacles, then it becomes like everyone's idea. So, then we can get on with it, or I get some feedback that: 'Come on, don't mess around, that's totally irrelevant' or 'That's not okay' or 'Let's think about it...' and so on...*" (7). In these schools, joint discussions were organised and decision making on new initiatives and activities was broad-based.

In the interviews, all of the school leaders mentioned regular meetings for sharing daily information, but in two schools there was also a longer meeting for joint discussions. In the words of one school leader: "*…last year we created a collaboration day /.../ when everyone said that we don't have time to meet /.../ and don't find the place, then we'll do it this way by making a collaboration day, and this collaboration day is Monday; on this day nobody leaves the school before half past three*" (2). It is important for the school leader to create opportunities for collaboration and provide teachers with the time and space for sharing experiences and planning their activities together. This not only creates a favourable environment for collaborative learning, but also supports joint decision making. When teachers are involved in the decision-making process and understand the goals of planned activities and consider them meaningful, they are more likely to contribute to these goals. When the strategic goals are unclear, the teachers seem to try everything (e.g., participate in projects, test different methods, etc.), but the usefulness and sustainability of it all remains weak. "*Then we've had learning via Skype /.../, integrated learning, we ordered an external lecturer for that... Then we've tried open education... We've also done lots of these kind of days of integrated learning to get teachers working together, to get them to mix and mingle, participating in projects, so we've had project days with different schools and we've also done a lot of outdoor learning days, internally, kind of collaboration days /.../ so we try to bring in these new methods every way we can...*" (5). Agreeing with every small change without any clear goal fragments teachers' capacity to work and decreases their motivation.

Only three of the school leaders reported involving teachers in implementing changes in which teachers' opinions were important. In these schools, teachers were also more proactive, and the leaders were more satisfied with teacher collaboration.

Interestingly, only one of the school leaders described how students were also involved in leading the school and how realising the development goals of the school was supported through the student council: "*…through the student council, time after time we try to remember, that you know, we've got this development goal (students taking responsibility for learning), and think by yourselves now how can you talk to the students, how can you first assume the right attitude yourselves and in this way influence the opinion of others...*" (9). The student council was involved in establishing differentiated pay for teachers: "*I allowed the students to discuss how the school leader should assess teachers contributions...*" (9).

To sum up, if the vision is vague (or there is none) and values are not dealt with, they are not sufficiently discussed and shared. This results in teachers feeling less responsible for changes and showing little initiative for collaboration, while their activities are fragmented and random. Although learners' development and shaping an environment conducive to learning was important for all of the school leaders, the development goals in most of the schools were unclear, and leadership was shared only among the management team. The interviewees described many different activities and projects, but it remained unclear why and how these undertakings helped to achieve the development goals. In addition, it seems that although there were systems for sharing everyday information, few of the school leaders had created favourable conditions for teachers' collaborative learning; for example, effective forms such as mentoring and co-teaching were not mentioned by the interviewees.

Improving the quality of teaching and learning

The quality of education depends on teachers' competence; therefore, improving the quality of teaching and learning is one of the most essential tasks for a school leader. However, it seems that only three of the nine school leaders included in the survey had a systematic approach to supporting teachers' professional development. These leaders mentioned appraisal interviews, lesson observations, giving feedback, and interviews based on teacher's self-assessment (the latter were also used for choosing training courses). Even though all of the school leaders talked about teachers' training needs and joint training, it remained unclear how these decisions were taken, or as one school leader said: "*...this kind of constant training and this is what we do, but I am not saying that it is very systematic and planned or something...*" (4).

Most of the school leaders emphasised teachers' freedom in choosing training courses. Nevertheless, too much autonomy may not support teachers if they lack a feeling of competence or connection: "*…people are kind of autonomous in the choices about self-development, like you can do what you want, you are a professional /…/ But sometimes, what comes along with it is that people feel a bit left alone, in the sense that the whole subject-field is on your shoulders. This is your business. And then when something gets weird, then there will be this kind of insecurity, I mean this system suits strong personalities...*" (6). It remained unclear how teachers made their choices, as most of the schools included in the study did not have a feedback system for teachers and lacked collaboration, while only four of the school leaders mentioned having developmental discussions.

A learning-centred school culture presupposes the constant monitoring of learning processes, data collection and analysis. However, only one of the school leaders highlighted the necessity for monitoring and collecting data to establish future directions in order to improve the learning process: "*…it's a matter of long-term perspective /.../ what are the things I can already work on, look at the number of students in the county, trends, determine how we accept students into our school, how many students there are, what they study, what they actually will need to know when they finish school, what the teaching methods could be, the content /…/ or look at how old my teachers are, uhm, what their training needs are, how to support them, help them, when some of them might say that okay, we're going to retire now, whether I've got some younger teachers who could take their place or I have to start looking for a replacement...*" (7). With regard to developing everyday practices: "*…this is again that example about surveys that when there is direct feedback that we've got a problem with collaboration, then we start to think, what are those points where we could improve that /.../ For* *improving that (teaching quality), you need feedback. This is where these surveys come into play, and it's no secret that examination results, contests, competitions etc. give an insight into such things*" (7). It can be said that the interviewed school leaders had a limited awareness about evidence-based leadership in school development; decisions and choices were made based on instincts rather than analysis of data.

Constant improvement of learning assumes that teachers receive feedback on their work. Among the studied school leaders, three stood out: in these schools, teachers received feedback via developmental discussions and lesson observations. In this regard, one of the school leaders, who was talking about developing a feedback system, commented: "*The thing with feedback systems is that you've got to be very careful when implementing them because at a certain point people get the feeling that they're going to check us. It's very much like a kind of fine mechanics...*" (6).

We can nevertheless conclude that improving the quality of teaching and learning was generally implemented in a chaotic manner in the schools studied. Processes of monitoring trends and collecting evidence in a more informed manner took place in only one school. Providing feedback for teachers and supporting their development seems to be random or largely the teachers' own responsibility.

Modelling and building trusting relationships

All of the school leaders emphasised the significance of relationships and had made efforts to improve relationships in their schools. However, the leaders' behaviour differed, and those in organisations where good relationships were valued clearly stood out from the others. The team was harmonious and teachers showed considerable initiative: "*...our teachers-staff members are really great leaders, as a teacher should be, and they often love to take the leader's role and implement their ideas*' (9); "*...I absolutely love these moments when we are all together and do some things together, I mean the kind of nice things of being and doing things together, whether it's sports day or some excursion or our school assemblies...*" (7).

However, several of the school leaders expressed scepticism regarding teachers' readiness and willingness to collaborate: "*...collaboration between people is so scattered*" (3). Others admitted that these processes had not gained ground as expected: "*But I think that one thing we haven't got running very well is the kind of working groups that do something throughout the whole year*" (6).

Top-to-bottom initiative does not increase teachers' trustful relations: "*…project day, this forces them again into some kind of collaboration. But well,* *we do have some bottlenecks, where nobody really wants to take responsibility for organising these project days /…/ so what we do at least is that we plan these things until the end of the school year, who does what, and topics are also more or less settled, but who leads this, who takes responsibility, like everyone does it, but in the end, it's the same people who do it'* (4). For motivation, it is essential that teachers sense the support of colleagues and management, the necessity and usefulness of changes in their daily work, and freedom of choice and opportunities, i.e., autonomy. In order to inspire teachers to develop their practice and try new things, the environment must be conducive to learning, as one school leader emphasised: "*…we allow a child to make mistakes, right, because through mistakes we learn; the same applies to all of us, that we all still make mistakes from time to time*" (8).

The importance of recognition for a supportive school climate was especially stressed by one of the school leaders: "*…we talk about things a lot. And it starts with the fact that we have long information meetings twice a week, 40 minutes, and there we haven't got this kind of boring and dry information sharing, but rather people talk about what they have done, where they have been with their students, how these events have succeeded, what results students have achieved in competitions, and the head of the school also has an opportunity to praise everyone, and by the way, the employees really like it, the fact that they are recognised right there in front of their colleagues*" (7).

It is also crucial for a school leader to be a pedagogical role model. As one of the school leaders said, his/her goal was to show that achieving certain development goals is possible for teachers: "*…firstly, I became a class teacher. I will show them that appraisal interviews and forming bonds is my first priority*" (1). Another leader emphasised the need to be updated and not to lose contact with teachers and students: "*…I want to be in front of the class for exactly this reason so that I would not lose touch with students or with teachers…otherwise, we would become distant*" (7).

Leaders should contribute to the development of trusting relationships by serving as an example when communicating with colleagues and students. As one of the school leaders noted: "*We are all different /…/ it isn't only important that we say that we care; it also needs to be obvious, people need to feel that they are cared for. Whether you are a child or an employee or a parent…*" (8).

Based on the analysis of the interviews, it seems that the relationships are better, and the willingness to collaborate is greater, in schools where the goals are clear and shared among the staff. Clear goals increase trust in the leader and collaborative activities are more focused, thus increasing teachers' willingness to take the initiative in the school's development. It could be assumed that school leaders' work experience or school size could influence the quality of leadership activities and the ability to create a shared vision and a collaborative learning-centred school culture. Surprisingly these patterns did not emerge in the results. The results of our study confirm those of other studies that found that there are more differences within the studied school leader group than between groups with a different level of experience (e.g., Barnett et al., 2012). Specifically, the views of the three school leaders with only 1–2 years of experience differed significantly: only one of them viewed the role of the school leader as mainly managerial, while the other two viewed the role primarily as a shaper of collaborative school culture and a substantial leader in school development. As this study focused primarily on school leaders' perceptions of the leader's role, further research is needed to determine the extent to which the ideas reflected in the interviews are realised in everyday school management activities.

Concluding remarks and further perspectives

A collaborative school culture presupposes shared goals and responsibility, encouraging teamwork and relationships, supporting everyone's development, and providing feedback and recognition in an atmosphere of trust. To summarise the findings, it appeared from the interviews that even though all of the nine school leaders included in the study valued collaboration and high-quality teaching and learning, only three of them had a clear vision and goal that was actually used as the basis for everyday decisions. The leaders who had a clear vision for school development supported collaboration in a more systematic and diverse manner, and described a well-functioning team and close collaboration in the school. An unclear and blurred perception, however, leaves space for uncertainty and irresponsibility. Therefore, the leaders who did not have a clear vision were mainly sceptical about the relationships between the teachers and described teamwork in their school as meagre. Even though these school leaders believed that teamwork was crucial, it remained unclear why that was the case and what could be done to improve the situation. The ability to create and communicate a shared vision, and to give meaning to it through actions, is considered one of the characteristics of an effective school leader (Leithwood et al., 2008; Kruse & Louis, 2009).

The role of leadership in shaping a collaborative school culture should not be underestimated, as it encourages sharing responsibility (Vangrieken et al., 2015) and promotes the sustainability of changes (Harris, 2005). From the interviews, it appeared that all of the school leaders exercised shared leadership to some extent, although it was frequently used in a top-to-bottom manner and without much thought on how to increase teachers' motivation and involvement in daily school management. If responsibility is shared, it is more likely that the organisation can guarantee professional development at the organisational as well as the individual level. Forming working groups inside the organisation is one way to develop relationships and increase the effectiveness of collaboration, and to support teachers' professional development (Kruse & Louis, 2009). All of the leaders pointed out that team training sessions and excursions were organised in their school to a greater or lesser extent. At the same time, supporting teachers' professional development was unsystematic, and the formation of working groups tended to be a management initiative. There needs to be an increase in school leaders' knowledge and skills regarding the creation and support of learning communities, and the strengthening of teacher agency and motivation to work in such communities.

If the goal is to change the school culture, it is essential for leaders to have a better understanding of their role in leading the main process of a school, that is, the learning process. Leading and developing teaching and learning is the main focus of a pedagogical leader; in the present study, this was most clearly seen in the case of three of the school leaders. Moreover, the data of TALIS 2013 indicated that school leaders who focused more on supporting teaching and learning were more likely to take specific measures to influence teachers' work and careers (Übius et al., 2014). Although the school leaders in the present study felt that changes in the learning process were necessary, most of them delegated the task of developing the learning process to the head teacher. It is very likely that this is because such tasks have traditionally been the duty of the head teacher in Estonian school culture.

In order to develop the learning process in school, leaders should gather and analyse data to make decisive changes. Only one of the school leaders talked about the importance of research, evidence and data analysis in making decisions about long-term development strategies, as well as everyday teaching and learning. Nowadays, capacity building, inquiry-oriented practice, and data-driven decisions are considered to be central themes of educational improvement (Hargreaves & Shirley, 2012; Louis & Stoll, 2007). Teachers should therefore also be informed about what kind of evidence they should collect so that the effectiveness of the practices implemented can be analysed (Datnow, 2011).

Practical implications for strengthening collaborative learning-centred school leadership

Based on the interviews, school leaders need support and training in two particular areas: 1) creating a shared school vision and putting it into practice;

and 2) acquiring the knowledge and skills required to systematically improve the quality of learning and teaching.

Development programmes for school leaders should support them in developing a pedagogical vision and should provide tools for collecting data and analysing the progress towards the established goals. With regard to practical measures, implementing school leaders' self-assessment as a part of internal assessments of schools should be considered, in addition to training and development opportunities. There are different self-assessment models available aimed at helping educational leaders to assess their personal mastery (e.g., Gregorzewski et al., 2018) and to analyse their activities with regard to the school's development goals, in order to gain feedback on their work and professional development.

Another aspect that should receive more attention in the training of school leaders is how to enhance teachers' professional development so as to increase school capacity and accomplish organisational goals. According to the present study, some school leaders seemed to believe that their role was mainly to find resources to implement a variety of innovations, but they failed to see that this kind of approach – introducing multiple forms of new initiatives without a clear vision – leads to a fragmented set of chaotic activities that are unable to support teacher motivation, commitment and learning. Development programmes for school leaders should therefore place more emphasis on how to enhance schools' capacity to build a shared commitment to school goals, and to create structures and work conditions that systematically promote teacher collaboration and learning. The dimensions and practices associated with collaborative and learning-centred leadership in this study may offer a starting point in this regard.

The limitations of the present study clearly suggest an opportunity for further research: the interviewees' answers may not have reflected all of the activities undertaken in their schools. A more broad-based study should therefore be conducted among school leaders in order to gain a deeper insight into the main challenges for school leaders in developing a collaborative school culture.

Acknowledgements

This research was supported by the Tallinn University Research Fund. We would like to thank all of the school leaders who participated in the study.

References

Angelides, P. (2010). The efficacy of small internal networks for improving schools. *School Leadership & Management, 30*(5), 451–467.

Barber, M., Whelan, F., & Clark, M. (2010). *Capturing the leadership premium. How the world's top school systems are building leadership capacity for the future*. McKinsey & Company. https://www.mckinsey.com/~/media/mckinsey/industries/social%20sector/our%20insights/capturing%20the%20leadership%20premium/capturing%20the%20leadership%20premium.ash

Barnett, B. G., Shoho, A. R., & Oleszewski, A. M. (2012). The job realities of beginning and experienced assistant principals. *Leadership and Policy in Schools*, *11*(1), 92–128.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101.

Bruggencate, G., Luyten, H., Scheerens, J., & Sleegers, P. (2012). Modeling the influence of school leaders on student achievement: How can school leaders make a difference? *Educational Administration Quarterly, 48*(4), 699–732.

Bush, T., & Glover, D. (2014). School leadership models: What do we know? *School Leadership & Management*, *34*(5), 553–571.

Campo, C. (1993). Collaborative school cultures: How principals make a difference. *School Organisation, 13*(2), 119–127.

Datnow, A. (2011). Collaboration and contrived collegiality: Revisiting Hargreaves in the age of accountability. *Journal of Educational Change, 12*(2), 147–158.

Day, C., & Sammons, P. (2006). *Successful leadership: A review of the international literature*. CfBT Education Trust. https://www.researchgate.net/publication/238717790_Successful_School_Leadership_What_it_Is_and_How_It_Influences_Pupil_Learning

Eisenschmidt, E., Reiska, E., & Oder, T. (2015). Algajate õpetajate tajutud juhtkonna tugi ning selle seosed kooliarendusse kaasatuse ja õpetajate koostööga [Novice teachers perceived leadership support and its relation to involvement in school development and teacher collaboration]. *Eesti Haridusteaduste Ajakiri, 3*(1), 148–172.

Engels, N., Hotton, G., Devos, G., Bouckenooghe, D., & Aelterman, A. (2008). Principals in schools with a positive school culture. *Educational Studies, 34*(3), 159–174.

Estonian Ministry of Education and Research. (2014). *Eesti elukestva* õppe *strateegia 2020* [Estonian lifelong learning strategy 2020]. https://www.hm.ee/sites/default/files/strateegia2020.pdf

Estonian Parliament. (2010). *Põhikooli- ja gümnaasiumiseadus* [Basic School and Gymnasium Act]. https://www.riigiteataja.ee/akt/13332410?leiaKehtiv#jg7

Fullan, M. G., & Hargreaves, A. (1991). *What's worth fighting for? Working together for your school*. Ontario Public School Teachers' Federation.

Gregorzewski, M., Schratz, M., & Wiesner, C. (2018). Exploring the personal mastery of educational leaders: FieldTransFormation360 and its validation in the Austrian Leadership Academy. *CEPS Journal, 8*(3), 59–78.

Hallinger, P. (2011). Leadership for learning: Lessons from 40 years of empirical research. *Journal of Educational Administration, 49*(2), 125–142.

Hallinger, P., Piyaman, P., & Viseshsiri, P. (2017). Assessing the effects of learning-centered leadership on teacher professional learning in Thailand. *Teaching and Teacher Education*, *67*(1), 464–476.

Hargreaves, A., & Shirley, D. (2012). *The global fourth way: The quest for educational excellence.* Corwin.

Harris, A. (2005). Leading or misleading? Distributed leadership and school improvement. *Journal of Curriculum Studies, 37*(3), 255–265.

Heck, R. H., & Hallinger, P. (2009). Assessing the contribution of distributed leadership to school improvement and growth in math achievement. *American Educational Research Journal*, *46*(3), 659–689.

Innove. (2016). *Haridusasutuse juhi kompetentsimudel* [The principal's competence model of the educational institution]. https://www.innove.ee/wp-content/uploads/2019/03/Haridusasutuse-juhi-kompetentsimudel4.pdf

Kohm, B., & Nance, B. (2013). Creating collaborative cultures. *Educational Leadership, 67*(2), 67–72.

Kruse, S. D., & Louis, K. S. (2009). *Building strong school cultures: A guide to leading change*. Corwin.

Lai, C., Li, Z., & Gong, Y. (2016). Teacher agency and professional learning in cross cultural teaching contexts: Accounts of Chinese teachers from international schools in Hong Kong. *Teaching and Teacher Education, 54*, 12–21.

Leithwood, K., Patten, S., & Jantzi, D. (2010). Testing a conception of how school leadership influences student learning. *Educational Administration Quarterly, 46*(5), 671–706.

Leithwood, K., Harris, A., & Hopkins, D. (2008). Seven strong claims about successful school leadership. *School leadership and management, 28*(1), 27–42.

Li, L., Hallinger, P., & Walker, A. (2016). Exploring the mediating effects of trust on principal leadership and teacher professional learning in Hong Kong primary schools. *Educational Management Administration & Leadership, 44*(1), 20–42.

Liu, S., Hallinger, P., & Feng, D. (2016). Supporting the professional learning of teachers in China: Does leadership make a difference? *Teaching and Teacher Education, 59*, 79–91.

Louis, K., & Stoll, L. (2007). *Professional learning communities: Divergence, depth and dilemmas*. Open University Press & McGraw Hill Education.

Murphy, J., Elliott, S. N., Goldring, E., & Porter, A. C. (2007). Leadership for learning: A research-based model and taxonomy of behaviours. *School Leadership and Management, 27*(2), 179–201.

Nevgi, A., & Korhonen, V. (2016). Pedagoginen johtaminen yliopiston keskijohdon johtamistyössä [Pedagogical leadership in university leadership management]. *Kasvatus, 47*(5), 429–433.

OECD. (2014). *A teachers' guide to TALIS 2013: Teaching and learning international survey*, TALIS, OECD Publishing. https://doi.org/10.1787/9789264216075-en

OECD. (2019). *PISA 2018 Results (Volume I): What students know and can do.* PISA, OECD Publishing. https://doi.org/10.1787/5f07c754-en

Opfer, V. D., Pedder, D. J., & Lavicza, C. (2011). The role of teachers' orientation to learning in professional development and change: A national study of teachers in England. *Teaching and Teacher Education, 27*, 443–453.

Robinson, V. M. J., Hohepa, M., & Lloyd, M. (2009). *School leadership and student outcomes: Identifying what works and why, best evidence synthesis.* Ministry of Education.

Robinson, V., Lloyd, C., & Rowe, K. (2008). The impact of leadership on student outcomes: An analysis of the differential effects of leadership types. *Educational Administration Quarterly*, *44*(5), 564–588.

Sales, A., Moliner, L., & Amat, A. F. (2017). Collaborative professional development for distributed teacher leadership towards school change. *School Leadership & Management, 37*(3), 254–266.

Tschannen-Moran, M. (2009). Fostering teacher professionalism in schools: The role of leadership orientation and trust. *Educational Administration Quarterly, 45*(2), 217–247.

Übius, Ü., Kall, K., Loogma, K., & Ümarik, M. (2014). *Rahvusvaheline vaade* õpetamisele *ja* õppimisele. *OECD rahvusvahelise* õpetamise *ja* õppimise *uuringu TALIS 2013 tulemused* [An international view of teaching and learning. Results of the OECD International Study on Teaching and Learning TALIS 2013]. SA Innove.

Vangrieken, K., Dochy, F., Raes, E., & Kyndt, E. (2015). Teacher collaboration: A systematic review. *Educational Research Review, 15*, 17–40.

Woodland, R., Kang Lee, M., & Randall, J. (2013). A validation study of the Teacher Collaboration Assessment Survey. *Educational Research and Evaluation, 19*(5), 442–460.

Youngs, P., & King, B. M. (2002). Principal leadership for professional development to build school capacity. *Educational Administration Quarterly, 38*(5), 648–670.

Yukl, G. (2012). Effective leadership behavior: What we know and what questions need more attention. *Academy of Management Perspectives, 26*(4), 66–85.

Zahed-Babelan, A., Koulaei, G., Moeinikia, M., & Sharif, A. R. (2019). Instructional leadership effects on teachers' work engagement: Roles of school culture, empowerment, and job characteristics. *CEPS Journal, 9*(3), 137–156.

Biographical note

Katrin Poom-Valickis, PhD, is a professor of Teacher Education at Tallinn University. She has been a project leader and co-coordinator in several projects related to the development of Teacher Education and member of the expert group who prepared the implementation of the induction year for novice teachers in Estonia. The main research interest is primarily focused on the teacher's professional development, more precisely how to support future teachers' learning and development during their studies and first years of work.

Eve Eisenschmidt, PhD, is a professor of Educational Leadership at Tallinn University. Her research areas are the professional development of teachers, school development and leadership. She has been active in development of teacher education policy on the national level, e.g. preparing the implementation of the induction programme for beginning teachers in Estonia and developing initial teacher education programmes at the university.

Ann Leppiman – M.A. in Educational Leadership. Ann was working 2017-2020 at Tallinn University as development manager for academics. Currently she serves as an adviser for the minister of Ministry of Education and Research. Ann's research interest is connected with school leadership and school culture, pedagogical development and support for academics.

doi: 10.26529/cepsj.958

Integrating Assessment for Learning into the Teaching and Learning of Secondary School Biology in Tanzania

Albert Tarmo[1]

The paper is about a study that investigated how the integration of assessment for learning enhances learning achievement among secondary school biology students in Tanzania. A quasi-experimental design involving pre-test and post-test of non-equivalent control and experimental groups was used to ascertain how the integration of assessment for learning into teaching and learning processes enhances students' learning achievement. Two boarding secondary schools located in the suburbs of Dar Es Salaam were selected. Students in the two schools had maintained equivalent performances in national examinations in previous years. The results showed that the students taught using teaching and learning processes integrating assessment for learning outperformed those taught using conventional approaches. The integration of assessment for learning is likely to have contributed to the higher learning achievement in the experimental group. The study contributes to our understanding of how teachers in resource-constrained classrooms can integrate assessment for learning techniques into their day-to-day lessons, thereby harnessing the power of assessment to enhance learning and raise standards.

Keywords: assessment, educational assessment, learning achievement, Tanzania

1 School of Education, University of Dar es Salaam, Tanzania; tarmo50@outlook.com.

Vključevanje ocenjevanja za učenje v poučevanje in učenje pri pouku biologije v Tanzaniji

Albert Tarmo

Prispevek predstavlja raziskavo, ki je preučevala, kako vključevanje ocenjevanja za učenje izboljša učne dosežke učencev biologije v Tanzaniji. Uporabljen je bil kvazieksperimentalni model, ki je vključeval pred- in potest neekvivalentne kontrolne in eksperimentalne skupine z namenom ugotavljanja, kako vključevanje ocenjevanja za učenje v proces poučevanja in učenja izboljša učne dosežke. Izbrani sta bili dve šoli iz predmestja Dar Es Salaam. Učenci obeh šol so imeli enake dosežke pri nacionalnem preverjanju znanja v letu pred izvedbo raziskave. Rezultati so pokazali, da so bili uspešnejši učenci iz skupine poučevanja z uporabo ocenjevanja za učenje kot tisti, ki so bili deležni poučevanja s konvencionalnimi pristopi. Vključevanje ocenjevanja za učenje je verjetno prispevalo k boljšim učnim dosežkom eksperimentalne skupine. Raziskava prispeva k razumevanju, kako lahko učitelji v z viri omejenih okoliščinah vključujejo ocenjevanje za učenje v vsakodnevno poučevanje ter tako izkoristijo moč ocenjevanja za izboljšanje učenja in dvig standardov.

Ključne besede: ocenjevanje, izobraževalno ocenjevanje, učni dosežki, Tanzanija

Introduction

Assessment is widely considered a powerful tool for enhancing students' learning achievement when embedded in the teaching and learning process (Black & Wiliam, 2018; Ellegaard et al., 2018; Wiliam et al., 2004; Wiliam, 2011). When integrated into the teaching and learning process, assessment serves to elicit evidence about students' learning progress. For example, assessment during instruction provides opportunities for students to display their understanding and uncover the strengths and weaknesses of their thinking (Greenstein, 2010). Teachers and learners can use such evidence to make decisions about subsequent learning steps (Wiliam, 2011). For instance, teachers can tailor subsequent lessons in response to students' learning needs and support their students' cognitive growth (Greenstein, 2010).

Moreover, when teachers ask questions to elicit students' prior experiences at the start of a new lesson, they generate evidence that becomes readily available to inform instructional decisions. Teachers can use information about students' prior knowledge to determine students' learning needs for a new lesson. For both teachers and students, learning needs are the bases for planning lessons and setting learning objectives and expectations. When such objectives are made explicit to students, they can take charge of their learning and work towards meeting these expectations (Wiliam, 2011). Most importantly, information about students' prior knowledge helps both teachers and students make connections between previous lessons and new topics to enhance meaningful learning (Greenstein, 2010).

Assessment during teaching allows teachers to continuously check students' learning and adjust instructional processes to meet learners' just-in-time needs (Wylie & Lyon, 2015). For example, teachers may intersperse their verbal descriptions with questions-and-answers to test students' comprehension of a topic. As students respond to questions, teachers can spot individual learners who are struggling to learn certain concepts or skills. In such cases, teachers may provide feedback that identifies gaps in students' thinking and redirect learning by showing the next steps students need to follow (Greenstein, 2010).

Using feedback mechanisms, teachers can focus students' attention on the areas in which they have demonstrated learning success and those that require more practice. Moreover, teachers can support learners to devise learning plans for achieving desired learning outcomes. Teachers facilitate students in directing their learning and make them active participants when they supply the required information about learning progress and provide support for subsequent learning steps. Most importantly, evidence of learning success motivates students and enhances their resilience when faced with learning difficulties that are within their capabilities (Berry, 2008). Generally, assessment shapes subsequent instruction and learning when teachers and students have continuous access to evidence showing learners' current levels of learning.

Formative assessment (FA) and assessment for learning (AfL) are two closely related and widely used concepts to describe the use of assessment to enhance future teaching and learning (Black & Wiliam, 2009; Wiliam, 2011). The meanings of these concepts remain widely debated (Jonsson et al., 2015; Hopfenbeck, 2018; Wiliam, 2011). Black and Wiliam (2005) defined FA as activities by teachers and students aimed at generating information about students' leaning progress and the use of such information as feedback to modify teaching and learning processes to meet learners' needs. In the contemporary literature, however, the term "assessment for learning" rather than "formative assessment" is favoured for describing assessment that promotes learning (Black & Wiliam, 2018; Broadfoot et al., 2002; Hopfenbeck, 2018). This is because FA is used in diverse ways. For example, in some contexts, FA is conceived as "early warning summative" assessment that provides information about the "likely performance of students on the state mandated tests" (Wiliam, n.d., p. 4). Feedback is given to students telling them the items they got right and wrong regardless of the use they make of such feedback (Wiliam, n.d.). In the Tanzanian context, FA often means regular monthly, terminal and annual testing to reduce overdependence on the single final examination that students sit at the end of each education cycle (Kyaruzi et al., 2018). On the other hand, the Assessment Reform Group defined AfL as "the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there" (Broadfoot et al., 2002, p. 2). The present study uses the term "assessment for learning"' and draws on the key attributes of assessment that enhance learning as summarised by the Assessment Reform Group (Broadfoot et al., 2002).

The idea of integrating assessment into instruction to enhance learning has been widely embraced at the national and regional levels (Hopfenbeck & Stobart, 2015), to the extent that its adoption has been described as a "research epidemic" (Steiner-Khamsi, 2004, p. 2). AfL is a tool for enhancing learning by making learning expectations explicit to students and providing them with continuous feedback in order to inform them about their learning progress and the next steps they need to take to improve their learning achievement (Hopfenbeck, 2017). Cases of large-scale implementation include Sweden (Jonsson et al., 2015) and four high-needs US districts (Wylie & Lyon, 2015), where AfL successfully transformed assessment practices and improved the collection of evidence about students' learning through questions-and-answers. In the Tanzanian context, AfL has received policy attention despite the scarcity of exemplary implementation practices at the classroom level, as further discussed below.

Learning Assessment in Tanzania

In the latest revision of the secondary education curriculum, the government of Tanzania stressed the need to integrate assessment activities with everyday instruction using authentic approaches such as practical tasks, project work, portfolios and verbal questioning (Ministry of Education and Vocational Training (MoEVT), 2007). The aim was to widen the range of learning achievement that could be assessed and use the information to guide and improve teaching and learning processes. Such assessment is aimed at promoting learning through building confidence and developing students' belief in their capacity to attain learning success. This assessment is envisioned to be formative in nature, as it monitors learning progress throughout a given education cycle (MoEVT, 2007). Generally, the curriculum calls for a change in assessment approach by adopting AfL and minimising overdependence on paper-and-pencil tests. However, local research suggests that efforts to improve learning achievement rarely make use of assessment as a means of raising standards (Kira et al., 2013; Kitta & Tilya, 2010; World Bank, 2008). High-stake, large-scale and centrally administered examinations, which are used for certification and placement purposes, remain dominant in Tanzania (Kyaruzi et al., 2019). Such examinations have lasting effects on students' life chances because the results are used to select students for highly valued places in further education and workplaces.

The government introduced Continuous Assessment (CA) to reduce overdependence on high-stake examination, assess students on a continual basis, and combine results with those obtained in final examinations to determine students' final grades (Kyaruzi et al., 2019). However, studies suggest that teachers often do not implement CA in such a way that the information collected could be used to improve instruction (Lema & Maro, 2018). Instead, teachers' assessment practices largely mimic the system-wide high-stake examinations. At the classroom level, paper-and-pencil assessment through quizzes, tests and examinations, which assesses memorisation and test-taking skills, dominates. Classroom observation studies suggest that during actual teaching, teachers largely ask closed questions and favour single answers, often known beforehand (UNICEF Tanzania, 2018). Classroom questioning often involves inviting students in turns to give answers until the correct answer that the teacher favours is provided. Teachers either do not provide feedback or provide only general feedback indicating the gaps in students' knowledge that made them give incorrect answers (Lema & Maro, 2018). Furthermore, paper-and-pencil assessment provides limited useful information for teachers and students to adjust instructional processes in ways that can improve achievement (Kippers et al., 2018). Paper-and-pencil assessment provides scores and grades, which are not particularly useful in guiding instructional improvements.

Since school success is typically judged based on students' performance in high-stake examinations, teachers are often compelled to resort to teaching to the test instead of promoting meaningful learning (O-saki & Njabili, 2003). They train students' techniques for answering examination questions instead of facilitating the development of higher-order skills as stipulated in the curriculum. Often teachers do not teach topics that are not tested in the national examination, or give them only marginal attention (World Bank, 2008). Moreover, the emphasis on grades as a determinant of access to higher education and employment often drives students to strive for higher grades instead of a deeper understanding of school subjects. When classroom cultures reward "gold stars" through grades or ranks, "students often play dirty to score higher grades" (Black & Wiliam, 2005).

Generally, the envisioned transformation in assessment practice through the adoption of assessment techniques that enhance learning achievement remains largely unrealised. Most importantly, the curriculum lacks practical examples showing how assessment reforms can be implemented in classrooms. Moreover, teacher education courses often focus on standardised assessment methods and how to enhance their psychometric properties (Kyaruzi et al., 2019). In this context, where teachers often lack assessment skills, the most logical option for teachers is to rely on traditional assessment approaches, mainly the tools provided by textbooks and instructional material publishers (Lema & Maro, 2018), which often replicate high-stake national examinations. Furthermore, there are relatively few studies on how teachers can integrate AfL into classroom lessons in the Tanzanian context (Kyaruzi et al., 2018, 2019; Lema & Maro, 2018). Thus, there is scant evidence regarding how teachers in resource-constrained classrooms can integrate AfL into their lessons and how AfL contributes to students' learning achievement in such contexts (Kyaruzi et al., 2019). It is therefore imperative that more research focusing on this be conducted.

Two studies conducted by Kyaruzi et al. (2018, 2019) explored teachers' and students' perceptions of FA and how these perceptions predicted self-professed feedback use and student performance. The results suggested that the perceived quality of teacher feedback predicted feedback use and student performance. Moreover, teachers claimed to formatively use assessment information for self-reflection, improving their approaches, correcting errors and conducting remedial classes to support weaker students. They further reported summative use of assessment information such as ability grouping, accountability reporting and reprimanding low achievers. These findings are limited, however, as no attempts were made to observe whether teachers' favourable perceptions of FA and their avowed use of feedback manifested in actual practice.

Ethnographic studies of teachers' practice in Tanzania suggest that while teachers may verbally commit to innovative pedagogies, their actual classroom practices often contrast with their perceptions (Vavrus, 2009; Vavrus & Bartlett, 2012). Indeed, findings from classroom observations by Lema and Maro (2018) and UNICEF Tanzania (2018) contradict teachers' and students' avowed use of assessment information, as reported by Kyaruzi et al. (2018, 2019). Lema and Maro (2018), for example, observed that teacher feedback constituted exclamatory verbal comments such as "excellent", "very good", "good try" and "that's fair" for students who answered questions correctly, whereas for those who got questions wrong teachers commented "work hard", "lazy" and "poor". Similarly, UNICEF Tanzania (2018) reported that teachers often gave very general feedback to explain why students made mistakes or answered questions incorrectly. Together these studies suggest that teachers lack skills for providing constructive feedback to help students improve their learning.

It was against this background that the present study redesigned biology teachers' lessons, integrating AfL techniques into the teaching and learning process to exemplify how teachers in resource-constrained schools in Tanzania can use AfL in actual lessons (see section 2.2). The aim was to assess the contribution of integrating AfL into the instructional process to students' learning achievement in biology. The question addressed was: What is the contribution of integrating AfL techniques into the teaching and learning process to students' learning achievement?

Method

A quasi-experimental design involving pre-test and post-test of non-equivalent control and experimental groups was used to establish how the integration of AfL into the teaching and learning of secondary school biology enhances students' learning achievement. Non-equivalent control and experimental group design is a form of quasi-experimental design in which the participants cannot be assigned randomly into experimental and control groups simply because the researcher has no control over the randomisation of treatment, unlike in true experimentation (Mitchell & Jolley, 2010). This was the most feasible design for the school context in which students were organised in intact streams. In such a setting, the random placement of students into control and experimental groups was restricted, as it could have caused learning disruption. Therefore, two intact streams of students, each from a different school, were randomly designated as experimental group (N = 44) and control group (N = 45) by tossing a coin. The use of existing streams also maximised the ecological validity of the findings.

Research setting

The setting was two boarding secondary schools located in the suburbs of the metropolitan city of Dar Es Salaam, Tanzania. Over the previous five years, both schools had maintained an overall Grade Point Average of 4.6 in the national Certificate of Secondary Education Examination. Thus, the students in the two schools had equivalent academic performance. Furthermore, the schools had similar learning environments because both were located in different parts of the same ward, had relatively similar student populations, and had class sizes of 40–45 students. Both were government schools and thus had similar timetabling, teacher recruitment, remuneration and supply of resources. The matching of the groups based on various characteristics, as well as their random assignment into control and experimental groups, sought to further strengthen the equivalence (Mitchell & Jolley, 2010).

The study involved form one students aged 13–14 years. These students were about to begin learning the topic *Cell Structure and Organisation* (MOEC, 2005). This topic comprises abstract content, which makes it among the most difficult school biology topics for form one students to comprehend (Ozcan et al., 2014). In their study of students' perceptions of difficult biology topics, the researchers found that topics related to the cell, cell division, heredity, DNA and genetic code were among the most difficult to comprehend. The intervention procedures of the current study are described next.

Procedures

Designing lessons

The literature covering the key principles of AfL (Black & Wiliam, 2009; Broadfoot et al., 2002) and exemplary practices in various contexts (Hopfenbeck, 2018; Jonsson et al., 2015; Wylie & Lyon, 2015) was surveyed to identify guidelines for lesson design. Copies of lesson plans from previous years for the topic of Cell Structure and Organisation were then requested from biology teachers at ten schools in the same district. These were analysed to establish whether they reflected any of the principles and practices of AfL. Moreover, the teachers' lessons other than those covering Cell Structure and Organisation were observed and detailed notes were written to establish whether AfL practices were incorporated in their actual lessons.

Overall, the lesson plans had similar patterns and did not reflect any AfL practices (see Appendix I). Typically, the lessons began with an introduction in which the students reviewed the previous topic. The teacher-directed presentation of new content was interspersed with illustrative visuals and observations, followed by questions-and-answers. The lessons concluded with a summary of key points and instructions for the next lesson. Teachers predominantly asked closed questions requiring single-word or simple affirmative factual answers. Moreover, they mainly gave affirmative feedback using words such as "okay", "correct" or "exactly" to approve students' responses. These observations were consistent with recent research on teachers' assessment practices in Tanzania (Lema & Maro, 2018; UNICEF Tanzania, 2018). After the lesson analysis and observation, the lesson plans were redesigned to incorporate AfL techniques. Verbal questions were added with increased wait-time, rubrics, small project reports, observational checklists, presentations and worksheets in order to broaden the range of assessment formats. Opportunities for the collaborative setting of learning objectives, self- and peer review of work before submission, sharing of assessment criteria in the form of rubrics, and provision of written and verbal feedback were also included (see Appendix II). Assessment tools were constructed, such as worksheets, rubrics and observational checklists, which were used at different stages of the lesson during the intervention (see Appendix III). Finally, two lesson plan formats were established: plans with AfL techniques integrated and the original lesson plans the teachers provided.

The AfL techniques embedded in the redesigned lessons reflected the research-based principles of AfL in various ways. For example, the teachers in the experimental group assisted the students using questions to identify the learning objectives and activities they needed to perform. In addition, they provided assessment rubrics showing different levels of performance when they assigned class work. Such practices reflect the principle of AfL that states that lesson planning should include "strategies to ensure that learners understand the goals they are pursuing and the criteria that will be used to assess their work" (Broadfoot et al., 2002, p. 2). In this case, the collaborative setting of learning objectives was a strategy to help learners understand the learning goals and rubrics were intended to communicate the assessment criteria. The redesigned lesson plans were used with the experimental group and the original lesson plans that the teachers had provided were used with the control group.

Teacher training on the use of AfL

Four biology teachers from the school designated as the experimental group were invited to a week-long workshop on the principles and practice of AfL. In addition to in-depth discussion about AfL, its core principles and exemplary practices, the workshop involved orienting the teachers on how to implement the redesigned lesson plans and the challenges they were likely to face when implementing AfL techniques in their classroom contexts. Finally, the teachers were given copies of the redesigned lesson plans to implement according to their school subject timetables.

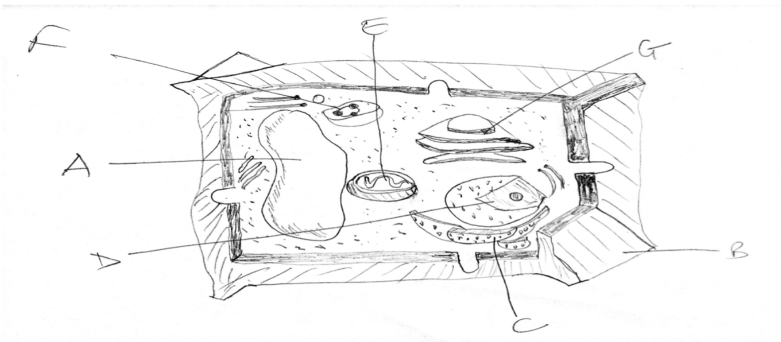
Designing the achievement test

Although the purpose of the AfL approach is to enhance authentic learning achievement (Wiliam et al., 2004), the students in both the control and experimental groups would eventually sit the National Form Two Examination, which largely tests their knowledge and understanding of biology concepts (Hakielimu, 2012). While AfL may have contributed both to the students' authentic learning and academic performance, the present study aimed to establish its contribution to their academic performance only. An achievement test was therefore constructed and used to measure the students' knowledge and understanding of Cell Structure and Organisation.

The test questions measured all of the learning objectives, covering definitions, characteristics, types and parts of cells, as listed in the syllabus under the topic Cell Structure and Organisation. The test was reviewed for content validity and error reduction by two experienced biology teachers. The necessary amendments were made following the review and the test was piloted in a secondary school comparable to the sampled schools. Immediately after the test, a reflective discussion focusing on the test's item clarity, difficulty and timing was held with ten randomly selected students from the pilot class. The test was then revised to create the final version, which was used as a pre-test and post-test. A typical test item is provided in Figure 1.

Figure 1
Typical Test Item

6. Carefully study the figure below and answer questions a-c @ **2 marks.**



a. This cell is typical of ________________cells

b. What evidence is there to support your answer in ‘a’ above?

    i. ______________________

    ii. _____________________

    iii. _________________________

c. Name the cell parts labeled A, B, C and F

The intervention

The experimental and control groups were pre-tested using the designed test to assess the prior learning achievement of the students before the topic Cell Structure and Organisation was taught. One teacher who had not participated in the training on AfL then taught the control group using the conventional lesson plan. Meanwhile, one of the four teachers who had participated in the training on AfL taught the experimental group using the redesigned lesson plans. The teachers who taught the control and experimental groups respectively were selected after carefully matching their demographics. They each held a Bachelor of Science with Education degree, had eight years of teaching experience, and were at the same salary level. These teachers had no other commitments apart from teaching and serving as class teachers.

In order to enhance the external validity of the results, the teaching in both groups followed the official syllabus and the school timetables. As per the syllabus, the topic Cell Structure and Organisation is supposed to be taught over four 80-minute periods (MoEC, 2005). These four periods cover three weeks of instructional time according to the school timetables. With an additional week for pre-testing and post-testing, the intervention lasted one month.

The researcher monitored teaching in the experimental group to verify that the redesigned lessons were implemented as intended. After teaching, the post-test was administered to both the experimental and control groups using the test described above. Pre-test and post-test scores were used to assess the difference in learning achievement between the control and experimental groups.

T-test analysis

The variation in the students' performance from pre-test to post-test in both the control and experimental groups was assessed using a paired sample t-test. Moreover, an independent sample t-test was used to ascertain whether the difference in mean scores between the experimental group and the control group was significant ($p < .05$). The aim was to establish whether the experimental group had higher learning achievement as a result of the treatment. Furthermore, the qualitative data that had been collected during the teachers' professional development and the monitoring of lesson implementation was analysed thematically following the example of Miles, Huberman and Saldana (2014). However, the present paper is based on the quantitative data.

Results

The study redesigned biology teachers' lesson plans to integrate AfL techniques into the teaching and learning process. Furthermore, it assessed the contribution to students' learning achievement of embedding AfL techniques in the instructional process. The results are presented next.

Difference in learning achievement for the control group before and after teaching

A paired sample t-test was used to compare the mean pre-test and post-test scores in order to determine whether the control group had a statistically significant difference in learning achievement before and after teaching using the conventional lesson plans. The results are presented in Table 1.

Table 1

Mean Pre-Test and Post-Test Scores for the Control Group

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pre-Test	14.12	45	4.82	.72
	Post-Test	18.62	45	4.46	.66

The results in Table 1 show that the mean post-test score was higher than the mean pre-test score, with a difference of 4.5. A paired sample t-test was performed to determine the statistical significance of the difference in mean scores between the pre-test and post-test. The results show a statistically significant increase in test scores from pre-test (M = 14.12, SD = 4.82) to post-test (M = 18.62, SD = 4.46), t (44) = -8.18, p < .001. The eta squared statistic (.6) indicated a large effect size. This suggests that the control group achieved some learning when the conventional lesson plans were used.

Difference in learning achievement for the experimental group before and after teaching

A paired sample t-test was used to compare the mean pre-test and post-test scores to determine whether the experimental group had a statistically significant difference in learning achievement before and after teaching. The results are presented in Table 2.

Table 2

Mean Pre-Test and Post-Test Scores for the Experimental Group

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pre-Test Scores	14.7	44	5.12	.77
	Post-Test Scores	33.18	44	9.21	1.38

The results (see Table 2) show that the mean post-test score was higher than the mean pre-test score with a mean difference of 18.48. A paired sample t-test was performed to determine the statistical significance of the difference in mean scores between the pre-test and post-test. The results show a statistically significant increase in test scores from pre-test (M = 14.7, SD = 5.12) to post-test (M = 33.18, SD = 9.21), t (43) = -14.995, p < .001. The eta squared statistic (.83) indicated a large effect size. This suggests that the experimental group achieved learning with the use of AfL-integrated lessons.

In both the control and experimental groups, there were gains in learning achievement. However, AfL-integrated lessons appear to have had a higher impact (eta squared = .83) compared to conventional lessons (eta squared = .6). In order to ascertain whether the difference in learning achievement between the experimental and control groups was statistically significant, an independent t-test was run to compare the mean post-test scores of the two groups. The results are presented next.

Difference in learning achievement between the control and experimental groups

Pre-test results for the experimental and control groups

In order to establish whether the students in both the control and experimental groups had the same level of prior knowledge and understanding of Cell Structure and Organisation, the mean pre-test scores of the two groups were compared. Table 3 shows the mean pre-test scores of the control and experimental groups.

Table 3

Pre-Test Mean Scores for the Experimental and Control Groups

	Group	N	Mean	Std. Deviation	Std. Error Mean
Pre-Test Scores	Experimental Group	44	14.7	5.12	.77
	Control Group	45	14.12	4.81	.72

The results displayed in Table 3 show that the experimental group had a mean score of 14.7 while the control group had a mean score of 14.12, with a mean difference of .57. In order to ascertain whether the mean difference in the pre-test scores between the experimental and control groups was statistically significant, an independent sample t-test was performed. The independent sample t-test for equality of means found no statistically significant difference in the mean scores between the experimental group (M = 14.7, SD = 5.12) and the control group (M = 14.12, SD = 4.81), t (87) = .553, p = .582. The magnitude of the difference in the means (mean difference = .58, 95% CI [-1.51, 2.67]) was very small (eta squared = .003). The results suggest that prior to the treatment, both the experimental and control groups had the same level of knowledge and understanding of Cell Structure and Organisation. The post-test was administered after teaching the topic using AfL-integrated lessons in the experimental class and conventional approaches in the control group. The post-test results are presented next.

Post-test results of the experimental and control groups

The mean post-test scores of the experimental and control groups were compared to assess the contribution to students' learning achievement of integrating AfL techniques into the teaching and learning process. The post-test results of the experimental and control groups are summarised in Table 4.

Table 4

Post-Test Mean Scores for the Experimental and Control Groups

	Class	N	Mean	Std. Deviation	Std. Error Mean
Post-Test Scores	Experimental Group	44	33.25	9.13	1.37
	Control Group	45	18.62	4.46	.66

The results in Table 4 show that the experimental group, which was taught using AfL-integrated lessons, had a mean score of 33.25, while the control group, which was taught using conventional lessons, had a mean score of 18.62. The mean difference between the two groups was 14.63.

In order to assess whether the mean difference in the post-test scores between the two groups was statistically significant, an independent sample t-test was carried out. The results showed that there was a statistically significant difference in the post-test scores between the experimental group (M = 33.25, SD = 9.13) and the control group (M = 18.62, SD = 4.46), t (62.12) = 9.569, $p <$ .001. The magnitude of the difference in means (mean difference = 14.63, 95% CI [11.57, 17.68]) was very large (eta squared = .51). The experimental group had a higher mean score compared to the control group. This suggests that the experimental group achieved higher learning compared to the control group. Higher learning achievement by the experimental group is likely to have been the result of the intervention, which involved integrating AfL techniques into the teaching and learning process, as discussed next.

Discussion

The most significant finding from this study is the higher learning achievement observed in the experimental group. Previous studies (Wiliam et al., 2004) show that teachers' use of AfL techniques in secondary school science and mathematics leads to increased quality of learning, and subsequently to higher learning achievement. The findings from the present study, which involved biology teachers in resource-constrained schools in the suburbs of Dar es Salaam, confirm those of previous studies. It is likely that embedding AfL techniques in biology lessons enhanced the learning achievement of the students in the experimental group in various ways.

First, asking open-ended, thought-provoking questions such as those indicated in the lesson plan (see Appendix II) is likely to have enhanced the students' mental engagement through classroom interactions and dialogues. Classroom interactions provide context for students to comment on each other's work, which makes them feel positive about their learning (Webb & Jones, 2009). In Tanzania, teachers often ask closed, factual questions with very brief wait-times (Kira et al., 2013). When no students volunteer to answer or when none answer as expected, teachers either seek answers from bright students or provide the correct answers themselves. This often limits classroom interactions to routinised, factual questions-and-answers with limited learning value (Hardman et al., 2012). The teachers in the experimental group allowed relatively more time for the students to think and generate well-thought-out ideas. In this way, these teachers demonstrated that they valued elaborate, well-thought-out contributions, as opposed to the short affirmative responses that characterise classroom questioning in Tanzania (Kira et al., 2013).

Second, although the lessons in both the control and experimental groups began with activities aimed at eliciting the students' prior knowledge of the topic, the teachers in the experimental group explicitly used the evidence of prior learning to plan the next learning steps (see Appendix II). In the control group, the teachers mostly adhered to the rigid lesson plans, regardless of the students' learning needs. Unlike those in the control group, the teachers in the experimental group not only shared the lesson objectives as indicated in the syllabus, but also collaborated with the students to adapt the lesson objectives in light of the students' prior knowledge of and experience with the topic. This collaborative setting of learning objectives enabled the students to understand what they were supposed to learn and to self-assess their progress accordingly (Wiliam et al., 2004). In this way, the teachers in the experimental group best served the students' learning needs. When students are involved in setting learning objectives, they adopt relevant strategies to learn and improve their achievement in spelling and punctuation (Black & Wiliam, 2018).

Third, while the teachers in the control group were mainly concerned with the correctness of the students' responses and taught to help the students produce correct answers known beforehand, the teachers in the experimental group asked questions to encourage thinking, and thus their students produced more thoughtful answers. The teachers in the experimental group were concerned with what they could learn from the students' answers and how they could provide feedback to help the students adjust their learning pathways. To this end, the teachers in the experimental group provided constructive feedback to enhance the students' confidence, optimism and determination. Such feedback specified the learning outcomes that the students had or had not achieved and the learning pathways they needed to follow. The quality of interactive feedback is a critical feature in determining the quality of the learning activity (Black & Wiliam, 2006).

Fourth, by engaging the students in self- and peer assessment of their work, the teachers motivated them to improve the standard of their work. Peer assessment provides opportunities for students to serve as instructional resources for one another (Black & Wiliam, 2006). The students in the experimental group were receptive to the comments made by their peers, probably because the comments were in a language they could relate to. By building on their peers' comments, the students in the experimental group were able to adjust their learning beyond what they would have done if they had not engaged in self- and peer assessment. Consequently, the students in the experimental group seemed to believe more strongly in their own learning success (Black & Wiliam, 2009).

Lastly, by sharing assessment criteria, the teachers made the students aware of the achievement benchmarks from the start of the topic. Therefore, both teachers and students could monitor learning progress based on the shared assessment benchmarks and lesson objectives. They planned the next learning pathways and managed their learning advancement. When learners participate in setting success criteria, they are able to monitor their thinking, performance and understanding (Davies, 2003). In other words, they use the assessment criteria to monitor their learning.

Conclusion

The present study set out to redesign biology teachers' lessons to integrate AfL techniques into the teaching and learning process. It further assessed the effect of integrating AfL techniques on students' learning achievement in form one biology. Independent sample t-tests revealed that the form one students in the experimental group exhibited higher performance than those in the control group on a test measuring understanding of the topic Cell Structure and Organisation. This suggests that the students in the experimental group, which was taught using AfL-integrated lessons, achieved higher learning compared to those in the control group, which was taught using conventional approaches. Overall, these results strengthen the idea that the integration of AfL into teaching and learning enhances students' learning achievement (Ellegaard et al., 2018; Wiliam et al., 2004). The present study contributes to our understanding of how teachers in resource-constrained classrooms such as those in sub-Saharan Africa can integrate AfL techniques into their day-to-day lessons, thereby harnessing the power of assessment to enhance learning and raise standards.

The findings of the study are, however, limited in some important ways. The most important limitation lies in the fact that the sample was small, which limits the generalisability of the findings. Furthermore, the training itself may have motivated the teachers in the experimental group to provide better and novel instruction. Consequently, the students in the experimental group may have benefited from such novelty (Mertens, 2010). Lastly, although efforts were made to match the control group and the experimental group along several key variables, including age, learning environment, teacher demographic, learning achievement, etc., the two groups were not equivalent. This is because the groups were not randomly assigned (Mitchell & Jolley, 2010). If interpreted cautiously, however, the findings may still prove useful in supporting the conclusions.

Future research could assess how students benefit from different AfL techniques and whether each of the techniques contributes equally to students' learning achievements. For example, research comparing the contribution of constructive feedback with the contribution of peer-assessment is needed. Furthermore, in resource-constrained classroom contexts, a follow-up study assessing the sustained use of AfL techniques by teachers in the intervention group is imperative. Such a study is needed in order to establish whether or not teachers continue to use AfL techniques after participating in continuous professional development aimed at improving their assessment practices.

References

Berry, R. (2008). *Assessment for learning*. Hong Kong University Press.

Black, P., & Wiliam, D. (2005). *Inside the black box: Raising standards through classroom assessment*. Granada Learning.

Black, P., & Wiliam, D. (2006). Developing a theory of formative assessment. In J. Gardner (Ed.), *Assessment and learning* (pp. 81–100). Sage.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, *21*(1), 5–31. https://doi.org/10.1007/s11092-008-9068-5

Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, *25*(6), 551–575. https://doi.org/10.1080/0969594X.2018.1441807

Broadfoot, P. M., Daugherty, R., Gardner, J., Harlen, W., James, M., & Stobart, G. (2002). *Assessment for learning: 10 principles.* University of Cambridge, School of Education.

Davies, A. (2003). Learning through assessment: Assessment for learning in science classroom. In J. M. Atkin & J. E. Coffey (Eds.), *Science educators essay collection: Everyday assessment in science classroom* (pp. 13–26). NSTA Press.

Ellegaard, M., Damsgaard, L., Bruun, J., & Johannsen, B. F. (2018). Patterns in the form of formative feedback and student response. *Assessment & Evaluation in Higher Education*, *43*(5), 727–744. https://doi.org/10.1080/02602938.2017.1403564

Greenstein, L. (2010). *What teachers really need to know about formative assessment*. ASCD.

Hakielimu. (2012). *School children and national examinations: Who fails who?: A research report on the relationship between examination practice and curriculum objectives in Tanzania.* Hakielimu. https://learningportal.iiep.unesco.org/en/library/school-children-and-national-examinations-who-fails-who-a-research-report-on-the

Hardman, F., Abd-Kadir, J., & Tibuhinda, A. (2012). Reforming teacher education in Tanzania. *International Journal of Educational Development*, *32*(6), 826–834. https://doi.org/10.1016/j.ijedudev.2012.01.002

Hopfenbeck, T. N. (2018). Classroom assessment, pedagogy and learning – twenty years after Black and Wiliam 1998. *Assessment in Education: Principles, Policy & Practice*, *25*(6), 545–550. https://doi.org/10.1080/0969594X.2018.1553695

Hopfenbeck, T. N. (2017). Balancing the challenges of high-stakes testing, accountability and students' well-being. *Assessment in Education: Principles, Policy & Practice, 24*(1), 1–3.

Hopfenbeck, T. N., & Stobart, G. (2015). Large-scale implementation of assessment for learning. *Assessment in Education: Principles, Policy & Practice*, *22*(1), 1–2. https://doi.org/10.1080/0969594X.2014.1001566

Jonsson, A., Lundahl, C., & Holmgren, A. (2015). Evaluating a large-scale implementation of Assessment for Learning in Sweden. *Assessment in Education: Principles, Policy & Practice*, *22*(1), 104–121. https://doi.org/10.1080/0969594X.2014.970612

Kippers, W. B., Wolterinck, C. H. D., Schildkamp, K., Poortman, C. L., & Visscher, A. J. (2018). Teachers' views on the use of assessment for learning and data-based decision making in classroom practice. *Teaching and Teacher Education*, *75*, 199–213. https://doi.org/10.1016/j.tate.2018.06.015

Kira, E., Komba, S., Kafanabo, E., & Tilya, F. (2013). Teachers' questioning techniques in advanced level chemistry lessons: A Tanzanian perspective. *Australian Journal of Teacher Education*, *38*(12), 66–79.

Kitta, S., & Tilya, F. (2010). The status of learner-centred learning and assessment in Tanzania in the context of the competence-based curriculum. *Papers in Education and Development, 29,* 77–91.

Kyaruzi, F., Strijbos, J.-W., Ufer, S., & Brown, G. T. L. (2018). Teacher AfL perceptions and feedback practices in mathematics education among secondary schools in Tanzania. *Studies in Educational Evaluation*, *59*, 1–9. https://doi.org/10.1016/j.stueduc.2018.01.004

Kyaruzi, F., Strijbos, J.-W., Ufer, S., & Brown, G. T. L. (2019). Students' formative assessment perceptions, feedback use and mathematics performance in secondary schools in Tanzania. *Assessment in Education: Principles, Policy & Practice*, *26*(3), 278–302. https://doi.org/10.1080/0969594X.2019.1593103

Lema, G., & Maro, W. (2018). Secondary school teachers' utilization of feedback in the teaching and learning of mathematics in Tanzania. *Papers in Education and Development*, *36*, 162–184.

Mertens, D. M. (2010). *Research and evaluation in education and psychology*. SAGE Publications.

Ministry of Education and Culture. (2005). *Biology syllabus for secondary schools.* Ministry of Education and Culture.

Ministry of Education and Vocational Training. (2007). *Syllabus for biology subject I-IV.* Ministry of Education and Vocational Training.

Mitchell, M. L., & Jolley, J. M. (2010). *Research design explained*. WADSWORTH.

Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). SAGE.

O-saki, K., & Njabili, A, (2003). *Secondary education sector analysis*. Ministry of Education and Culture and the World Bank.

Ozcan, T., Ozgur, S., Kat, A., & Elgun, S. (2014). Identifying and comparing the degree of difficulties biology subjects by adjusting it is reasons in elementary and secondary education. *Procedia - Social and Behavioral Sciences*, *116*, 113–122. https://doi.org/10.1016/j.sbspro.2014.01.177

Steiner-Khamsi, G. (Ed.). (2004). *The global politics of educational borrowing and lending*. Teachers College, Columbia University.

UNICEF Tanzania (Ed.). (2012). *Cities and children: The challenge of urbanisation in Tanzania*. UNICEF Tanzania.

Vavrus, F. (2009). The cultural politics of constructivist pedagogies: Teacher education reform in the United Republic of Tanzania. *International Journal of Educational Development*, *29*(3), 303–311. https://doi.org/10.1016/j.ijedudev.2008.05.002

Vavrus, F., & Bartlett, L. (2012). Comparative pedagogies and epistemological diversity: Social and materials contexts of teaching in Tanzania. *Comparative Education Review*, *56*(4), 634–658. https://doi.org/10.1086/667395

Webb, M., & Jones, J. (2009). Exploring tensions in developing assessment for learning. *Assessment in Education: Principles, Policy & Practice*, *16*(2), 165–184. https://doi.org/10.1080/09695940903075925

Wiliam, D. (2004, June). Keeping learning on track: Integrating assessment with instruction [Invited address]. 30th Annual Conference of the International Association for Educational Assessment, Philadelphia, PA. https://www.dylanwiliam.org/Dylan_Wiliams_website/Papers_files/IAEA%2004%20paper.pdf

Wiliam, D., Lee, C., Harrison, C., & Black, P. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education: Principles, Policy & Practice*, *11*(1), 49–65. https://doi.org/10.1080/0969594042000208994

Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, *37*(1), 3–14. https://doi.org/10.1016/j.stueduc.2011.03.001

World Bank. (2008). *Curricula, examinations and assessment in secondary education in sub-Saharan Africa* (World Bank Paper no. 128). World Bank.

Wylie, E. C., & Lyon, C. J. (2015). The fidelity of formative assessment implementation: Issues of breadth and quality. *Assessment in Education: Principles, Policy & Practice*, *22*(1), 140–160. https://doi.org/10.1080/0969594X.2014.990416

Biographical note

Albert Tarmo is a lecturer in the field of science education at the School of Education, University of Dar es Salaam, Tanzania. His main areas of research include educational assessment, biology education, science teacher education, inquiry-based science teaching, and learner-centred pedagogy.

Appendix I: Lesson Plan for Control Class

Preliminary Information

Subject	Date	Stream	Period	Time	Number of students	
					Registered	Present

Main Topic: Cell Structure and Organisation

Sub Topic: The concept of a cell

General Objective: Students should understand the concept of a cell

Specific objectives: By the end of 80 minutes, each student should be able to:

- explain the meaning of a cell correctly;
- mention at least four characteristics of a cell correctly;
- differentiate various types of cells.

Resources: Charts showing various types of cells, biology text book.

Lesson Development

Stage	Teaching Activities	Learning Activities	Assessment Procedures
Introduction (5 min.)	Introducing a new lesson. Asking questions about the meaning of cell. • What is a cell?	Listening. Answering questions.	Verbal questions.
Presentation (50 min.)	Describing the concept of a cell. Guiding students to observe charts showing different types of cells. Asking students to identify and write down the characteristics of various types of cells. Writing notes on the chalkboard.	Listening. Observing charts showing various types of cells. Identifying and writing down the characteristics of various types of cells. Taking notes in exercise books.	Verbal questions.
Reinforcement (10 min.)	Provide a reading activity for students to differentiate various types of cells.	Reading biology textbooks in groups to differentiate various types of cells.	
Reflection (10 min.)	Guiding students to discuss the use of the knowledge learned in their daily life.	Discussing the use of the new knowledge in their daily life.	Verbal questions.
Conclusion (5 min.)	Guiding students to summarise the lesson learned.	Summarising the lesson learned.	

Appendix II: Lesson Plan for Intervention Class

Preliminary Information

Subject	Date	Stream	Period	Time	Number of students	
					Registered	Present

Main Topic: Cell Structure and Organisation

Sub Topic: The concept of a cell

General Objective: Students should understand the concept of a cell

Specific objectives: By the end of 80 minutes, each student should be able to:
- explain the meaning of a cell correctly;
- mention at least four characteristics of a cell correctly;
- differentiate various types of cells.

Resources: Charts showing various types of cells, biology text book.

Lesson Development

Stage	Teaching activities	Learning Activities	Assessment Procedures	Feedback
Introduction (15 min.)	Asking open-ended related questions on the concept of a cell: • What are living things like plants made of? • What are these parts, e.g., a leaf, made of? • If you tear up a leaf blade into the smallest units/parts, what will you end up with? • If you divide a living organism into the smallest units, what will you end up with? Guide students to formulate objectives and activities to pursue based on the last two questions.	Brainstorming and answering questions asked in order to elicit their prior thoughts. Suggest objectives and activities to pursue.	Verbal questions (3+ minutes wait time). Verbal responses.	Constructive verbal feedback; teacher scaffolds, follow-up questions.
Presentation (30 min.)	Leading students in groups of five to observe charts showing various types of cells. Provide rubrics to guide peer assessment. Commenting on the group work and peer comments. Clarifying any misconceptions and queries arising from group activity.	Observing charts showing various types of cells. Identifying characteristics of various cells. Writing down meaning and characteristics of cells. Exchanging work between groups. Commenting on peer work based on the rubric provided.	Assessment by peers on the exchanged work focusing on weaknesses, strengths and points for further improvement.	Written comments by peers. Verbal comments by teacher.

Stage	Teaching activities	Learning Activities	Assessment Procedures	Feedback
Reinforcement (15 min.)	Provide reading activity for students to differentiate various types of cells. Providing worksheets.	Reading biology textbooks in groups of five to differentiate various types of cells. Attempt questions on worksheets. Exchange worksheets between groups for peer comments.	Assessment of worksheets by peers identifying weaknesses, strengths and points for further improvement.	Written comments by peers. Verbal comments by teacher.
Reflection (10 min.)	Leading plenary discussion on the application of knowledge of cells in daily life, e.g., in sickle cell screening.	Discuss in plenary the application of knowledge of cells in daily life	Verbal questions (3+ minutes wait time).	Verbal comments by teacher on individual responses.
Conclusion (10 min.)	Guiding students to revisit the objectives set and summarise the major concepts learned.	Revisiting objectives set. Summarising major concepts learned.	Verbal questions. Self-assessment to determine what they have learnt.	

Appendix III

A. Rubric on observing charts with various cell types

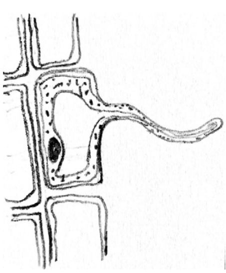
Task	Good	Better	Best
Defining cell	*Defined a cell fairly correctly but missed the most relevant keywords, thus meaning conveyed is vague.*	*Defined a cell correctly using some relevant keywords, thus the meaning conveyed is fairly clear.*	*Defined a cell correctly using relevant keywords, thus conveying a clear meaning.*
Identifying the characteristics of cells	*Identified at least four characteristics without explaining or giving examples.*	*Identified at least four characteristics giving a fairly clear explanation without examples.*	*Identified at least four characteristics giving a clear explanation and examples.*

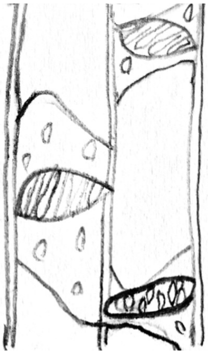
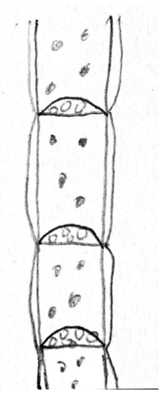
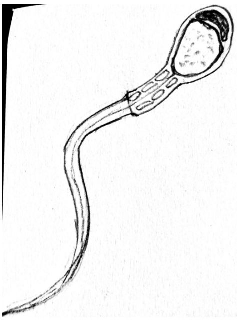
B. Worksheet

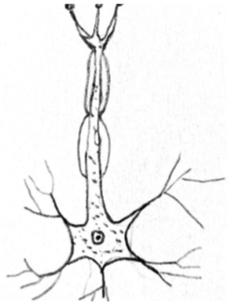
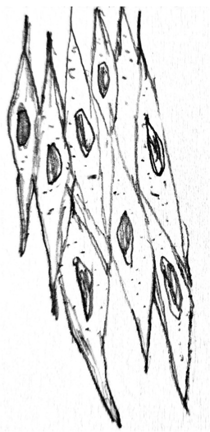
Instructions:

- Carefully read question I–III and fill in the blanks as required.
- Once you complete your work, give it to your neighbouring group, who will also give you their work.
- Once you receive the work of your peers, read it carefully and provide constructive comments highlighting both the strengths and weaknesses, and suggesting what they might do to further improve their work.

Question I

Cell type	Description
A. Root hair 	1. Name ________________________________ 2. Location ________________________________ 3. Functions ________________________________ ________________________________ ________________________________ ________________________________ 4. Adaptive features ________________________________ ________________________________ ________________________________ ________________________________

B. Xylem 	1. Name____ 2. Location ____ 3. Functions ____ 4. Adaptive features ____
C. Phloem 	1. Name____ 2. Location ____ 3. Functions ____ 4. Adaptive features ____
D. Sperm cell 	1. Name____ 2. Location ____ 3. Functions ____ 4. Adaptive features ____

<table>
<tr><td>E. Nerve cell
</td><td>1. Name____________________
2. Location ____________________
3. Functions ____________________
____
____
____
4. Adaptive features ____________________
____
____
____________________</td></tr>
<tr><td>F. Muscle cell
</td><td>1. Name____________________
2. Location ____________________
3. Functions ____________________
____
____
____
4. Adaptive features ____________________
____
____
____________________</td></tr>
</table>

Question II

Cells "A" are found in the specialised region of __________ called_______. The cells in this region produce a ___________________that helps them burrow into the soil more easily.

Question III

Special types of proteins called ___________and _________________ make up cells F. Due to the presence of such proteins, ______________ can slide past each other.

doi: 10.26529/cepsj.1025

Mind the Gap: Age-Related Differences in Students' Perceptions of English Foreign Language Teacher and Motivation

Morana Drakulić[1]

The present paper addresses the age issue in the context of learning English as a foreign language in instructional settings. Our attention has been directed towards the examination of potential differences in students' perceptions of their foreign language teacher and motivation in relation to age. A total of 592 participants attending higher grades of elementary school participated in the research. The results have shown that students' perception of English language teacher characteristics and competences varies in relation to age. Although elementary school students perceive their language teacher to be the most competent in the area of instructional competences, younger students seem to put more emphasis on teacher's personal characteristics over professional competences. The study also indicates differences in motivation, with an accentuated decline in relation to students' age. The study offers a valuable information for teachers and policy makers and emphasizes the need for further adjustment of teaching methodology to various age groups.

Keywords: age, foreign language teacher, motivation, students' perceptions

1 Faculty of Teacher Education, University of Rijeka, Croatia; morana.drakulic@ufri.uniri.hr.

Upoštevanje vrzeli: s starostjo povezane razlike v učenčevih dojemanjih učiteljev angleščine in motivacije

Morana Drakulić

V raziskavi smo preučevali vprašanje starosti v kontekstu učenja angleščine v učnih okoljih. Pozornost smo posvetili preučevanju mogočih razlik v dojemanju njihovih učiteljev tujega jezika in motivacije v povezavi s starostjo. Skupno je v raziskavi sodelovalo 592 učencev višjih razredov osnovne šole. Izsledki so pokazali, da učenci različno dojemajo značilnosti učiteljev angleščine in njihove kompetence glede na starost. Čeprav osnovnošolci dojemajo svoje učitelje angleščine kot najsposobnejše na področju kompetenc poučevanja, pa se je izkazalo, da mlajši učenci dajejo več poudarka osebnostnim značilnostim kot profesionalnim kompetencam. Raziskava kaže tudi na razlike v motivaciji s poudarjenim upadanjem glede na učenčevo starost. Raziskava daje pomembne vpoglede za učitelje in odločevalce ter poudarja potrebo po nadaljnjem prilagajanju metod poučevanja različnim starostnim skupinam.

Ključne besede: starost, učitelj tujega jezika, motivacija, dojemanje učencev

Introduction

Learning English as a foreign language (EFL) at an early age has become a more prominent topic in the last two decades or so as both policy makers and field practitioners have begun to realise the potential benefits an early start may bring. This raising awareness is partly the consequence of national politics that advocate the necessity of communication in at least one FL (European Commission, 2006) and the assumption that younger children learn better as well as easier and that the extended period of learning a language leads to higher levels of proficiency and attainment.

The impact of the global pressures to learn English from an earlier age gave rise to uncertainties regarding effective programme design for this age group (Enever & Moon, 2009, p. 9). This has, as a consequence, again initiated extended research on the age-related issues in the context of early FL learning in formal educational settings.

The age factor in L2 attainment has long been a subject of a debate whether differences in the FL achievement may be explained in terms of maturational constraints (Critical Period) or in terms of the effect of a wider set of factors. In regard to the latter, a substantial amount of research has turned their focus to the examination of a complex interplay between age and a plethora of variables that may affect the quality of acquisition. However, as Cenoz (2003) points out, much of the studies on, what they call, age issue, have been conducted in second language situations where children are exposed to a FL both in and outside of the classroom. Such a surrounding is very different from a FL setting and calls for s new research perspective which should focus on those factors which are relevant to the classroom teaching and learning context. In this connection, one line of the research recognized the importance of investigating contextual factors as well as learner-related variables. Within this research paradigm, an emphasis is put on a more detailed examination of students' subjective perceptions of the language learning environment especially in relation to younger students whose subjective interpretations and views on the classroom activities as well as on the FL teacher work and behaviour may strongly affect the level of their FL learning motivation.

Motivated by these arguments, the aim of the present study is threefold. First, we wanted to determine effective FL teacher competences as perceived by elementary school students. Guided by the premise that students' subjective perceptions serve as an affective base for the development of attitudes towards the learning situation (Gardner, 2010), we also wanted to explore the nature of the relationship between students' subjective perceptions and language learning motivation. Lastly, the study is to shine a new light on age-related debates through the examination of potential differences in students' perceptions and, consequently, their level of motivation in relation to their age.

Issues related to optimal age for learning a FL

The issue of the ideal or optimal age for successful language acquisition has been initiated approximately sixty years with the Critical Period Hypothesis (CPH), which has since been a topic of controversy and critic. Differences in the interpretation of its main fundamental postulates, such as relevant age span (Vanhove, 2013) and varying research results that argue for and against CPH, led to reasoned as to whether age-related differences in attainment may be explained in terms of a Critical Period. In this context, much attention has been drawn to the notion of ultimate attainment, which has been set on a level of native-like proficiency. Despite empirical evidence that supports the assumption that exposure to a FL from an early age facilitates the level of proficiency (Butler, 2015; Chen et al., 2020; Domínguez & Pessoa, 2005), the question of whether reliance on this particular comparison is really the best way of exploring age effects and maturational issues has been raised (Muñoz & Singleton, 2011). Arguments made by those taking this perspective seem to suggest that a native speaker model in the FL teaching context is unfounded, since there is no reason why a FL learner's ultimate attainment should be the same as that of a native speaker (Muñoz & Singleton, 2011). Furthermore, research focused on late beginners attaining native-like proficiency suggested that the differences in FL acquisition between younger and older learners may be attributed to a large set of factors, and not only to biological maturation constraints as CPH implies (Johnstone, 2002). Moreover, as Mihaljević Djigunović (2014) points out, this new approach has come to a realization that age is difficult to disentangle from other variables and its impact on SLA can be better understood if we take into account its interactions with other factors (Mihaljević Djigunović, 2014, p. 420).

With the aim of focusing their work on promoting new research paradigms in language education, recent empirical studies have involved a plethora of relevant variables that may facilitate and debilitate the process of FL learning. Thus, the attention has been drawn to factors such as attitudes and motivation (Gardner, 2010; Dörnyei & Muir, 2019; Dörnyei, 2020; Muñoz & Singleton, 2011), quality and the amount of home support (Enever, 2011; Muñoz & Lindgren, 2011) the role of the teacher (Borg, 2006; Dewaele et al., 2019); Drakulić, 2018; Sert, 2019; Tragant & Lundberg, 2011), the learners self-concept

(Mihaljević Djigunović & Lopriore, 2011), willingness to communicate (Dewaele & Dewaele, 2018; Dewaele, 2019) etc. In recent overview chapters of the age effects, it is possible to discern a growing interest in the investigation of the learners' subjective perceptions of the language learning situation (Ahn & West, 2016; Enever, 2011; Jakominić & Mihaljević Djigunović, 2004; Mihaljević Djigunović, 1998; Mihaljević Djigunović, 2014; Nikolov, 1999; Nikolov, 2002; Yuksel & Halici, 2010), i.e. the relationship between the learner's perception, age and the ultimate attainment. In this line of work, further evidence of age-related difference has been reported with regard to students' perceptions of their FL teacher work and behaviour as one aspect of attitudes towards the learning situation. However, this particular issue has often been investigated alongside the wider set of contextual variables and has, unfortunately, rarely been formulated as the main research aim of the empirical research conducted so far.

Motivation and attitudes in FL learning

Motivation refers to a factor that provides an initial impetus, which later turns into a drive that pushes learners throughout the whole learning process. Attitudes refer to the learner's feelings about the FL community and the learning situation and, as such, represent the affective base for the motivation to develop.

Research on FL learning motivation has been heavily influenced by Gardner's motivational theory which has been built around two key principles: integrative orientation (i.e. identification with L2 community), and instrumental orientation (i.e. practical value of learning a language). Despite its dominance in the social-psychological period, the integrative-instrumental dichotomy was argued to be insufficient to describe motivation on other language learning contexts, such as a classroom context, where learners do not have a direct contact with a target language community (Dörnyei, 2009). Therefore, Dörnyei proposed the L2 Motivational Self-System, consisting of three components: (1) the ideal L2 self (i.e. learner's imagined ideal future self as an L2 speaker); (2) the ought-to L2-self (i.e. learner's beliefs on the attributes they should possess to meet expectations or avoid possible negative outcomes); and (3) L2 learning experience (i.e. motivating factors of the learning situation such as the teacher, the curriculum, the peer group etc.).The component addressing the language context acknowledges the fact that the two-self guides do not impact the learning process in an isolated way (Csizér, 2019, p. 73).

More recently, the research has also pointed out to the dynamic nature of motivation and its sensitivity to temporal dimension which, in consequence, gave rise to longitudinal approach to the investigation of this multifaceted phenomenon (Mihaljević Djigunović, 2012). In the context of early FL learning, Mihaljević Djigunović and Nikolov (2019) proposed the framework of young learners' motivation which illustrates the complex interplay between motivation and valued others (i.e. family members, the teacher, and peers) along three stages. At stage one (6-8 years) the teacher holds the central motivating role and is the primary source of learners' motivation. The impact of the family in this stage is quite significant as well. At the second stage (9-11 years) the teacher and the family still have an effect on the level of motivation, whereas peers' role gains more importance. Finally, in the third stage (12-14 years) the impact of the family and the teacher declines, giving rise to peers who take over the function of main motivator and a role model.

Research on young language learners' motivation, attitudes and age

Research on the relationship between age and other factors relevant for the instructed language learning settings has coherently pointed to the importance of individual learners' characteristics. Research to date suggests that young learners' motivation is a specific phenomenon due to its sources and dimensions, as well as to the complex and dynamic interaction it enters into with other language learning variables (Mihaljević Djigunović & Nikolov, 2019). However, Nikolov and Mihaljević Djigunović (2006) warn how these claims should be interpreted with caution, since superficial considerations may lead to already widely accepted assumption that early FL instruction will, as a rule, contribute to children's favourable attitudes (Nikolov & Mihaljević Djigunović, 2006, p. 246). Positive language learning experiences, in other words, do not develop *per se*, or just because of the fact that FL instruction has been introduced at an early age. This shift in the perspective has consequently resulted in research based on a more situated approach which showed that classroom and learning processes are the primary sources of formation and variation of young learners' attitudes and motivation (Mihaljević Djigunović & Nikolov, 2019).

Concerning the relation between motivation and attitudes at the one hand, and immediate learning environment on the other, there seems to be an agreement that the teacher has an important, multidimensional motivational role. Insights from classroom studies suggest that teachers' choice of activities as well as their personality may raise and maintain students' motivation but may also cause demotivation (Mihaljević Djigunović & Letica Krevelj, 2010).

One good illustration of the variety of sources of motivation in the EFL classroom, and their variability in relation to age, is the study conducted by Nikolov (1999) who examined possible age effects on primary school students' attitudes and motivation. The study showed that the most important motivating factors for elementary school learners are attitudes towards the teacher and intrinsically motivated materials, tasks and activities (Nikolov, 1999, p. 53). This was especially evident among the youngest (6-8 years old) participants, while intrinsic motivation was gradually replaced by instrumental-knowledge motives for learning around puberty. Similar results were found by Fenyvesi (2020) who investigated young learners' attitudinal and motivational changes between two age groups across one year. The results revealed that children's dependence on significant others decreased significantly over time. The study also confirmed more positive attitudes among the early starters at the end of the first year, but revealed that after the second year of instruction their attitude levels dropped to the approximately same level as the late starters' attitudes after the first year (Fenyvesi, 2020, p. 708).

Perhaps one of the most extensive studies exploring the variety of both contextual and learner-related factors within the context of language learning at primary level is the trans-national, longitudinal project Early Language Learning in Europe (ELLiE). The principal aim was to investigate children's perceptions on their language learning environment, while simultaneously looking at their attitude and motivational changes throughout the four-year period. The project reports on highly complex interactions of learner characteristics (attitudes, motivation, and self-concept) with the other factors, contextual ones in particular. The analysis of children's responses shows that young learners display favourable attitudes and high level of motivation for FL learning which emerge from age-appropriate, game-like activities that, in turn, provide a feeling of achievement and language confidence. When it comes to language achievement, the results indicate a strong interdependence between this variable and motivation, attitudes and self-concept. Finally, the findings also suggest that language achievement becomes increasingly more associated with the learners' individual characteristics as the children mature (Mihaljević Djigunović & Lopriore, 2011).

Mihaljević Djigunović (2007) investigated the relationship between the learners' affective characteristics in predicting achievement in English as a FL among primary and secondary school students. The comparison of the two age groups revealed certain differences in terms that older students have a more positive affective profile and that this profile serves as the best predictor of achievement. Somewhat similar results were also confirmed in a project conducted by the research group at the University of Basque County. One of the aims was to investigate the effect of the introduction of EFL at different ages on the rate of achievement, the development of attitudes and motivation between the groups of learners who have had the same amount of exposure to language (Cenoz, 2003, p.81). The results indicate that young learners obtain significantly lower results in all of the measures of English proficiency except for the mechanics in writing, where no significant differences were found. On the other hand, younger learners 'outperformed' their older colleagues in the affective domain since they displayed significantly more positive attitudes and higher levels of motivation.

Data analysis from these and other extensive studies point to several important issues as well as drawbacks found in the current theoretical and empirical research. It suggests that FL learners' attitudes and motivation are highly dependent on their language learning experience, that is, on their subjective perceptions of the language learning environment. Acknowledging the fact that teachers represent a cornerstone and the prerequisite for the development of the quality of the language learning experience, one would expect a plethora of studies devoted to this specific issue. However, the research on how the (perceived) quality of the teacher contributes to the language learning in general is still rather scarce, except for few noticeable exceptions (Beaudrie et al., 2004; Brosh, 1996; Brown, 2009; Kadha, 2009; Shishavan & Sadeghi, 2009; Zamani & Ahangari, 2016). For this reason, Nikolov & Mihaljević Djigunović (2006) accentuate the necessity for exploring classroom practices over time and the triangulation of data collected from learners, teachers and observers.

Method

Research aim

The study aimed at getting an insight into the potential differences in students' perceptions of teacher's professional competences and intrapersonal and interpersonal characteristics and skills and the way the potential differences in perceptions affect students' level of motivation in relation to age. We expect that there will be a statistical relevant difference in the students' perceptions of their language teacher as well as in their level of motivation in relation to age.

Participants

The study was carried out in the Croatian socio-educational context and involved a total of 592 participants out of which 302 (51%) were male and 290 (49%) were female. Participants attended higher grades of elementary school (age: 10/14 years) i.e. 139 students (23.5%) attended fifth grade, 152 students (25.7%) attended sixth grade, 162 students (27.3%) attended seventh grade and, finally, 139 (23.5%) students attended eighth grade of elementary school. All students follow the same curriculum, that is, they learn English as the FL from the very beginning of their compulsory education.

Instruments

The participants' perceptions of their FLT work and behaviour were measured by the questionnaire designed for the purpose of this study. The questionnaire examined students' views of their FL teacher professional competences and intrapersonal and interpersonal characteristics and skills, to what extent the competences and characteristics were present in students' real or actual language teacher. The final version of the instrument, after the pilot study was conducted, consists of 27 positively worded items distributed across three components assessing FL teacher work and behaviour: Intrapersonal and interpersonal characteristics and skills (12 items), Intercultural competence (6 items), and Competences related to FL classroom instruction (9 items). Items are rated on a five-point Likert scale ranging from (1)'*Not important at all* to (5) *Very important*) (see Appendix). Prior to the construction, we examined some of the already constructed and available instruments (Barnes & Lock, 2010; Beaudrie et al., 2004; Bell, 2005; Carvalho, 1999; Sakurai, 2012) and took into consideration those items which best suited our instrument purpose. However, taking into account the importance of the specificity of the socio-educational context (Gardner, 2010), the primary source which we consulted was the competence description from the Competences of Primary School FL Teachers in the Republic of Croatia (Radišić et al., 2007).

Information about learners' motivation for learning EFL was gathered through Motivation questionnaire which is comprised of 30 items and assessed on a Likert-type scale ranging from strong disagreement (1) to strong agreement (7). The instrument is the adapted version of Gardner's international version of Attitude Motivation Test Battery (Gardner, 2010). It consists of three subscales assessing affect-based constructs of motivation: motivational intensity, desire to learn English, and the attitudes towards the learning situation. Since the instrument is originally written in English, the order of the items was taken from the original whereas the translation to Croatian was provided by Gardner and Mihaljević Djigunović (2003).

Data analysis

We used the following statistical procedures: descriptive statistics, correlational analysis, and one-way analysis of variance. Descriptive statistics (mean and standard deviation values) was used with the aim to inspect the distribution of results within each component and construct. On the basis of descriptive statistics it was also determined which aspects of FL teacher were perceived as more or less favorable. The level of students' motivation was also observed and explored at the descriptive level. Correlational analysis (Pearson correlation coefficient) was carried out with the aim to determine the degree and the type of relationship between all the aforementioned variables. As minimal correlation coefficient the value of 0,30 was used. The analysis of variance was used to investigate whether there are statistical significant differences between mean values of variables included in this research.

Results

The distribution of the statistical data was first investigated through the variables' mean and standard deviation values for each instrument separately. Potentially statistically significant differences in relation to age were then investigated through the analysis of variance.

Table 1

Composite scores for components assessing ELT competences and characteristics

Component		Mean	St. deviation
Intrapersonal and interpersonal characteristics and skills	Sum score	44.10	9.58
	Scale average	3.67	.79
Intercultural competence	Sum score	20.10	6.12
	Scale average	3.35	1.02
Competences related to FL classroom instruction	Sum score	33.80	8.45
	Scale average	3.75	.93

The analysis of the questionnaire assessing students' perceptions of their FL teacher (Table 1) indicates that the component assessing language teacher's instructional competence has the highest sample mean score (x = 3.75, SD = .93), followed by the component assessing the teacher's personal characteristics (x = 3.67, SD = .79), and by the component assessing the teacher's intercultural competence (x = 3.35, SD = 1.02).

In order to determine if there is any statistical difference in students' perceptions in relation to their age, we used the analysis of variance (Table 2). The analysis confirmed statistically significant differences between the means of all groups under the analysis (p < .01). Regarding the teacher's intrapersonal and interpersonal characteristics and skills, a statistically significant difference of perceptions was found between the sixth (x = 44.88, SD = 9.34), and seventh (x = 42.79, SD = 7.94) graders, and between the sixth and eighth (x = 42.00, SD = 10.41) graders. These age groups differ in relation to their perceptions of their actual teacher intercultural competences since significant differences were found between sixth (x = 22.22, SD = 4.80) and seventh (x = 19.14, SD = 6.61) graders and between sixth and eighth (x = 18.61, SD = 5.53) graders as well. Differences in perceptions were the most evident in relation to students' assessment of competences related to FL classroom instruction. Here, statistical significant difference was found between fifth and seventh graders; fifth and eighth graders; sixth and seventh graders and, finally, between sixth and eighth graders.

Table 2

Analysis of variance for ELT perceptions in relation to age

Component	Grade	Mean	SD	df	F	sig.
PER	5th	44.88	10.04	3.58	7.55	.00
	6th	46.72	9.34			
	7th	42.79	7.94			
	8th	42.00	10.41			
CUL	5th	20.43	6.72	3.58	10.74	.00
	6th	22.22	4.80			
	7th	19.14	6.61			
	8th	18.61	5.53			
INSTR	5th	36.07	7.92	3.58	16.95	.00
	6th	36.14	7.05			
	7th	32.50	8.17			
	8th	30.46	9.31			

Note: PER – personal characteristics and skills; CUL – intercultural competence; INSTR – competences related to FL classroom instruction

The results obtained from the preliminary analysis of *Motivation* questionnaire reveal that the motivation for learning English is present within our sample (Table 3).

Table 3

Mean and standard deviation values for items assessing motivation

	Items	Mean	SD
1	I don't pay much attention to the feedback I receive in my English class.	2.65	1.74
2	Learning English is really great.	5.02	1.91
3	I have a strong desire to know all aspects of English.	5.63	1.69
4	I make a point of trying to understand all the English I see and hear.	5.74	1.44
5	Knowing English isn't really an important goal in my life.	1.87	1.64
6	I hate English.	1.95	1.66
7	I don't bother checking my assignments when I get them back from my English teacher.	2.43	1.83
8	I really enjoy learning English.	4.72	1.93
9	If it were up to me, I would spend all of my time learning English.	3.27	1.99
10	I keep up to date with English by working on it almost every day.	3.70	1.86
11	I sometimes daydream about dropping English.	2.79	2.19
12	I would rather spend my time on subjects other than English.	3.03	1.97
13	I put off my English homework as much as possible.	2.37	1.96
14	English is a very important part of the school programme.	6.32	1.21
15	I want to learn English so well that it will become natural to me.	5.94	1.54
16	When I have a problem understanding something in my English class, I always my teacher for help.	4.56	2.03
17	I'm losing any desire I ever had to know English.	2.16	1.74
18	Learning English is a waste of time.	1.66	1.40
19	I tend to give up and not pay attention when I don't understand my English teacher's explanation of something.	2.36	1.66
20	I plan to learn as much English as possible.	5.97	1.48
21	I would like to learn as much English as possible.	6.11	1.41
22	I really work hard to learn English.	5.50	1.65
23	To be honest, I really have no desire to learn English.	1.62	1.35
24	I think that learning English is dull.	2.34	1.79
25	I can't be bothered trying to understand the more complex aspects of English.	2.17	1.62
26	I love learning English.	4.84	1.96
27	I wish I were fluent in English.	6.45	1.21
28	When I am studying English, I ignore distractions and pay attention to my task.	4.80	1.83
29	I haven't any great wish to learn more than the basics of English.	1.92	1.50
30	When I leave school, I will give up the study of English because I am not interested in it.	1.65	1.36
	Sum score	111.54	50.55
	Scale average	3.71	1.68

The analysis of the learner's responses reflect the presence of motivational intensity, i.e. the effort, persistence, and consistency in learning. Students, in other words, invest an effort to learn and understand English as much as possible (#4, #22,), are willing to concentrate on tasks and to the feedback they receive from their teacher (#1, #28) but are, to a certain extent, reluctant to work on the improvement of English on daily basis (#7, #10). High mean values obtained for the items within a questionnaire implicate the existence of positive attitudes and a desire to learn English as well. Learning English is perceived as enjoyable activity and the possession of the knowledge of English is considered as important competence to acquire both in the present and in the future.

Possible statistically significant differences in the level of motivation in relation to the students' age were explored on the basis of ANOVA results and post hoc analysis.

Table 4
Analysis of variance for motivation in relation to age

Construct	Grade	Mean	SD	df	F	sig.
MOT	5th	172.97	30.21	3.58	7.04	.00
	6th	165.87	32.28			
	7th	166.31	28.02			
	8th	156.60	29.05			

As it is apparent from the Table 4, there was significant difference found in relation to age and level of motivation ($F(3.58) = 7.04$, $p < .05$) between the means of all groups under the analysis. Post hoc tests revealed that eighth graders statistically significantly differ from the fifth, sixth and seventh graders. It appears that the level of motivation decreases with age.

The final set of analysis investigated possible significant differences in relation to gender on the basis of an independent sample t-test.

Table 5 shows that the level of motivation significantly differs in relation to students' gender, that is, that female learners display higher levels of motivation in comparison to male learners.

Table 5
T-test on gender differences in relation to motivation

Construct	Gender	Mean	SD	Test for Equality of Variances		t-test		
				F	p	t	df.	p
MOT	M	160.21	31.68	6.722	.010	4.38	583.08	.00
	F	170.98	27.92					

Contrary to that, no statistically significant differences were found in relation to students' perception and their gender. As can be seen from Table 6, the perceptions of ELT competences and personal characteristics are rather similar in male and female students since p-values are all greater than 0.05.

Table 6

T-test on gender differences in relation to Actual ELT perceptions

Component	Gender	Mean	SD	Test for equality of variances		t-test		
				F	p	t	df.	p
PER	M	43.66	9.86	.68	.40	1.13	587.56	.25
	F	44.55	9.27					
CUL	M	20.27	6.10	.06	.80	.66	589	.50
	F	19.94	6.14					
INSTR	M	34.03	8.18	1.57	0.21	.70	582.90	.48
	F	33.54	8.73					

After the examination and the analysis of the results obtained by each instrument separately, we wanted to determine whether there is a relationship between all variables included in our research. For the estimation of possible correlations between the variables, a Pearson correlation coefficient was used.

Table 7

Correlational analysis between motivation, age, gender and ELT perceptions

Variables	GEN	AGE	MOT	PER	CUL	INSTR
GEN	1					
AGE	.026	1				
MOT	.178**	-.173**	1			
PER	.047	-.147**	.359**	1		
CUL	-.027	-.157**	.236**	.617**	1	
INSTR	-.029	-.266**	.342**	.762**	.667**	1

*p < .05
**p < .01

The results of this investigation also show that there is correlation between all variables assessing good English teacher competences. Thus strong positive and significant correlation was found between PER and CUL (r =

0.61, $p < .01$), PER and INSTR ($r = .76$, $p < .01$) and CUL and INSTR ($r = .66$, $p < .01$). Motivation was found to correlate moderately and positively with students' perceptions of language teacher intrapersonal and interpersonal characteristics ($r = .35$, $p < .01$), with perceptions of teacher's competences related to FL classroom instruction ($r = .34$, $p < .01$), and weakly and positively with gender ($r = .1$, $p < .01$).

The results indicate a weak, negative but significant correlation between age and all variables under the analysis, especially between age and the perceptions of the teacher's instructional competences ($r = -.2$, $p < .01$) and between motivation ($r = -.7$, $p < .01$).

Discussion

This study set out with the aim of assessing the importance of relationship between students' age and other contextually relevant variables in the instructional FL learning setting, namely the students' perceptions of their ELT and motivation.

Generally speaking, students, regardless age, ascribed more importance to the teachers' intrapersonal and interpersonal characteristics and skills than to other assessed aspects of their profession. These findings are in line with the qualitative research conducted by Drakulić (2018) who found the more importance was ascribed to teacher's personal characteristics than to competences related to FL classroom instruction. Learners accentuated characterictis such as calmness, rationality, emphaty, flexibity and sense of humor as attributes that positively affect their motivation. Conversely, overt reaction to criticism, as well as overt criticism towards the learners, were perceived as sources of demotivation (for more details see Drakulić, 2018).

In accordance with our initial assumptions, students perceived their language teacher differently with regard to their age, i.e. grade they were attending. Younger students, those attending 5th and 6th grade, assessed their teacher's personal characteristics more positively than their older peers did. The inspection of mean values indicate that the difference between younger (5th and 6th graders) and older (7th and 8th graders) is significant, with the younger students evaluating teacher's personal characteristics more highly. The same pattern may be observed for students' evaluation of their teacher's intercultural competence as well. The present findings seem to be consistent with that of Mihaljević Djigunović (1998) who found that the positive perceptions of teacher's ability to inspire decrease as students get older. In her study on reasons for learning English, Nikolov (2002) also states that young learners (aged 8-11) report teacher's personal characteristics more frequently and in a more enthusiastic way when compared to the older learners (aged 11-14).

The greatest differences in students' perceptions were found in assessment of competences related to FL classroom instruction. In this connection, teacher was perceived as more competent by younger students and less competent by older ones. These findings are in agreement with those of Mihaljević Djigunović (1998) which showed that younger students tend to assess teacher's competence more highly than the older students do. Same observation was also reported by Nikolov (2002) who found that students tend to give fewer positive statements related to teacher's qualities as they get older.

As can be seen, students' age affects the way students perceive their ELT competences and characteristics. This conclusion is further supported by the results of correlational analysis which indicate a rather weak but significant negative correlation between students' age and all three components assessing language teacher competences and characteristics. While younger students tend to perceive their teacher more positively, it seems that this tendency weakens as they enter the last two grades of their compulsory education. A possible explanation for this result is that younger students perceive emotional link to the teacher as a relevant factor within the learning context and, thus, attach more importance to teacher's personal characteristics than to some other aspect of teacher profession. Another possible explanation may be, as Mihaljević Djigunović (1998) points out, young students' tendency to perceive a person who knows more than they do as a very competent person. The older they get, the more proficient they become in the language which may affect their perception of the teacher's expertize and competence negatively. Finally, the age-related decline in perceptions could be related to the growing difficulties imposed by the more demanding syllabus. We may assume that students who experience difficulties in mastering the content would probably perceive their language teacher as too demanding and strict and the use of his methods and procedures as inappropriate. This may further be enhanced by the critical attitude towards adults and education in general which usually appears around puberty (Pfenninger & Singleton, 2019) and potentially leads to, as Mihaljević Djigunović and Nikolov (2019) argue, demotivation.

One surprising finding was that the 6th graders, not the 5th, view their FL teacher as more competent in all three components. A possible explanation for these results may be related to the fact that 6th grade students know their teacher better and are more familiar with both his personality and instruction methodology than the 5th graders are. After the 6th grade, it seems that a sort of 'shift' in perception occurs, since the positivity in perception weakens all the way to the 8th grade, for which the lowest assessments were found. A possible interpretation for these results is that students attending 6th grade have probably not yet entered the sensitive time of puberty which is generally characterized by (too) critical attitude towards adults and school in general. For these reasons, it may not be surprising that a shift in perception occurs after the 6th grade, after which learners enter the preadolescent stage.

An unanticipated finding of this study was that boys and girls share a rather consistent views on the importance of FL teacher competences and personal characteristics. This is rather surprising since female students have often been reported to exert general sensitivity towards a language learning situation and, as such, tend to emphasize interpersonal characteristics of a teacher (Mihaljević Djigunović, 1998; Shishavan, 2010). This may be explained by the non-existence of teachers' gender-biased attitudes and expectations (Rúa, 2006). Dewaele et al. (2018), on the other hand, draw attention to the teacher's emotional intelligence which has shown to contribute to the effective working in heterogeneous classrooms.

In line with our expectations, the results have also shown differences in the level of motivation in relation to students' age. The comparison of mean values revealed that the 5th graders exert the highest level of motivation contrary to the 8th graders who statistically significantly differ from other age groups. In accordance with the present results, previous studies have also demonstrated age-related differences in motivation. In this connection, Mihaljević Djigunović (1998) found significant differences in the level of motivation across three different age groups which all go in favour of the youngest (elementary) school population. Much of the same results were reported by Williams et al. (2002) and Fenyvesi (2020) who also found a decrease of motivation with age. In their study younger students had a higher integrative orientation, more positive attitude to their teacher as well as perception of their ability, success and invested effort.

Building on these findings, it seems possible to state that younger students generally start FL learning with very positive attitudes and motivation which are later shaped by the language learning environment and the experience. In this connection, one of the possible explanations may be that students, as they gain more learning experience, become more aware of likes and dislikes of the various elements of the learning process (Ahn & West, 2016).Closely connected to this is a strong critical attitude towards adults and learning in general which typically appears around puberty and, as such, may explain a decrease in motivation among older students in our sample. Another possible explanation could be related to growing difficulties with the language learning itself. A more demanding syllabus as well as pressure about achieving good grades as a condition for enrolment in secondary school may result in losing desire to learn altogether (Nikolov & Mihaljević Djigunović, 2019). All in all, it could be conceivably hypothesized that age-related differences in motivation are affected by various aspects of formal learning environment but at the same time they are also related to a more general adolescent dissatisfaction with the whole process of education.

Finally, the analysis conducted on gender differences in relation to motivation confirmed previous findings which often report higher levels of motivation among female learners (Bećirović, 2017; Mihaljević Djigunović, 1998; Heinzmann, 2009). In an attempt to clarify possible reasons behind gender differences, Yashima et al. (2017) argue that gender-biased responses are socially and culturally conditioned. This tendency is, as the authors imply, strongly linked to the Ought-to self, that is, to the learners' experience of the strong external pressure to succeed academically, and internal pressure to measure up to who they wish to be as educated persons (p.705). The aforementioned is closely linked to a folk belief that girls are better at languages than boys, i.e. that (foreign) languages are female subjects, whereas maths and physics are male subjects (Heinzmann, 2009, p. 20). These stereotypic beliefs about gender-biased aptitude may not only give rise to unequal social expectations, but may also impact students' beliefs about language learning and, consequently, their level of motivation (Horwitz, 1999).

Although the presented research is correlational in nature and, as such, does not allow causal conclusions, the interpretation of the potential causality may be interpreted within the context of the socio-educational model of second language acquisition (Gardner & MacIntyre, 1993). In this connection, attitudes towards the learning situation are represented through students' subjective evaluative reaction to the teacher's work and behaviour. Since attitudes serve as an affective base for motivation to develop (Gardner, 2010), it may be assumed that one source of higher levels of motivation are positive perceptions of the language teacher. This notion is further supported by moderate, positive and significant correlations found between motivation and all three constructs assessing language teacher competences and characteristics. In fact, as Hattie (2003) points out, teachers are the greatest source of variance in students' achievement and it is our responsibility to concentrate on enhancing these sources of variance to truly make a difference (Hattie, 2003, p. 1).

Conclusion

Despite its exploratory nature, the findings of the present study suggest that age is a powerful factor which affects students' perceptions of the language learning situation in the interrelation with other individual learner characteristics and contextual factors. Bearing in mind teachers' powerful motivational role, it seems plausible to assume that students' subjective perceptions on FL teacher competences and personal characteristics may be one of the explanatory factors for the variation in language learning motivation. It has been shown that younger students attach more value and importance to intrapersonal and interpersonal characteristics and skills such as empathy, flexibility and a sense of humour than to competences related to FL classroom instruction. Although age is just one of the plethora of factors which may attribute for the differences in students' preferences, we can say, with almost complete certainty, that young children are not as similar to one another as it is often assumed (Nikolov, 2002).

These results are very indicative and offer valuable information and guidelines for teachers and policy makers who need to be aware that only competent teachers who are able to adjust their methodology to various age groups can really contribute to the quality of the teacher-student relationship and the quality of the learning process as a whole.

Due to practical constraints, this study cannot provide a comprehensive review of other potentially relevant variables that may have also affected the observed differences. A deeper insight into the various constituents of language learning motivation will surely contribute to a more comprehensive understanding of age-related differences found in this research. Moreover, additional clarifications, such as a control type of rating, may potentially delineate the reasons behind a negative answering strategy of the older learners. Thus, the future research should be directed towards the incorporation of a wider set of methodological procedures as well as variables relevant for the FL learning in instructional settings.

References

Ahn, S-Y., & West, G.B. (2016). Unmasking young learners' perceptions of good language teacher identity. *English Teaching, 71*(1), 73–95.

Barnes, B. D., & Lock, G. (2010). The attributes of effective lecturers of English as a foreign language as perceived by students in a Korean University. *Australian Journal of Teacher Education, 35*(1), 139–152. https://doi.org/10.14221/ajte.2010v35n1.2

Beaudrie, S., Brown, A., & Thompson, G. (2004). Multiple perspectives on teacher evaluation in the foreign language classroom. *Arizona Working Papers in SLAT, 11*, 57–80.

Bećirović, S. (2017). The relationship between gender, motivation, and achievement in learning English as a foreign language. *European Journal of Contemporary Education, 6*(2), 210–219. https://doi.org/10.13187/ejced.2017.2.210

Bell, T. R. (2005). Behaviors and attitudes of effective foreign language teachers: Results of a questionnaire study. *Foreign Language Annals, 38*(2), 259–270. https://doi.org/10.1111/j.1944-9720.2005.tb02490.x

Borg, S. (2006). The distinctive characteristics of foreign language teachers. *Language Teaching Research, 10*(1), 3–31. https://doi.org/10.1191/1362168806lr182oa

Brosh, H. (1996). Perceived characteristics of the effective language teacher. *Foreign Language Annals, 29*(2), 125–136. https://doi.org/10.1111/j.1944-9720.1996.tb02322.x

Brown, A. V. (2009). Students' and teachers' perceptions of effective foreign language teaching: A comparison of ideals. *The Modern Language Journal, 93*(1), 46–60. https://doi.org/10.1111/j.1540-4781.2009.00827.x

Butler, Y. G. (2015). English language learning education among young learners in East Asia. A review of current research. *Language Teaching, 48*(3), 303–342. https://doi.org/10.1017/S0261444815000105

Carvalho, I. (1999). The effective language teacher. *New Routes, 8*, 12–18. https://sites.google.com/site/estesinversos/Home/curriculo-lattes/publicacoes-academicas/the-effective-language-teacher

Cenoz, J. (2003). The influence of age on the acquisition of English: General proficiency, attitudes and code-mixing. In M. P. Garciá Mayo & M. L. Garciá Lecumberri (Eds.), *Age and the acquisition of English as a foreign language* (pp. 77–93). Multilingual Matters.

Chen, S., Zhao, J., de Ruiter, L., Zhou, J., & Huang, J. (2020). A burden or a boost. The impact of early childhood English learning experience on lower elementary English and Chinese achievement. *International Journal of Bilingual Education and Bilingualism*. https://doi.org/10.1080/13670050.2020.1749230

Csizér, K. (2019). The L2 motivational self system. In M. Lamb, K. Csizér, A. Henry & S. Ryan (Eds.), *The Palgrave handbook of motivation for language learning* (pp. 71–98). Palgrave.

Dewaele, J-M. (2019). The effect of classroom emotions, attitudes toward English, and teacher behavior on willingness to communicate among English foreign language learners. *Journal of Language and Social Psychology, 38*(4), 523–535. https://doi.org/10.1177/0261927X19864996

Dewaele, J-M., & Dewaele, L. (2018). Learner-internal and learner-external predictors of willingness to communicate in the foreign language classroom. *Journal of the European Second Language Association, 2*(1), 24–37. http://doi.org/10.22599/jesla.37

Dewaele, G., & Mercer (2018). Do ESL/EFL teachers´ emotional intelligence, teaching experience, proficiency and gender affect their classroom practice? In M. Agudo & J. de Dios (Eds.), *Emotions in second language teaching* (pp. 125–141). https://doi.org/10.1007/978-3-319-75438-3_8

Dewaele, J-M., Magdalena, A. F., & Saito, K. (2019). The effect of perception of teacher characteristics on Spanish EFL learners' anxiety and enjoyment. *The Modern Language Journal, 103*(2), 412–427. https://doi.org/10.1111/modl.12555

Domínguez, R., & Pessoa, S. (2005). Early versus late start in foreign language education: Documenting achievements. *Foreign Language Annals, 38*(4), 473–483. https://doi.org/10.1111/j.1944-9720.2005.tb02514.x

Dörnyei, Z. (2009). The L2 motivational self-system. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 9–42). Multilingual Matters.

Dörnyei, Z. (2020). *Innovations and challenges in language learning motivation*. Routledge.

Dörnyei, Z., & Muir, C. (2019). Creating a motivating classroom environment. In X. Gao (Ed.), *Second handbook of English language teaching* (pp. 719–736). Springer International Handbooks of Education. https://doi.org/10.1007/978-3-030-02899-2_36

Drakulić, M. (2018). Young learners' perceptions on effective foreign language teacher characteristics – A qualitative study from Croatia. In M. Orel & S. Jurjevič (Eds.), *EDUvision 2018 'Modern approaches to teaching the coming generations'* (pp. 284–300). Ljubljana, Slovenia.

Drakulić, M. (2019). Exploring the relationship between students' perceptions of the language teacher and the development of foreign language learning motivation. *Theory and Practice in Language Studies, 9*(4), 364-370, http://dx.doi.org/10.17507/tpls.0904.02

Enever, J. (Ed.). (2011). *ELLiE. Early language learning in Europe*. The British Council.

Enever, J., & Moon, J. (2009). New global contexts for teaching primary ELT: Change and challenge. In J. Enever, J. Moon, & U. Raman (Eds.), *Young learner English language policy and implementation: International perspectives* (pp. 5–21). Garnet Education.

European Commission. (2007). *Key competences for lifelong learning European reference framework*. Office for Official Publications of the European Communities. https://www.voced.edu.au/content/ngv:59967#

Fenyvesi, K. (2020). English learning motivation of young learners in Danish primary schools. *Language Teaching Research, 24*(5), 1–24. https://doi.org/10.1177/1362168818804835

Gardner, R. C. (2010). *Motivation and second language acquisition. The socio-educational model.* Peter Lang.

Gardner, R. C., & MacIntyre, P.D. (1993). A student's contribution to second-language learning. Part II: Affective Variables. *Language Teaching, 26*, 1–11. https://doi.org/10.1017/S0261444800000045

Gardner, R. C., & Mihaljević Djigunović, J. (2003). *AMTB*. University of Western Ontario, Canada and University of Zagreb, Croatia.

Hainzmann, S. (2009). "Girls are better at language learning than boys": Do stereotypic beliefs about language learning contribute to girls' higher motivation to learn English in primary school? *Bulletin VALS-ASLA (Swiss association of applied linguistics), 89*, 19–36. https://core.ac.uk/download/pdf/20644378.pdf

Horwitz, E. (1999). Cultural and situational influences on foreign language learners' beliefs about language learning: A review of BALLI studies. *System, 27*(4), 557–576. https://doi.org/10.1016/S0346-251X(99)00050-0

Hattie, J. (2003). *Teachers make a difference: What is the research evidence? Paper presented at the Building Teacher Quality: What does the research tell us*. ACER Research Conference, Melbourne, Australia. http://research.acer.edu.au/research_conference_2003/4/

Jakominić, N., & Mihaljević Djigunović, J. (2004). Učenički stavovi prema nastavniku stranoga jezika [Students' attitudes towards a foreign language teacher]. In D. Stolac, N. Ivanetić & B. Pritchard (Eds.), *Suvremena kretanja u nastavi stranoga jezika* (pp. 187–194). HDPL.

Johnstone, R. (2002). *Addressing 'the age factor': Some implications for language policy. (Guide for the development of language education policies in Europe. From linguistic diversity to plurilingual education. Reference study)*. Council of Europe. https://rm.coe.int/addressing-the-age-factor-some-implications-for-languages-policy-/1680886e92

Kadha, H. M. (2009). What makes a good English language teacher? Teachers' perceptions and students' conceptions. *Humanity and Social Sciences Journal, 4*(1), 1–11.

Mihaljević Djigunović, J. (1998). *Uloga afektivnih faktora u učenju stranoga jezika* [The role of affective factors in foreign language learning]. Filozofski fakultet Sveučilišta u Zagrebu.

Mihaljević Djigunović, J. (2007). Croatian EFL learners' affective profile, aspirations and attitudes to English classes. *Metodika, 8*(1), 115–126. https://hrcak.srce.hr/26944

Mihaljević Djigunović, J. (2012). Attitudes and motivation in early foreign language learning. *Center for Educational Policy Studies Journal, 4*(3), 55–74.

Mihaljević Djigunović, J. (2014). L2 learner age from a contextualized perspective. *Studies in Second Language Learning and Teaching, 4*(3), 419–441. https:doi.org/10.14746/ssllt.2014.4.3.3

Mihaljević Djigunović, J., & Nikolov, M. (2019). Motivation of young learners of foreign languages. In M. Lamb, K. Csizér, A. Henry, & S. Ryan (Eds.), *The Palgrave handbook of motivation for language learning* (pp. 515–533). Palgrave Macmillan.

Mihaljević Djigunović, J., & Lopriore, L. (2011). The learner: Do individual differences matter? In J. Enever (Ed.), *Early Language Learning in Europe* (pp. 43–61). The British Council.

Mihaljević Djigunović, J., & Letica Krevelj, S. (2010). Instructed early SLA – Development of attitudes. *Studia Romanica et Anglica Zagrabiensia, 54*, 137–156.

Muñoz, C., & Lindgren, E. (2011). Out-of-school factors: The home. In J. Enever (Ed.), *Early Language Learning in Europe* (pp. 103–124). The British Council.

Muñoz, C., & Singleton, D. (2011). A critical review of age-related research on L2 ultimate attainment. *Language Teaching, 44*(1). 1–35. https://doi.org/10.1017/S0261444810000327

Nikolov, M. (1999). 'Why do you learn English?' 'Because the teacher is short.' A study of Hungarian children's foreign language learning motivation. *Language Teaching Research, 3*(1), 33–56.

https://doi.org/10.1177/136216889900300103

Nikolov, M. (2002). *Issues in English language education*. Peter Lang.

Nikolov, M., & Mihaljević Djigunović, J. (2019). Teaching young language learners. In X. Gao (Ed.), *Second handbook of English language teaching* (pp. 1–23). Springer.

Nikolov, M., & Mihaljević Djigunović, J. (2006). Recent research on age, second language acquisition, and early foreign language learning. *Annual Review of Applied Linguistics, 26*, 234–260. https://doi.org/10.1017/S0267190506000122

Pfenninger, S. E., & Singleton, D. (2019). Making the most of an early start to L2 instruction. *Journal of Language Teaching for Young Learners, 1*(2), 1–35. https://doi.org/10.1075/ltyl.00009.pfe

Radišić, M., Pavičić Takač, V., & Bagarić, V. (Eds.) (2007). *Competences of primary school foreign language teachers in the Republic of Croatia*. Faculty of Teacher Education in Osijek.

Rúa, P. L. (2006). The sex variable in foreign language learning: An integrative approach. *Porta Linguarium, 6*, 99–114. https:/doi.org/10.30827/Digibug.30663

Sakurai, Y. (2012). Learners' perceptions of 'good' foreign language teachers: A quantitative analysis between native and non-native teachers. *Electronic Journal of FL Teaching, 9*(1). https://www.researchgate.net/publication/286273612_Learners%27_perceptions_of_good_foreign_language_teachers_A_quantitative_analysis_between_native_and_non-native_teachers

Sert, O. (2019). Classroom interaction and language teacher education. In S. Walsh & S. Mann (Eds.), *The Routledge handbook of English language teacher education* (pp. 216–238). Routledge.

Shishavan, H. B. (2010). The relationship between Iranian English language teachers' and learners' gender and their perceptions of an effective English language teacher. *English Language Teaching, 3*(3), 3–10. https:/doi.org/10.5539/elt.v3n3p3

Shishavan, H. B., & Sadeghi, K. (2009). Characteristics of an effective English language teacher as perceived by Iranian teachers and learners of English. *English Language Teaching, 2*(4), 130–143. https://doi.org/10.5539/elt.v2n4p130

Tragant, E., & Lundberg G. (2011). The teacher´s role: What is its significance in early language learning? In J. Enever (Ed.), *ELLiE. Early Language Learning in Europe* (pp. 81–102). British Council.

Vanhove, J. (2013). The critical period hypothesis in second language acquisition: A statistical critique and a reanalysis. *PLoS One, 8*(7). https://doi.org/ 10.1371/journal.pone.0069172

Williams, M., Burden, R., & Lanvers, U. (2002). 'French is the language of love and stuff': Student perceptions of issues related to motivation in learning a foreign language. *British Educational Research Journal, 28*(4), 504–528.

Yashima, T., Nishida, R., & Mizumoto, A. (2017). Influence of learner beliefs and gender on the motivating power of L2 selves. *The Modern Language Journal, 101*(4), 691–711. https://doi.org10.1111/modl.12430

Yuksel, H., & Halici, Y. (2010). Motivating young EFL learners through effective classroom management. In P. Hüseyin (Ed.), *2nd International Symposium on Sustainable Development* (pp. 105–115). International Burch University. https://core.ac.uk/download/pdf/153447149.pdf

Zamani, R., & Ahangari, S. (2016). Characteristics of an effective English language teacher (EELT) as perceived by learners of English. *International Journal of Foreign Language Teaching and Research, 4*(14), 69–88. http://jfl.iaun.ac.ir/article_563432.html

Biographical note

Morana Drakulić, PhD, is Assistant Professor in the field of foreign language teaching methodology currently employed at the Faculty of Teacher Education, University of Rijeka. Her main areas of interest are individual differences in foreign language learning, foreign language teacher competences, the relationship between students' perspectives of foreign language teacher and the affective variables within the elementary school context.

Appendix

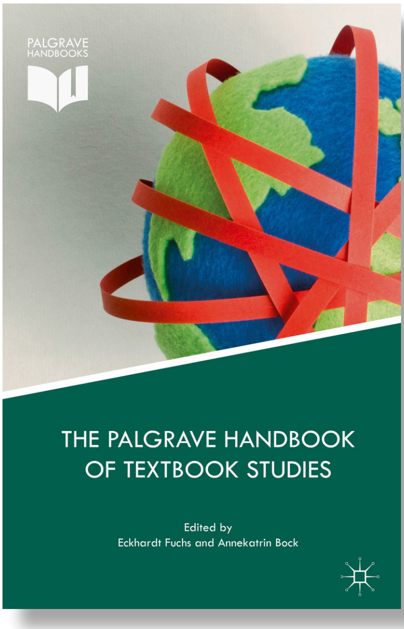
Current language teacher questionnaire

My real languge teacher...					
Gives me enough time to answer a question.	1	2	3	4	5
Is patient.	1	2	3	4	5
Is fair.	1	2	3	4	5
Accepts my opinion and my ideas.	1	2	3	4	5
Has a sense of humor.	1	2	3	4	5
Respects me.	1	2	3	4	5
Explains the teaching content in a manner that is clear and comprehensible to me.	1	2	3	4	5
Is a lenient grading teacher.	1	2	3	4	5
Understands me and establishes a good (friendly) relationship.	1	2	3	4	5
Speaks English in a manner comprehensible to me.	1	2	3	4	5
Speaks about her/his personal experiences.	1	2	3	4	5
Enjoys teaching English.	1	2	3	4	5
Often teaches about the culture and customs of the English language speaking countries.	1	2	3	4	5
Emphasizes the similarities and differences between Croatian culture and the English language speaking cultures.	1	2	3	4	5
Is well familiar with the culture of the English language speaking countries.	1	2	3	4	5
Uses different materials and activities during lessons.	1	2	3	4	5
Chooses interesting texts which describe different customs of the English language speaking countries.	1	2	3	4	5
Develops an understanding for cultural and social customs which differ from those in Croatia.	1	2	3	4	5
Sets clear rules which I should follow.	1	2	3	4	5
Examines the content that has been covered and practiced during lessons.	1	2	3	4	5
Cares about how much I have learned.	1	2	3	4	5
Always knows what we did and what has to be done.	1	2	3	4	5
Holds English language lessons in which I can work in peace.	1	2	3	4	5
Makes the lesson pleasurable and relaxing for work.	1	2	3	4	5
Encourages me to work and to actively participate during lessons.	1	2	3	4	5
Explains what I did well and what I need to work on in a manner comprehensible to me.	1	2	3	4	5
Checks understanding of the content by asking me questions.	1	2	3	4	5

doi: 10.26529/cepsj.1481

Eckhardt Fuchs and Annekatrin Bock (Eds.), *The Palgrave Handbook of Textbook Studies*, Palgrave Macmillan, New York: 2018; 432 pp.: ISBN: 978-1-137-53141-4

Reviewed by Mark Jupiter Užmah[1]



Textbooks and textbook-oriented studies are generally caught between two areas of research: publishing and education. Printed textbooks have played a vital role in the Western education system, at least until the beginning of the twentieth century, having been deemed the main source of knowledge. As a consequence of educational textbook usage, children and young adults established regular reading habits, but the many social, cultural and technological cha(lle)nges of the last century mean that textbook use, content, format and production have changed significantly. The present essay collection offers a much-needed insight into the field of textbooks, which it also broadens with vital input from social studies and the humanities.

The Palgrave Handbook of Textbook Studies consists of 30 essays, assembled and edited by two established researchers in the field, both affiliated with the Georg Eckert Institute for International Textbook Research (GEI) in Braunschweig, Germany. Eckhardt Fuchs is a historian and has been the director of the Institute since 2015. His research interests lie in globalisation processes related to education, European education policy, the history of transnational education relationships, the global history of textbook revision and the history of teaching materials (GEI: Prof. Dr Eckhardt Fuchs, n.d.). Annekatrin Bock is head of the Media in Schools research team at the Institute and is best known

1 Faculty of Arts, University of Ljubljana, Slovenia; mark.uzmah@gmail.com.

for her research work in educational media and school and classroom, as well as in reception and appropriation in media and communication studies. Dr Bock's recent work focuses on digital transformation and the change in learning culture (GEI: Dr Annekatrin Bock, n.d.).

The 36 contributors to this volume come from various fields connected to textbook studies, including publishing, digital media, education, curriculum studies, history, geography, linguistics and discourse studies, social anthropology and sociology. The editors thus succeed in producing a diverse and comprehensive volume, shedding light on textbooks from a variety of angles and discussing the historical, methodological, political and other dilemmas in the field.

As the editors explain in the *Introduction*, textbook studies are a substantially diverse area of research that demands an exploration of the common ground between the chapters' perspectives. They outline that "the conceptual focus revolves explicitly around textbook-related research questions informed by a cultural studies perspective" (p. 2) and further clarify that the field "centers on approaches from sociological and cultural studies, such as memory studies, sociology of knowledge, discourse theory, and media theory, alongside those drawn from the social and educational sciences" (ibid).

The book is arranged in four main parts, opening with an insight into the history, theory and methods of textbook research (Part I, Chapters 2–4), before continuing with a focus on the context in which textbooks are produced (Part II, Chapters 5–12), textbook content (Part III, Chapters 13–26), and the use, effects and practices of textbooks (Part IV, Chapters 27–29). The editors characterise Parts II to IV as "production, product, and reception" (p. 3). The book is rounded out with a conclusion, Part V, which contains reflections on future directions. Due to the length and complexity of the Handbook, not all of the chapters can be discussed here. Instead, the content of the four sections will be outlined and the chapters considered important for the conceptual framework of this issue of the CEPS Journal will be described.

Part I, entitled *History, Theory, and Methods of Textbook Research,* consists of three essays that focus on the origins and development of textbook studies through to the present. The first two chapters, *History of the School Textbook* by Steffen Sammler and *History of Textbook Research* by Eckhardt Fuchs and Kathrin Henne, are complementary in that they both discuss the transformation of school textbooks since the early modern period. Both pieces indicate the important conceptual turn that occurred in the last century when textbooks, initially considered books to be used in schools as well as in other situations, became an exclusively educational medium produced solely for school use. Additionally, technology has advanced considerably in the last decade, impacting both textbook usage in schools as well as the field of textbook research. The third essay, *Theories and Methods of Textbook Studies* by Annekatrin Bock, accordingly turns the reader's gaze to the methodological approaches to the multifaceted and manifold concept of textbook studies. Bock argues that the methodology of textbook research long had a reputation of being one-dimensional, focused solely on content analysis, but has expanded immensely in the last decade by using different approaches, such as narrative theory, interviews and computer-aided analysis. By outlining these numerous methods, both marginalised and established, the author paves the way for future research in the field. This first part of the Handbook could be seen as a prolonged introduction, since these three essays provide insight into the history, complexity and diversity of textbook research.

Part II, entitled *Textbooks in Their Context,* comprises eight chapters covering an array of topics relating to the social context of textbooks, from their publishing and authorship to their role in the education system in general and in particular areas, namely, the social sciences and the humanities, as well as science and geography. This section also addresses textbook quality by stressing the importance of both criteria and evaluation. Although the educational perspective is incorporated in the historically and methodologically engaged Part I, the connection between textbooks and education is even more apparent in Part II, which highlights the societal conditions around the production of textbooks and the decisions over their substance, the knowledge included in them. In the chapter *Textbooks and Education*, Eugenia Roldán Vera sheds light on the importance of social context in the process of textbook production, which is inherently connected to its content. Textbooks have changed considerably since the nineteenth century, when they played a crucial role in providing pupils with common representations of the past, thus building uniformity in the self-perception of nations. The atrocities of World Wars I and II triggered analysis and wider consideration of the role of education in the generation of conflict, and this had a considerable impact on textbook research and on understanding the role of textbooks in contemporary societies. Later developments brought additional challenges, such as the multiplication of knowledge sources and digitalisation. The latter is discussed in Christoph Bläsi's *Educational Publishers and Educational Publishing*, in which specific features of textbook publishing across disciplines in several Western and developing countries are explored.

The third part of the book, *Textbooks and Their Contents*, is the most extensive, containing 14 essays that build on the dynamics between context and content. As highlighted by the editors, Part III is intrinsically divided into three subsections, with the first focusing on national, transnational and regional identities, as well as on representations of class, race, gender and sexual orientation in textbooks. The second set of chapters puts a spotlight on textbook representations of the historical and cultural issues of religion, nationalism and the Holocaust, colonialism, socialism and, more generally, dictatorship and war. Sylvie Guichard's impactful *War in Textbooks* is cited in many of the Handbook's other chapters because it demonstrates how fundamentally textbook content was affected by the two world wars. The third section in Part III deals with political and economic concepts, such as human rights education in Patricia Bromley and Julia Lerch's *Human Rights as Cultural Globalisation: The Rise of Human Rights in Textbooks, 1890–2013)* and the climate crisis in Tobias Ide's *The Environment*.

Part IV, entitled *Textbook Use, Effects, and Practices*, centres on the usage of textbooks by different users in both formal and informal education. As mentioned above, textbooks were originally just books used in schools and other places, but they gradually transformed into specifically educational materials or tools. As Thomas Illum Hansen explains in *Textbook Use,* this development led to textbooks being used by students as well as by teachers. Today, textbooks remain the most regularly used learning materials. As several other authors in this Handbook note, Illum Hansen explores the emergence of digital textbooks and other digital teaching materials, referencing a systematic review of 16 open educational resource studies and concluding that there is little difference in the learning outcomes of digital and paper-based learning materials. Although students prefer printed textbooks to digital versions, they perceive their quality as comparable. The differences between reading (in general rather than specifically textbooks) on screens or on paper has been increasingly researched over the last several years, and meta-study findings suggest the opposite of Illum Hansen's first argument and affirm the latter (Singer and Alexander, 2017; Delgado et al., 2018; Clinton, 2019). Reading on a screen differs from reading on paper; it is harder to read long-form informational text on a screen because we remember less and it is more exhausting than reading from paper (Stavanger Declaration 2019). This topic relates to digitalisation, but it is also very closely connected to the complexity of content in books and textbooks, which leaves room for future studies.

The Handbook's conclusion is called *New Directions*, or Part V, and is written by Barbara Cristophe, Annekatrin Bock, Eckhardt Fuchs, Felicitas Macgilchrist, Marcus Otto and Steffen Sammler. It outlines the key trends presented in the Handbook and points out the central shortcomings of the field. The authors acknowledge the already significant broadening of the range of theoretical approaches to textbook studies, which has led to the area being "well enough grounded to make significant contributions to theoretical debate in several areas" (p. 414). Furthermore, they remark on the significant changes in the perception of textbooks, which are now studied as complex media rather than simple transmitters. This closing chapter also underlines possible ideas for future research, such as expanding into geographical regions that have been neglected in certain respects and inventing methods for analysing different teaching materials, especially in relation to digital textbooks and open resources.

The Palgrave Handbook of Textbook Studies sheds light on the development of textbook-related research, focuses on the importance of the social conditions in which textbooks are produced, and analyses the role of teaching with school textbooks in the process of achieving educational aims. It shows how the field of textbook studies has been established over time and how interconnected it is not only with education studies, but also with many disciplines in the social sciences and the humanities. The collection identifies many pressing questions relating to textbook studies, some of which are briefly outlined above, and leaves room for further reflection and analysis. For example, the severe reductions and changes in what students learn at school over the last few decades has become a burning issue, impacting how and to what extent we access the social stock of knowledge. It has also severely affected textbook production and usage. Whether future research focuses on historical or methodological perspectives or on another area of textbook studies, this Handbook provides a comprehensive review of existing theory and research and is therefore a vital contribution to a field that has had troubles in the past, but has recently laid its scholarly foundations.

References

Clinton, V. (2019). Reading from paper compared to screens: A systematic review and meta-analysis. *Journal of Research in Reading, 42,* 288–325. https://doi.org/10.1111/1467-9817.12269

Delgado, P., Vargas, C., Ackerman, R., & Salmerón, L. (2018). Don't throw away your printed books: A meta-analysis on the effects of reading media on comprehension. *Educational Research Review, 25,* 23–38. https://doi.org/10.1016/j.edurev.2018.09.003

GEI: Dr. Annekatrin Bock. (n.d.). http://www.gei.de/en/mitarbeiter/dr-annekatrin-bock.html

GEI: Prof. Dr. Eckhardt Fuchs. (n.d.). http://www.gei.de/en/mitarbeiter/prof-dr-eckhardt-fuchs.html

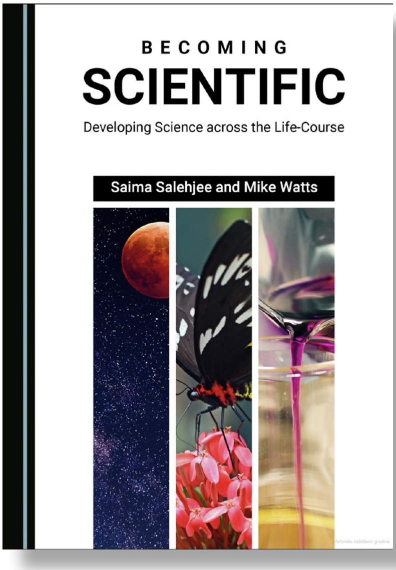
Singer, L. M., & Alexander, P. A. (2017). Reading on paper and digitally: What the past decades of empirical research reveal. *Review of Educational Research, 87*(6), 1007–1041. https://doi.org/10.3102/0034654317722961

Stavanger Declaration. (2019). https://ereadcost.eu/wp-content/uploads/2019/01/StavangerDeclaration.pdf

doi: 10.26529/cepsj.1480

Saima Salehjee and Mike Watts, *Becoming Scientific: Developing Science across the Life-Course*, Cambridge Scholars Publishing: 2020; 257 pp.: ISBN: 1-5275-5498-8

Reviewed by Keith S. Taber[1]



Being and becoming 'sciencey'

Saima Salehjee and Mike Watts have produced a very thought-provoking book considering the nature of peoples' relationships and identification with science over the life course. '*Becoming Scientific*' is not about science education *per se*, or at least not formal science education. Rather it considers the different influences (schooling included, but as just one among various elements) that might lead to people identifying with science, valuing science, rejecting science, and so forth, and considering themselves as 'sciencey', or not. The flavour (sic) of their mission is reflected in the culinary analogy they choose to use in the book, as in the following taster:

> …Do some people come 'sciencey-ready flavoured', or is it possible to 'science marinate' them over time? [Our view is, of course that both of these are possible.] How can we 'science' them? …A key follow-up question then might be: what exactly is the balance of ingredients - the formula of that 'science marinade' to help people become 'sciencey'? (pp. 2-3)

Salehjee and Watts consider the nature of identification with science or 'Sci-ID.' Although not limiting itself to formal education, the book clearly has great relevance for schools and schooling when we consider our purposes whether as educators or more generally as members of a society prescribing education. Clearly education policies and curricula vary from place to place, but it is very common that (a) young people, at least whilst considered children

1 Faculty of Education, University of Cambridge, United Kingdom; kst24@cam.ac.uk.

and often beyond, are expected (that is, usually required) to attend school, and (b) science is often part of a core curriculum. That is, contemporary societies generally consider that it is important for all citizens to be taught some science.

One rationale for this is that societies need a supply of scientists, engineers, medical professionals, technicians, and so forth. Yet it has long been argued that there is something questionable about a science education whose prime purpose is to service the career needs of a minority of those in the population who aspire to, and might be selected for, science-based occupations. Such a perspective can be considered disrespectful to the majority and perhaps undemocratic (a misuse of power, not considering the rights of the child) or even a kind of misplaced elitism. Of course, all children need to experience something of science if they are to make an informed choice about whether to proceed to science-based advanced study and/or careers, but that experience should be a science education for all, not just offering a foundation for elective higher level study. Moreover, even for the minority who will become scientists, a science curriculum focused on preparation for future study offers an impoverished science course.

The usual argument here looks at other perspectives on the purpose(s) of education besides the economic driver to provide personnel for employment. Education should be about supporting the development of the whole, well-balanced, person; about providing the basic skills required for adult life; about offering glimpses of, and pathways towards, different future possibilities (including, but not limited to, employment options); and induction into the culture(s) of the society. This latter strand may be considered a form of indoctrination (a term that need not necessarily carry negative associations), but also as an enablement or affordance. That is, something needed to take one's full place in society – just as how in a cultural tradition where community dancing is core to rituals and social activities a child would need to be inducted into dance; or as in a society where a canon of epic poetry was used as a key referent to discuss and understand social and political life then a full education must encompass examination of those poems.

Science is a key part of the cultures of modern democratic societies. Being cultured in such societies must therefore include a level of familiarity with science – its nature, and some of its products (e.g., models, theories, concepts) and applications. There was a time when this was not assumed, when, for most of those with influence in society, 'culture' meant music, literature, fine art, and so forth, and science was seen (if mistakenly) as a minority technical interest (Snow, 1959/1998). That stance is certainly not viable today. Full *engagement* in civic society is not possible without being able to enter into meaningful discourse about such issues as climate change, energy supply, deforestation, pollution, biodiversity, recycling, genetic medicine and so forth. Moreover, everyday media and public conversation encompasses such ideas as evolution, atoms, extinction, the double helix, nuclear power, and the exponential growth of infections. It is not a matter of passing examinations, but of making lifeworld decisions (for example, about healthcare), and being included in everyday discourse.

Educators may see this in terms of the different *drivers* for education – what society seeks to achieve by committing so much public resource, and indeed in somewhat restricting individual choice, in prescribing compulsory schooling. Salehjee and Watts's engaging book reminds us that there is another way to look at this: not what society collectively thinks is good for people, but what people themselves come to value, and engage with, in relation to science. After all, science is the study of the natural world writ large, and that is going to be relevant to all.

Salehjee and Watts draw upon a range of different studies to support the arguments they make in the book. Those studies have taken place at different times, and for different purposes, with foci such as emotional responses to science, and using student question-posing in science education, as well as the authors' more recent work. However, the authors marshal their materials in the cause of an overall argument and mission.

So, Salehjee and Watts explore how different people relate to, engage with, and feel about aspects of science – even when those people themselves may not be primarily framing this engagement as a 'science' interest or activity. In this regard, their book has some similarity with Joan Solomon's last book, '*Science of the People*' (Solomon, 2013) which offered an ethnographic account of how people in Market Town (an assumed name) thought about science. One strong impression from reading that book was "that most people have interests related to science (even if they do not always recognise this), but even so they seldom rely heavily on their learning from formal science education" (Taber, 2015, p. 111). The accounts offered by many of the adult contributors to Salehjee and Watts's study would seem to reinforce that view. Sandra, a primary teacher with an arts and humanities background, but who expresses awe when learning about topics such as black holes and anti-matter, even offers a vignette supporting Solomon's (1992) account of how often learners' formal science learning largely takes place in isolation from their everyday, their everyday, 'lifeworld' notions of scientific topics deriving from quotidian discourse and activities,

> "…I am left with two often contradictory strands of thought that exist simultaneously. At times I actually make little attempt to reconcile my own view with the scientific one because, in my head, they do not come into conflict. Each idea is logged in separate compartments, so for me there is no real contradiction." (p. 34)

Whereas Solomon's study sought to explore the thinking and discourse of a variety of people in one geographic place, Salehjee and Watts's book offers a series of snapshots from different contexts to collectively build up a picture across the life course. In that sense, the book might be considered loosely cross-sectional, but rather than being the report of a single coherent study with sampling of people at different ages, it is more a patchwork of accounts of related studies which collectively build up an overall picture. These different slices of data cannot be considered to be strictly comparable as in a true cross-sectional study (and that is never claimed to be the aim), but certainly complement each other to make a very readable and informative book. Perhaps the weakest link in terms of comparability is the 'slice' taken from a higher education institution (Brunel University in London), where the focus shifts somewhat from natural science as such to the wider notion of 'STEM' (science, technology, engineering and mathematics), both in terms of a survey of students and the classification of case studies of academics (so lecturers in computer engineering and mathematics education are classed as 'scientists' for the purposes of the book). STEM has gained a foothold in international education discourse, but is more an alliance of discrete disciplines with some common interests than a unified domain. This raises the issue of whether (or, perhaps better, when) STEM can be seen as a proxy for science (or vice versa), or even to what extent science identify should be seen as a unitary notion (rather than physics identify, biology identity, and so forth). There are surely substantive differences in interests, priorities, foci and motivations – even if there is seldom a complete demarcation – between doing science as a means to better understand the world and applying science as a means to meet practical ends. As one example, whilst science qualifications are needed to study medicine, it is not clear if something like Sci-ID is so well aligned with aspiring to be a medical doctor.

Largely, the book can be considered to offer accounts of naturalistic studies, although one of the projects (in Chapter 6) reports on secondary students' responses to a year-long intervention to offer curriculum enrichments to a class of 13-year olds in a Muslim girl's school. This was an independent (i.e., private, fee paying) school that limited the science studied in the curriculum because "parents like their daughters to opt for and spend more time in studying religious education and humanities-based subjects" (p. 105). Perhaps in part reflecting such parental views, and the school's accommodation of them, one of the students commented that "unlike religious studies, science does not make sense at all, and it is not what I see, hear and feel in my day-to-day life" (p. 106).

In many ways, one of the key messages from the book is how contingent so many careers are: having a significant relative in a particular job (or suffer some major health issue), a particular liking (or disliking) for a particular teacher, a response to some particular experience – such things can change the direction of a life. This, of course, becomes clear from the kind of conversational approach to research underpinning most of the studies drawn upon. An idiographic method that invites personal narratives reveals the idiosyncrasies of lives (and so lived realities) that tend to be obscured in approaches that seek to measure population variables.

Being 'sciencey' is, according to Salehjee and Watts, "the very nature of being a scientist" (p. 10), and is linked to what might be termed an enquiring nature about the natural world. They present a notion of a 'sciencey' person as someone not happy to adopt the natural attitude (Schutz & Luckmann, 1973) and just accept that the sun rises each day to illuminate the world, but to ask why, and what the sun actually is; a sciencey person is not content to take pleasure in the aesthetic response to the colours of flowers but rather enquires into why these phenomena have arisen. Yet, as these authors acknowledge, it is important to avoid representing science in science education as a cold, objective, rationale activity that is only interested in enquiry as a cognitive activity. For one thing, it would misrepresent the nature of science and scientists. Scientists may need to learn to bracket off the affective response to focus on the objective analysis of data during *some stages* of enquiry (and objectivity is a kind of ideal, which actual *human* scientists can at best only approach), but the aesthetic appeal of the natural world has been what has initially attracted many scientists to the focus of their enquiries, and – even if the lay person may not always appreciate this – for the scientist, understanding 'how' and 'why' often adds to the wonder of the natural phenomenon rather than simply explaining it away. So, to wonder at the complexity, subtlety and variety of human anatomy and to see it as evidence of a master craftsman creator, as William Paley (1802/2006) famously did, surely pales (sic) beside a neo-Darwinian account that seeks to explain how such exquisite 'design' can arise through contingency and natural mechanisms.

Moreover, to eliminate non-scientific values from science education would be completely at odds with the need for scientific literacy, as science education needs to prepare people to engage with science in the context of socio-scientific decision-making, where what is technically possible, and the likely consequences of different choices, need to be considered in the light of considerations external to science. So, when 13-year-old Vanessa is ambivalent about whether personal feelings should be valid features of science lessons, and suggests that "if you are dissecting an animal for data, you may feel sympathy for it and not do it" (p. 24), this should not be considered as an over-emotional child being unable to focus scientifically in the classroom, but rather as a human being engaging their personal values to decide to treat a specimen as a creature that can be harmed rather than just scientific (or actually, here, educational) source material (cf. Keller, 1983). That, surely, is a good sign. Many people who would not consider themselves 'sciencey' would associate the image of cold, logic-led scientists with such cultural referents as Hiroshima and Nagasaki; Three Mile Island and Chernobyl; Bhopal and Seveso; chlorine used in battle in World War 1, and Zyklon B used to kill innocent non-combatants in concentration camps in World War II; smoking beagles, and rabbits used as living indicators to test cosmetics; and so forth. A science education that does NOT encourage learners to ask questions about the ethical aspects of both scientific enquiry and technological applications of science is not only deficient, but also unfit for purpose.

Through the book, Salehjee and Watts build up a "theoretical model of science identify: Sci-ID" (p. 59), considering how individuals are influenced by a combination of societal, community and personal factors. They draw upon various theoretical considerations and a good deal of data. Perhaps, for some, this will be the book's main strength. Whilst in some sense a work of bricolage, drawing from here or there, this is not an uncritical accretion of disparate ideas – for example, the authors show some scepticism about work around the notion of 'science capital' as tending to stress "the power of extensionality – the role of macro forces – more than the power of intentionality in science identify formation and transformation" (p. 66). Archer and her colleagues have proposed this construct (science capital) as a label for the "science-related forms of cultural and social capital…[for use] as a theoretical lens for explaining differential patterns of aspiration and educational participation among young people" (Archer et al., 2015, p. 922). Whilst science capital has certainly proved effective as a rhetorical device for drawing attention to a wide range of factors that can influence a young person's career (in the broad sense of the term), it is questionable whether a construct encompassing aspects of a person's experiences, aspirations, beliefs, values, knowledge – with aspects of the attitudes and behaviours of others that interact with them mixed in – has the ontological coherence to be considered as more than a useful theoretical phantom.

It could be suggested that Salehjee and Watts's own model of Sci-ID suffers a similar flaw, but whereas Archer et al. (2015) see science capital as something that can be quantified in a single score through a multi-scaled questionnaire – so that in an English sample of "3658 secondary school students, aged 11–15 years" (p. 922) they identified that 5% had a "high science capital" score (p. 936)[2] – Salehjee and Watts offer an 'ecological' model of the layers of factors *impacting on* their Sci-ID, which they see as a variable characteristic of individuals: "the individual responds to both…his or her own internal sense of integrity, as well as to the external medium and large-scale forces that are in operation at any one time" (p. 73). As well exploring the stories their informants tell to conjecture about critical factors influencing Sci-ID, the authors are able to demonstrate the fluid nature of the construct with case studies of the shifts in the responses from schoolgirls who had been involved in the curriculum enhancement intervention.

Salehjee and Watts have brought together a great deal of testimony from a diverse group of people of different ages, inside and well outside formal education, illustrating just how variously people relate to and engage with (or sometimes disregard) science. This can be considered an important contribution, as although there are many scattered studies offering such glimpses, this modest volume gives voice to many different individuals who were prepared share their ideas, views, interests and responses to natural phenomena and science. If the mantra of 'science for all', which is supposed to be the mission of school science in so many countries, is to be taken as more than a slogan, then perhaps all those preparing to work as science teachers should read this book and reflect on how science lessons can genuinely help support everyone in the population to become more scientifically literate, and perhaps even encourage them to become more 'sciencey'.

2 It would seem arbitrary to assign any level of score as 'high' (or 'low' or whatever) *without* having some independent measure of 'science capital'. Archer et al. (2015: 936) are quite open that they obtained their result by considering the range of 'science capital' scores following statistical manipulation (a regression analysis informed by responses to five items assumed to reflect "future science job affinity"). As these processed scores ranged from 0 to 115, they simply divided the scale into three equitable regions: 'low' (0-34: 27% of the sample), 'medium' (35-6: 68% of the sample) and 'high' (70-105).

References

Archer, L., Dawson, E., DeWitt, J., Seakins, A., & Wong, B. (2015). "Science capital": A conceptual, methodological, and empirical argument for extending bourdieusian notions of capital beyond the arts. *Journal of Research in Science Teaching, 52*(7), 922–948. https://doi.org/10.1002/tea.21227

Keller, E. F. (1983). *A feeling for the organism: The life and work of Barbara McClintock*. W H Freeman and Company.

Paley, W. (1802/2006). *Natural theology: Or evidence of the existence and attributes of the deity, collected from the appearances of nature* (M. D. Eddy & D. Knight Eds.). Oxford University Press.

Schutz, A., & Luckmann, T. (1973). *The structures of the life-world* (R. M. Zaner & H. T. Engelhardt, Trans.). Northwest University Press.

Snow, C. P. (1959/1998). The Rede Lecture, 1959: The two cultures. *The Two Cultures* (pp. 1–51). Cambridge University Press.

Solomon, J. (1992). *Getting to know about energy – in school and society*. Falmer Press.

Solomon, J. (2013). *Science of the people. Understanding and using science in everyday contexts*. Routledge.

Taber, K. S. (2015). What little they remember: Understanding science in the life-world. *Studies in Science Education*, 1–12. https://doi.org/10.1080/03057267.2014.994933

Instructions for Authors for publishing in CEPS Journal (https://www.cepsj.si/index.php/cepsj/about/submissions)

Submissions

Manuscript should be from 5,000 to 7,000 words long, including abstract and reference list. Manuscript should be not more than 20 pages in length, and should be original and unpublished work not currently under review by another journal or publisher.

Review Process

Manuscripts are reviewed initially by the Editors and only those meeting the aims and scope of the journal will be sent for blind review. Each manuscript is reviewed by at least two referees. All manuscripts are reviewed as rapidly as possible, but the review process usually takes at least 3 months. The ceps Journal has an online-based review system via the Open Journal System. All submissions should be made via the OJS – http://cepsj.si/.
For more information visit our web page http://cepsj.si/.

Abstracting and indexation

Scopus | EBSCO - Education Source Publications | Cooperative Online Bibliographic System and Services (COBISS) | Digital Library of Slovenia - dLib | DOAJ - Directory for Open Access Journals | Academic Journals Database | ERIH PLUS | ERIC | Elektronische Zeitschriftenbibliothek EZB (Electronic Journals Library) | Base-Search | DRJI - The Directory of Research Journal Indexing | GSU - Georgia State University Library | MLibrary - University of Michigan | NewJour | NYU Libraries | OhioLINK | Open Access Journals Search Engine (OAJSE) | peDOCS: open access to educational science literature | ResearchBib | Scirus | Ulrich's International Periodicals Directory; New Providence, USA

Annual Subscription (4 issues). Individuals 45 €; Institutions 90 €. Order by e-mail: info@cepsj.si; postal address: CEPS Journal, Faculty of Education, University of Ljubljana, Kardeljeva ploščad 16, 1000 Ljubljana, Slovenia.

Online edition at **http://cepsj.si/**.

Number of copies printed: 150

Navodila za avtorje prispevkov v reviji (https://www.cepsj.si/index.php/cepsj/about/submissions)

Prispevek

Prispevek lahko obsega od 5.000 do 7.000 besed, vključno s povzetkom in viri. Ne sme biti daljši od 20 strani, mora biti izvirno, še ne objavljeno delo, ki ni v recenzijskem postopku pri drugi reviji ali založniku.

Recenzijski postopek

Prispevki, ki na podlagi presoje urednikov ustrezajo ciljem in namenu revije, gredo v postopek anonimnega recenziranja. Vsak prispevek recenzirata najmanj dva recenzenta. Recenzije so pridobljene, kolikor hitro je mogoče, a postopek lahko traja do 3 mesece. Revija vodi recenzijski postopek preko Open Journal System (OJS). Prispevek oddaje na strani: http://cepsj.si/.

Več informacij lahko preberete na spletni strani http://cepsj.si/.

Povzetki in indeksiranje

Scopus | EBSCO - Education Source Publications | Cooperative Online Bibliographic System and Services (COBISS) | Digital Library of Slovenia - dLib | DOAJ - Directory for Open Access Journals | Academic Journals Database | ERIH PLUS | ERIC | Elektronische Zeitschriftenbibliothek EZB (Electronic Journals Library) | Base-Search | DRJI - The Directory of Research Journal Indexing | GSU - Georgia State University Library | MLibrary - University of Michigan | NewJour | NYU Libraries | OhioLINK | Open Access Journals Search Engine (OAJSE) | peDOCS: open access to educational science literature | ResearchBib | Scirus | Ulrich's International Periodicals Directory; New Providence, USA

Letna naročnina (4 številke). Posamezniki 45 €; pravne osebe 90 €. Naročila po e-pošti: info@cepsj.si; pošti: Revija CEPS, Pedagoška fakulteta, Univerza v Ljubljani, Kardeljeva ploščad 16, 1000 Ljubljana, Slovenia.

Spletna izdaja na **http://cepsj.si/**.

Naklada tiskane izdaja: 150 izvodov

University of Ljubljana
Faculty of Education

C O N T E N T S

C·E·P·S Journal

Center for Educational Policy Studies Journal

Revija Centra za študij edukacijskih strategij

Vol.12 | Nº2 | 2022

http://cepsj.si

ISSN 1855-9719