14 **Learning Tech**

Tidsskrift for læremidler, didaktik og teknologi

LÆRE
MIDDEL
DK

AI og didaktik i uddannelse

14 **Learning Tech**
Tidsskrift for læremidler, didaktik og teknologi

AI og didaktik i uddannelse

Learning Tech – Tidsskrift for læremidler, didaktik og teknologi
Udgives af Læremiddel.dk

Learning Tech er et forskningstidsskrift, hvor alle artikler er forskerbedømt i form af dobbeltblindt peer review. Tidsskriftet bringer artikler, der rammer genstandsfeltet mellem læremidler, didaktik og teknologi, og hensigten er at spille en betydelig rolle som platform for den voksende skandinaviske læremiddelforskning.

Design
Trefold

Grafisk tilrettelæggelse
Kamilla Bjørnskov Olsen, UCL Erhvervsakademi og Professionshøjskole

ISSN 2445-6810 (ONLINE)

Kontakt
Læremiddel.dk, Niels Bohrs Allé 1, 5230 Odense M
www.laeremiddel.dk

LÆRE
MIDDEL
DK

135 Danske gymnasieelevers multimodale anvendelse i fysiske og digitale læringsrum

Et observationsstudie af danske gymnasieelever, med henblik på undersøgelse af deres faglige læsning.

Af Michael Juul Nielsen & Lasse Bo Jensen

LÆRE
MIDDEL
DK

Forord

AI-teknologi har de senere år gjort store fremskridt og har (måske) potentialet til at revolutionere måden, vi underviser og lærer på. Som så ofte før, når ny teknologi møder uddannelsessektoren, har debatten om AI's betydning for uddannelse delt vandene. Brugen af AI i undervisning kan potentielt personalisere læring, øge elevernes motivation og engagement, og hjælpe lærerne og underviserne med effektivt at evaluere elevernes progression. Omvendt tegner kritiske røster et billede af et uddannelsessystem med en større grad af social isolering. Frygten for en dehumanisering af arbejdsområder, der tidligere var menneskets domæne, samt en ikke-demokratisk og på anden vis uforholdsmæssig stor magt til diverse tech-giganter, er også markant. Fra et læremiddelperspektiv, vil dette temanummers tre redaktører hævde, at vi som udgangspunkt altid må prøve at forstå de nye teknologier og deres potentielle virkninger på undervisning og læring, før vi fælder dom over teknologierne. Dette skriver vi velvidende, at vi ikke kan forudse, hvordan for eksempel en ny teknologi som AI vil forandre de måder, hvorpå vi underviser og lærer - og de måder hvorpå vi overhovedet anskuer og problematiserer undervisning, evaluering og læreprocesser.

Vi kan opfatte teknologi som kultur i færd med at skabe sig selv. Den franske teknologifilosof Gilbert Simondon (1924-1984) beskriver teknologi som menneskets kapacitet til at være selvbetinget. Vi (om-) skaber vores eget miljø samtidigt med, at vi i processen reorganiserer os selv på en måde, så organisme og miljø forandrer de komplekse problemer, vi står overfor. Teknologi skal derfor heller ikke anskues som et neutralt middel, medium, artefakt eller værktøj, hvor kun vores intentioner bestemmer deres betydning. Teknologiens værdi er, på godt og ondt, båret af vores stående kapacitet til at vurdere og omsætte dem i handling, materialer, netværk og infrastruktur. Og det vi gør, er ikke altid det samme som det, vi vil, eller tror vi gør.

I disse dage er det AI, der skaber røre i uddannelsessektoren og udfordrer vores forståelser af didaktik, læring, evaluering og læremidler. I denne udgave af Learning Tech præsenterer vi fire artikler under temaet AI i didaktik og uddannelse.

I temanummerets første artikel, *Use of AI-powered technologies in upper secondary language learning*, præsenterer Bundgaard og Møller resultaterne af en survey-undersøgelse af gymnasieelevers brug af og holdninger til AI-baseret maskinoversættelse i fremmedsprogsundervisningen. Netop AI-støttede sprogteknologier er et eksempel på, at AI er udbredt og efterhånden har været i spil i undervisningsverdenen i et stykke tid, i hvert fald for elevernes vedkommende. Men at undervisere kan have undladt at forholde sig tilstrækkeligt til teknologien.

I artiklen *Between the Clicks – Student learning paths when interacting with an adaptive learning resource in 4th grade mathematics*, undersøger Gissel og Jørnø fjerde klasse-elevers interaktion med et adaptivt læremiddel til matematikundervisning. Artiklens ærinde er, at opnå en dybere forståelse af interaktionen mellem elev og et AI-baseret læremiddel gennem skærmoptagelser af elevernes adfærd i forhold til de faglige udfordringer, læremidlet præsenterer eleverne for. Dermed giver artiklen indsigt i, hvordan det adaptive læremiddel håndterer elever der agerer forskelligt og kommer med forskellige forudsætninger, samtidig med, at der tegnes et billede af, hvordan eleverne reagerer på at være i hænderne på den adaptive maskine, Rhapsode.

I temanummerets tredje artikel, *Scaffolding teaching – between learning objective and assessment when using AI chatbots*, undersøger Jensen, Gade og Dau etiske og pædagogiske perspektiver på anvendelsen af ChatGPT gennem et casestudie. Artiklen undersøger studerendes perspektiver på, hvordan undervisere kan stilladsere deres brug af ChatGPT i relation til arbejdet med læringsudbytter og undervisernes evalueringspraksis. Studiet er baseret på fokusgruppeinterviews med studerende på en videregående uddannelse. På baggrund af de empiriske data udledes en begrebsramme for anvendelse af generativ AI, som undervisere på de videregående uddannelser kan anvende til at stilladsere læringen hos de studerende. Begrebsrammen adresserer kognitive, pædagogiske og undervisningsmæssige perspektiver på aktiviteter og evaluering i relation til anvendelsen af ChatGPT i undervisning og læreprocesser.

I temanummerets fjerde artikel *Måling af flydende læsning med øjenbevægelser i skolen*, præsenterer Klerke og Engmose et studie, hvor der anvendes eye-tracking teknologi til at måle elevers læsning i et forsøg på at oversætte data til læsefaglige indsigter. Med teknologien kan nedbrud i elevernes flydende læsning indfanges, for at finde korrelationer mellem øjenbevægelser og læse-/stavetestresultater og fokusord. Studiet baserer sig på data fra 68 elever fra fem fjerde klasser. Hensigten er at omsætte sådanne mål til praktisk relevans, eksempelvis undervisningsevaluering.

Temanummerets femte og sidste artikel *Danske gymnasieelevers multimodale anvendelse i fysiske og digitale læringsrum* af Nielsen og Jensen bygger på observationsstudier fire gymnasiale retninger. I artiklen er der fokus på elevernes faglige læsning, og der identificeres fire forskellige digitale elevtyper, som hver for sig på forskellig vis anvender digitale læremidler. De fire identificerede digitale elevtyper er i artiklen karakteriseret som: Digitale passagerer, digitale drømmere, digitale solister og digitale sociale skabere. På baggrund af artiklens fund udledes der anbefalinger til designprincipper i relation til udvikling af digitale læremidler. Artiklen bidrager ligeledes med perspektiver på hvordan elevadfærden kan styrke fællesskaberne i de transformerede læringsrum og hvordan elevers multimodale blik kan understøtte ekspertise. Artiklen er udenfor tema.

Temaredaktionen
Rasmus Leth Jørnø, Susanne Dau & Stig Toke Gissel

Abstract

For many years, language technologies based on artificial intelligence (AI) have been influencing foreign language teaching. This is only expected to increase with the introduction of tools like ChatGPT. Until now, particularly machine translation (MT) has challenged Danish upper secondary language teachers, and the reaction has typically been to prohibit MT and thus not integrate it into teaching. However, this can be problematic if students use MT anyway, e.g. because it can result in inappropriate use of the technology. Likewise, today, we often come across machine translations and other AI-based texts online. This article presents the results of a survey conducted in language subjects of upper secondary students' use of and attitudes towards MT. The study shows a widespread use of MT, e.g. for homework and hand-in assignments. Against this background, and based on digital literacy theory, perspectives in incorporating AI-powered language technologies into foreign language teaching are discussed.

Sprogteknologier, der er baseret på kunstig intelligens (AI), har påvirket fremmedsprogsundervisningen i mange år, og dette forventes kun at tiltage med introduktionen af værktøjer som ChatGPT. Hidtil har særligt maskinoversættelse udfordret fremmedsprogslærere på danske ungdomsuddannelser, og reaktionen har oftest været at forbyde brugen og dermed ikke integrere teknologien i undervisningen. Det kan imidlertid være problematisk, hvis eleverne alligevel bruger maskinoversættelse, da det bl.a. kan resultere i, at de bruger teknologien på uhensigtsmæssige måder. Ligeledes sker det ofte i dag, at vi støder på maskinoversættelser og andre AI-baserede tekster på internettet. I denne artikel præsenteres en undersøgelse af gymnasieelevers brug af og holdninger til maskinoversættelse. Den indsamlede empiri er baseret på spørgeskemaundersøgelser i sprogfag på HHX og HTX. Undersøgelsen viser en udbredt brug af maskinoversættelse, bl.a. ifm. lektier og afleveringsopgaver. På denne baggrund og med udgangspunkt i teori om digital literacy diskuteres potentielle perspektiver i at inddrage AI-støttede sprogteknologier i fremmedsprogsundervisningen.

Use of AI-powered technologies in upper secondary language learning

Current tendencies and future perspectives

Kristine Bundgaard, Aalborg University, &
Anders Kalsgaard Møller, Aalborg University

Introduction

For several years, we have discussed the need to build competencies that can prepare us for the 21st century (Griffin & Care, 2014; González-Pérez & Ramírez-Montoya, 2022). The digitization of society is often referred to as the third industrial revolution and marks the transition to the 21st century, however, Schwab (2017) has suggested a fourth industrial revolution which deals with automation, artificial intelligence (AI) and the use of robots. While this revolution has slowly evolved over the past years, with e.g. the development surrounding autonomous vehicles, people seem to have only recently realized the extent to which these technologies have entered, and will increasingly enter, our work and civic life. This is particularly apparent with the current discussion of generative AI technologies such as ChatGPT (OpenAI, 2023a). With such developments we believe it is necessary to shift focus from discussing the more loosely defined competencies for the future to discussing the more specific competencies necessary to deal with the present solutions. Not least within education, a discussion of how we can start to develop these competencies is imperative.

AI-powered technologies have been changing the conditions for foreign language teaching and learning for several years (Lee, 2023). The most prominent example is machine translation (MT) which is the automatic translation of text, i.e. translation without human intervention. Since 2016, MT systems have primarily been based on AI in the form of large language models drawing on neural networks, and there is broad agreement that AI has caused a major leap forward in MT quality. Billions of words are translated by Google Translate every single day, and a multitude of other free online MT services are available. Thus, MT can be described as software that is everywhere or, in other words, as "everyware" (Cronin, 2010). MT is an example of the type of methods within Natural Language Generation (NLG) which are referred to as text-to-text generation (Gatt & Krahmer, 2018). ChatGPT is another example of text-to-text generation which covers "applications that take existing texts as their input, and automatically produce a new, coherent text as output" (Gatt & Krahmer, 2018, p. 65). Input texts might be brief prompts (e.g. "write a text about X") or texts that need to be processed in their entirety, for instance machine-translated, summarized or paraphrased.

Three recent reviews have examined research investigating the use of MT in foreign language learning (FLL) (Lee, 2023; Jolley & Maimone, 2022; Klimova et al., 2023). All three reviews conclude that MT has a positive impact on FLL. For instance, Klimova et al. found that MT is an efficient tool for developing both productive and receptive language skills and concluded that MT tools are "beneficial for FLL, the only question being the level of this beneficial impact" (Klimova et al., 2023, p. 677). However, the reviews also make it clear that teachers are still sceptical about using MT in their teaching, for instance because of limited trust in MT quality and in the effectiveness of using MT in FLL as well as a view of MT use as academic dishonesty or cheating. This often leads teachers to prohibit students from using MT and to teachers rarely addressing MT tools in their teaching in a constructive way. This picture is also clear in a Danish context, where a recent study of upper secondary language teachers' integration of digital technologies has shown that MT is only used to a limited extent (Caviglia et al., 2021). Specifically, Caviglia et al. found that only 7% of language teachers use MT in their teaching. 5% indicate that they are considering using MT, and 29% advise students against using it. 59% of teachers simply state that they do not use MT. The very limited integration of MT in upper secondary language teaching might also, at least partly, be a consequence of MT tools being prohibited at exams (Børne- og Undervisningsministeriet, 2023).

If upper secondary students are using MT tools although their teachers do not integrate it into their teaching and maybe even in spite of having been advised against it, it can be problematic. For instance, it can have negative implications for the student-teacher relationship, and students might be using MT tools in non-optimal and inappropriate ways which might hinder learning. Therefore, several scholars have argued that students should be trained in how MT works (e.g., O'Neill, 2019; Vold, 2018) equipping them with so-called *machine translation literacy* (Bowker, 2021a).

Surprisingly, we know very little about students' use of MT. In a university context, Dorst et al. (2022) have found that almost all Humanities students are very or extremely familiar with Google Translate, and that almost 70% use Google Translate regularly. However, to our knowledge, no studies have investigated the use of MT by upper secondary students. Hence, in this paper, we first present a survey of how upper secondary students use MT tools. This, among other things, shows a widespread use of MT tools. Against this backdrop, the purpose of the paper is to discuss current tendencies and future perspectives of using AI-powered language technologies in foreign language teaching.

In the next section, we present background literature on MT including different approaches and ways in which MT can be used. We then continue the background section with selected literature on AI and digital literacy. The background section is followed by a methods section where we describe how we conducted the survey. We then present the results and finally discuss the use of AI-powered technologies and the development of the necessary literacy in foreign language teaching.

Machine translation

MT is "a sub-field of computational linguistics (CL) or natural language processing (NLP) that investigates the use of software to translate text or speech from one natural language to another" (Liu & Zhang, 2015, p. 105). Within the field of NLP, it is, as mentioned above, an example of text-to-text generation within NLG. Around 2016, previous predominant approaches to MT, i.e. rule-based MT and statistical MT, were gradually replaced by so-called neural MT (NMT) drawing on AI. NMT has made it possible to train software to translate words and set of words in the context in which they are embedded instead of translating isolated words and recombining these in the target language (Pym, 2019), leading to increased output quality.

Apart from so-called "pull" situations where a user seeks out a machine-translated version of a source text him-/herself, e.g. through a free online MT service such as Google Translate, MT is also provided in many "push" situations where a translation is pushed to the user without him/her having actively requested one. This goes for e.g. automatic translations of user-generated content on social media such as Facebook and Instagram and on platforms such as Tripadvisor and Airbnb. This means that e.g. students may have a high consumption of machine-translated texts although they are not actively using MT tools.

Within human evaluation of MT quality, it is common to distinguish between the fluency and adequacy of machine-translated output, often a sentence (Koehn, 2010; O'Brien, 2017). Adequacy refers to "the meaning of the MT output in relation to the meaning of the source language segment" (O'Brien, 2017, p. 315), whereas fluency refers to the grammatical correctness of the MT output. With NMT, translations are often fluent, i.e. grammatically correct, which may lead a user to believe that they are also adequate renderings of the source text content. However, for example, parts of the source text content may have been left out, new content may have been added, and words may have been mistranslated (Klimova et al., 2023), leading to an inadequate rendering of the source text content. This has caused scholars to refer to NMT as deceivingly fluent (cf. e.g. Martindale and Carpuat, 2018). Other common problematic issues with machine-translated output include bias, formality levels, incoherence, and terminological inconsistency (Klimova et al., 2023).

If MT is used just to get the gist of the source text content, there is no need to make changes to the machine-translated text. It is simply used in its *raw* form (Way, 2013). This is typically the case with the use of MT for the translation of user-generated content. However, if MT is being used as a step in producing a text ready for publication, the MT output usually needs to be checked, and errors need to be corrected. This is an activity usually referred to as post-editing (O'Brien, 2022). Here, scholars typically distinguish between light and full post-editing. In light post-editing, focus is on ensuring that the translation is "an understandable reflection of the source-text content, but ignoring stylistic niceties" (Way, 2013, p. 4), whereas full post-editing involves producing a text "that is not only understandable, but also presented in a stylistically appropriate way" (Way, 2013, p. 4). This distinction is not unproblematic, but very common in the literature, and we will not delve further into the discussion here. Post-editing is a very common activity carried out by professional translators, and several studies on the integration of MT in FLL have focused on developing students' post-editing skills.

Digital literacies and learning

To be part of the 21st century civic life, it is important that students acquire the necessary skills to become knowledge creators rather than just passive consumers of information (Gretter and Yadav, 2016). This includes skills within creativity and critical analysis that help prepare students to both create and critically analyze digital materials. According to Gretter and Yadav (2016), a combination of computational thinking and media and information literacy (Grizzle et al., 2014) can help students achieve digital creativity and critical awareness in a globalized and hyper-connected world.

Media and information literacy (Wilson et al., 2013) is a term coined by UNESCO. Here, the need to build critical analytical skills with media and information consumers is emphasized. This has the purpose of empowering consumers and strengthening their critical sense and communication skills. This includes, among other things, knowledge of how to access and evaluate information, and understand how it can be used in an ethical manner, as well as being able to understand the media's role in our daily lives. Furthermore, it is important to be able to understand computer technologies and how the Internet influences the spread of information in a globalized and connected world (Grizzle et al., 2014). One of the widespread computer technologies that affect the information on the Internet today is MT where users often meet content that has been automatically translated ("push" situations) or use MT themselves ("pull" situations).

The actual use of online MT tools is very simple – it only requires writing or pasting a text into the source text window, choosing the languages involved and hitting the "translate" button. However, using such tools in an informed and critical way is less simple. Therefore, students in general (Bowker, 2021a) and language learners (Carré et al., 2022) should acquire literacy related to MT, not least because research has shown that students who are trained in how MT works perform significantly better when using MT for writing tasks (O'Neill, 2019). The fact that training matters is also emphasized by Vold who concludes that "training, scaffolding techniques and guidance from the teacher are of paramount importance" (Vold, 2018, p. 89) when integrating MT into foreign language teaching.

MT literacy is defined by O'Brien and Ehrensberger-Dow as "knowing how MT works, how it can be useful in a particular context, and what the implications are of using MT for specific communicative needs" (2020, p. 146). However, to our knowledge, no studies have developed didactic frameworks for the introduction of MT literacy in upper secondary education. In fact, studies conducted on the effectiveness of MT in language teaching and learning have predominantly focused on university students (see e.g. Lee, 2023; Klimova et al., 2023), and these studies have typically not included teacher interventions or student trainings for using MT but have rather used MT as an instrument.

In a university context, Bowker (2021b) has tested five different formats for MT literacy instruction to undergraduate students who are not studying to become language professionals. She suggests the following key elements of MT literacy instruction: (1) Understanding data-driven approaches to MT, (2) Transparency and MT use, (3) Risk assessment and MT, and (4) Interacting with MT. *Understanding data-driven approaches to MT* includes giving students a basic knowledge of how e.g. NMT works. This will make students understand why different MT tools may be more or less useful for different language pairs and text types, why different MT tools are likely to produce different results for the same source text, and why results might change from one translation of a source text to the next. Also, this element should make students aware of potential bias in MT output. The element *transparency and MT use* encourages students to be transparent about their use of MT, e.g. when citing material that has been translated. Also, students should be made aware that the use of MT "may be more or less appropriate depending on the learning objectives of the course" (Bowker, 2021b, p. 27) and on teacher preferences. Bowker also stresses that being transparent about the use of MT is important because it enables the readers of a machine-translated text to make a qualified decision as to how much to trust the content. The element *risk assessment and MT* has two overall dimensions: firstly, students should understand that the use of MT may carry a lower or higher risk depending on the use case. For instance, using MT to understand a friend's post on social media carries another risk than using MT to translate documents within legal or health care settings such as contracts or patient information leaflets. Secondly, students should learn that they should not enter sensitive information into a free online MT tool because data is kept and reused. Finally, *interacting with MT* includes facilitating that students get hands-on experience with improving MT quality, either by correcting the output (post-editing) or improving the source text (pre-editing).

The skills and competencies defined in relation to the development of media and information literacy and MT literacy are also important when evaluating text produced by generative AI tools such as ChatGPT. When working with this type of technology, a basic understanding of AI or so-called AI literacy is also required. Possessing AI literacy means to have the essential abilities that people need to live, learn, and work in our digital world with AI technologies (Ng et al., 2021).

Long and Magerko (2020) describe AI literacy in terms of 16 competencies needed to interact with and evaluate AI. These competencies include being able to recognize an AI as well as understanding how an AI works and what the strengths and weaknesses of an AI are. This includes competencies to identify different ethical issues such as privacy, transparency, diversity, bias, and accountability. Similarly, Ng and colleagues (2021) identify four aspects of AI literacy: know and understand, use and apply, evaluate and create, and ethical issues.

It can be difficult for people to spot AI technology when it is integrated into various applications. As an example, it is not always clear when something is written by an AI. This can lead to misunderstandings and make it more difficult to critically evaluate e.g. online content. It can also make interaction and collaboration more difficult if you are not aware that you are in fact interacting with an AI and not a living person (Long & Magerko 2020). Similarly, a lack of understanding of what AI is and how it works can lead to bad or incorrect use of the technology.

Research on how to teach and learn AI literacy is very sparse. A recent literature review on the topic (Ng et al., 2021) found 30 relevant papers. However, of these, many were preliminary results, and only eight were published in scientific journals. Among other things, Ng and colleagues (2021) pointed to a lack of research into how to use learning artefacts and integrate AI literacy effectively into K-12 classrooms to motivate students to learn about AI. It is also pointed out that teachers should update their AI knowledge to be able to offer better and more personalized learning to students.

Based on the above, it is clear that literature on MT literacy, media and information literacy, and AI literacy share many similar elements. However, specifically in relation to AI-powered language technologies in upper secondary language learning, our knowledge of necessary literacy elements is very limited. Therefore, in the discussion section, drawing on the insights from a survey of upper secondary students' use of MT and on this section on digital literacies, we will discuss the development of the necessary literacy in language teaching.

Methods

The empirical data for this article were collected through a survey conducted in August/September 2021 at different Danish upper secondary education institutions and in different language subjects. At two institutions, students attending The Higher Technical Examination Programme (HTX) were invited to join the study, and at one other institution, students attending The Higher Commercial Examination Programme (HHX) were invited. In the survey the students were asked questions about their use of MT, how often they use it and for what purposes. The students were asked to provide their study year and studied language. Other than that, no background information was requested.

The survey was created in SurveyXact, a software used to create and distribute questionnaire-based surveys. A link to the survey was distributed to students through their teachers in the specific language subjects. The fact that the survey was distributed to students through their teachers has some potential drawbacks. For instance, the teachers might not have given similar instructions, and the fact alone that teachers distributed the survey might have introduced a bias due to the attitudes of the teachers towards MT as perceived by the students. However, we attempted to address this by stressing in the introductory survey text that the data would be treated anonymously, and that it was important to provide honest answers even if the student had the perception that his/her teacher did not like MT. As a supplement to the introductory text, teachers were also provided with instructions which were used to orally prepare the students before filling out the survey. The collected data include responses from 55 second year HTX students in the subject English, 56 second year HHX students in the subject German and 36 HHX students (20 second year and 16 third year students) in the subject Spanish.

The 55 responses from English students represent three classes taught by two different teachers. The 56 German students represent two classes taught by the same teacher, and the 36 Spanish students represent two classes taught by two different teachers. Since we do not know the complete number of students in each class, unfortunately, we are not able to provide an exact response rate, however, we do know that classes are not allowed to surpass a class size ratio of 28 students. Thus, we can conclude that the response rate is relatively high, especially for German students. With the given sample size, an experimental error of a certain size is to be expected, especially when comparing the different language classes. Since only three upper secondary institutions and a limited number of classes and teachers were involved in the survey, the generalisability of the results is also limited. Further, since we did not include demographic background questions, we are not able to say anything about the differences between the use of MT by, for instance, different genders. However, the survey should be able to give us indications of the overall tendencies and attitudes towards MT.

Upper secondary students' use of machine translation

As is evident from Figure 1, 95% of English students (n=55) indicated that they had tried using MT. This number was 98% for German students (n=56) and 92% for Spanish students (n=36). In terms of frequency of use, as shown in Figure 2, 38% of English students answered that they use MT every week, 33% every month, and 29% answered less than once a month. For German students, this was 73%, 22% and 5%, and for Spanish students 67%, 27% and 6%, respectively. Thus, although almost all students have tried using MT, we see a more pronounced use of MT with German and Spanish students. Since in a Danish context, German and Spanish are typically the students' third languages (L3), a plausible explanation for this is a lower L3 proficiency compared to the L2 (typically English) and thus a greater need for assistance both during productive and receptive tasks.

Figure 1.
Percentage of students who have tried using machine translation.

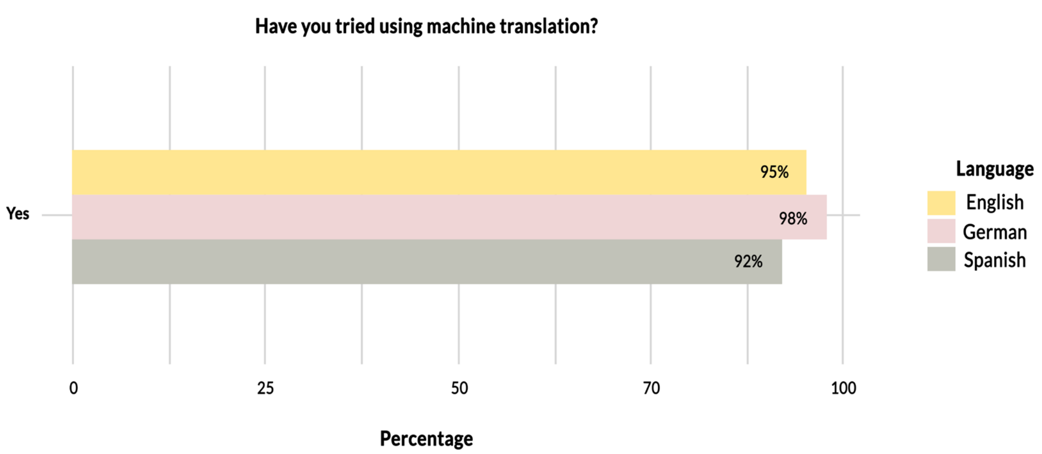
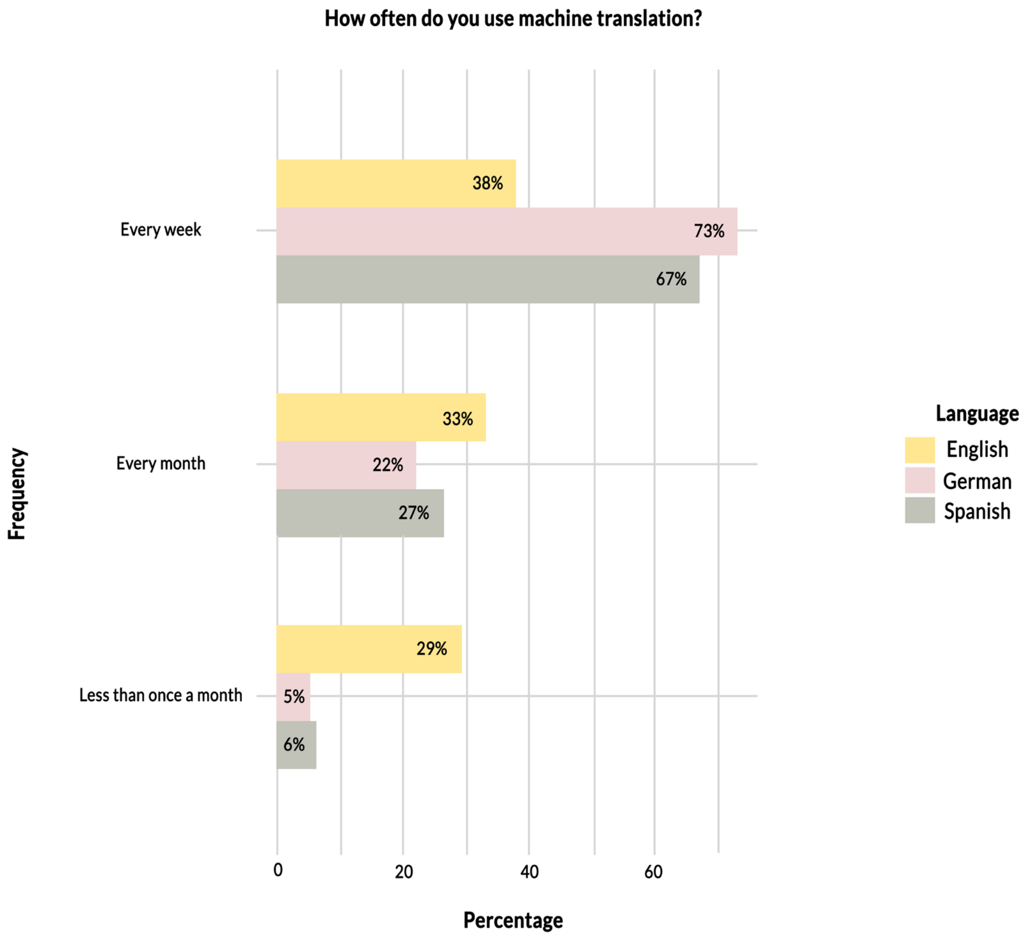
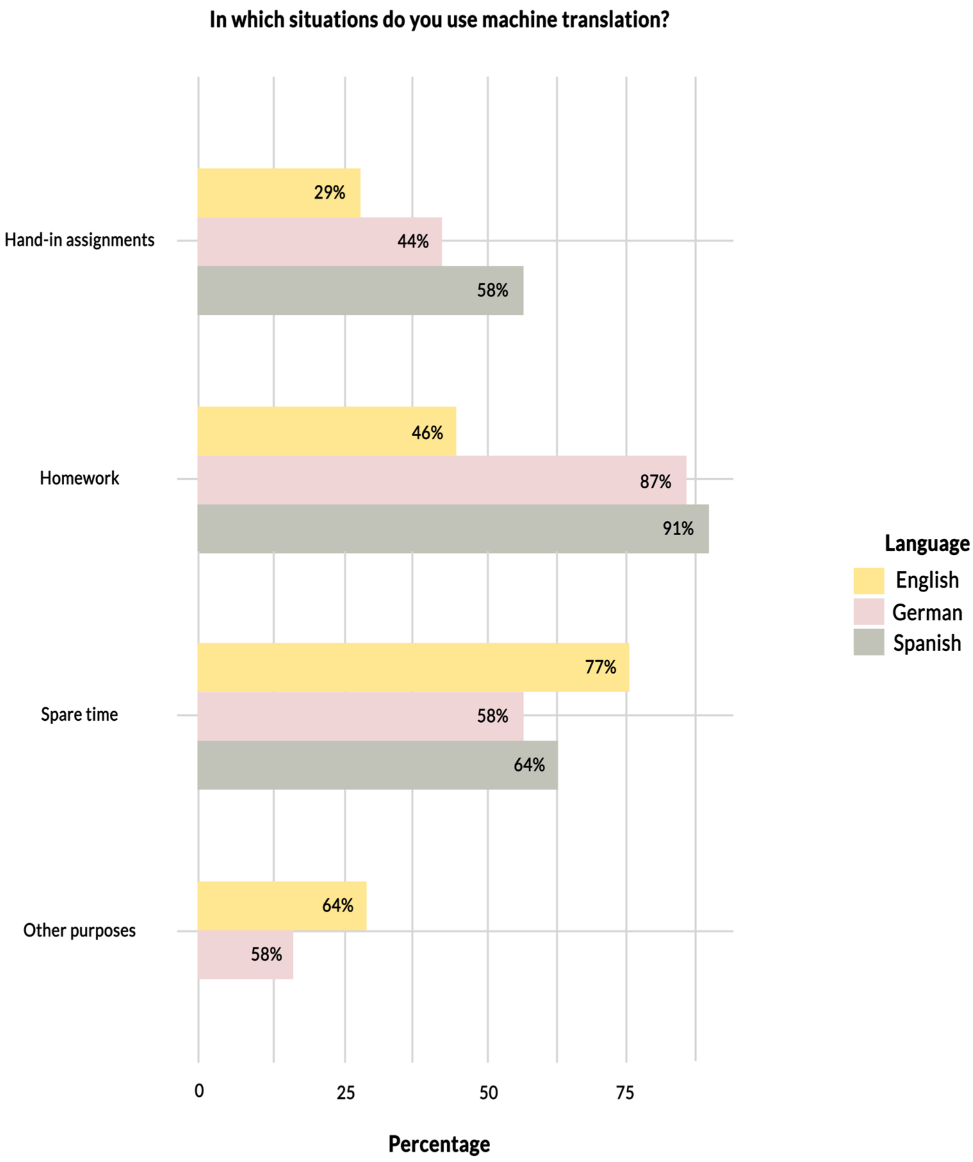


Figure 2.
Frequency of machine translation use.



As shown by Figure 3, when asked about the situations in which the students use MT, 29% of English students, 44% of German students and 58% of Spanish students answered that they use MT when preparing hand-in assignments to be evaluated by the teacher. 46% of English students, 87% of German and 91% of Spanish students responded that they use MT when doing their homework. 77% of English, 58% of German and 64% of Spanish students indicated that they use MT in their spare time. Finally, 29% of English, 16% of German and 0% of Spanish students used MT for other purposes. It is interesting to note here that German and Spanish students use MT tools for hand-in assignments and homework much more often than English students, whereas the use of MT in the spare time was more frequent for English students. Other purposes included the use of MT to check a sentence before saying something orally in class and to translate words from languages other than the ones the students were studying (such as Finnish or Romanian).

Figure 3.
Situations in which machine translation is used.



The students were also asked whether they use MT to translate single words, single sentences and/or whole paragraphs/texts. Here, 85% of English, 65% of German and 79% of Spanish students answered single words. 50% of English students used MT for translating single sentences; this was the case for 82% of German and 76% of Spanish students. Finally, 8% of English, 40% of German and 36% of Spanish students used MT to translate paragraphs/texts. These results show us that students from all language subjects most commonly use MT for the translation of single words or single sentences. The widespread use for single words might indicate that students use MT as an alternative to a bilingual dictionary; something that was also observed by e.g. Dorst et al. (2022). However, since NMT draws on context, this is not a recommended way of using MT (e.g. Ducar & Schocket 2018). Interestingly, the rather large number of German and Spanish students using MT to translate whole paragraphs or texts might indicate that students in L3 subjects use MT more frequently to assist with receptive tasks than students in L2 subjects.

Next, the students were to answer a question about how high or low they assessed the quality of MT output to be. Here, the answers were very similar. Between 44% and 48% of students neither assessed quality to be high or low, around 20% either found the quality high or low, and 0-5% found the quality very high or very low. Also, between 5% and 8% stated that it was difficult for them to assess. A possible explanation for the high number of students who either could not assess MT quality or judged it to be neither high nor low might be that it is difficult for upper secondary students to evaluate MT quality due to their language proficiency level. This would be in line with Chung's study (2020) of university students which found that more advanced students are more critical of MT quality than students with lower language proficiency. Another potential explanation is that the students have not received any instructions as to how to evaluate MT output, e.g. an introduction to the concepts of fluency and adequacy.

When asked whether they know their teacher's attitude towards MT, 60% of English, 95% of German and 73% of Spanish students answered "yes". Hence, most students believe to know the teacher's attitude. Again, this is more pronounced in L3 subjects. When they were asked to describe the attitude, there were different types of answers. A lot of students simply stated that the teacher did not like it (e.g. "She hates it" or "She is not a fan") or that MT should not be used (e.g. "She does not want us to use it" or "It is stupid to use it"). Other students specifically mentioned that their teacher did not want them to use MT tools such as Google Translate, but instead told them to use a dictionary (e.g. "She does not like Google Translate, but the dictionary is okay"). Some students also mentioned that the teacher thinks they are cheating themselves when they are using MT (e.g. “You are cheating yourself”). A low number of answers indicated that the teacher allowed the use of MT for small linguistic units or for certain tasks (e.g. “For single words it is fine – otherwise it is a ‘no thank you’” or “We can use it to understand things, but she doesn’t want us to use it for hand-in assignments since that would be a waste of our and her time”). Interestingly, one student pointed out that the teacher did not want the students to use MT and thinks she is catching them in using it even when they are not. This is an indication that prohibiting MT might lead to an atmosphere of suspicion in the FLL classroom. The general perception of students that the teachers do not like MT seems to correspond to the results by Caviglia et al. (2021) showing a very limited integration of MT in Danish upper secondary language learning.

Discussion

95% of all students in our survey indicated that they have tried using MT, and 57% use MT every week. These results point towards a regular use of MT tools among the students from all language subjects. The results also indicate that MT is primarily used for the translation of single words and sentences, and that it is used for a range of tasks, i.e. both for preparing hand-in assignments, homework and during the students’ spare time. Further, the findings showed that German and Spanish students use MT for school-related purposes more often than English students, and that English students are more inclined to use it in their spare time. Nonetheless, the widespread use of MT by students for educational purposes stands in sharp contrast to the teachers’ general reluctance and limited integration of MT as shown in the students’ statements in our survey and in (Caviglia et al., 2021; Lee, 2023; Jolley & Maimone, 2022; Klimova et al., 2023).

The widespread use of MT particularly by German and Spanish students might pose an issue since we anticipate that students in L3 subjects may find it more difficult to evaluate and identify errors in MT output, as they typically have fewer years of language learning compared to English students. This may especially be true for aspects such as fluency and adequacy, which require a nuanced understanding of the target language. However, our data do not deliver insights into whether L3 students used MT in a less critical way than L2 students, and we have not encountered any other studies comparing the use of MT in L2 and L3 subjects. Also, previous studies have shown in-consistent findings in terms of the influence of L2 proficiency on MT use (Lee, 2023). The quite widespread use of MT for preparing hand-in assignments is also noteworthy. When MT is prohibited by teachers, this is done secretly, and thus, it is difficult, if not impossible, for the teachers to know what the students have done themselves and what has been machine-translated. In this context, it is also interesting that with increasing MT quality, it is more often signs of untypically good text production that signal MT use rather than weaknesses in students' assignments (Jolley & Maimone, 2022; Ducar & Schocket, 2018). For example, the signs could be linguistic forms that students have not yet been introduced to. In other words, nowadays, teachers may detect MT use because of unusually high rather than because of low quality. Of course, from a teacher's point of view, it seems fruitless and maybe even frustrating to be evaluating an assignment produced entirely by MT, and an excessive and uncritical use of MT might lead to students not learning much. In such situations, MT can become a shortcut in the sense that the student simply "outsources" work (Dalsgaard et al., 2022) to the technology to avoid doing the work him- or herself. These are some of the understandable reasons for teachers' opposition to MT and their reaction of trying to prohibit MT use in language learning.

However, the lack of integration of omnipresent technologies such as MT as well as prohibiting students from using MT may not only create an atmosphere of suspicion in the language classroom, but also create a situation where students use the technology without any didactic introduction. They are thus left to themselves in terms of developing appropriate strategies when using MT. This could lead to students using MT in non-optimal ways such as using MT to translate single words which was frequent in our survey. In order to enable students to use MT tools as a "cognitive partner" where the technology is used productively and supports learning (Dalsgaard et al., 2022), they need to gain the necessary literacy. Along the same lines, Vold (2018) for instance emphasises the importance of teacher's guidance when learning to use MT, and O'Neill (2019) points out that learners who are trained, even briefly, in how MT works, write better compositions than those with no training.

Thus, teaching basic instructions on how to use MT tools seems to improve the students' use of MT, but it remains uncertain whether such instructions are sufficient considering the widespread use of MT in today's context. With new tools and technologies such as ChatGPT, which can also be used for MT, we expect the development to continue in the direction of more AI-generated text and not less. Thus, we will gradually come across more and more content that is generated or translated by a machine. For example, this may happen when we browse content on social media and when we read articles posted online. While it is often declared if a text is generated or translated by a machine, this is not always the case. Furthermore, given the speed at which we typically consume online content, we may not recognize machine-generated text as such. Also, AI-powered language tools are increasingly being integrated into various applications such as the Microsoft 365 suite (Microsoft, 2023) where we can use them to assist our own text production. The need to learn more about AI-generated content must therefore be seen both from a content consumer and a content creator perspective. Hence, we argue for a need to teach students competencies that can guide them with good practices and ethical considerations when they actively use MT and text generation technologies (i.e. in "pull" situations") as well as enable them to critically evaluate translated or generated content (in both "pull" and "push" situations). Since we see a fusion of various technologies in today's applications, we will draw on literature within the fields of both MT literacy, AI literacy, and media and information literacy to identify or fuse relevant competencies and practices that need to be taught to meet the beforementioned dualistic requirements for AI-powered language technologies.

According to UNESCO's report (Grizzle et al., 2014), critical analytical and critical awareness skills are important parts of developing media and information literacy. Akin to the importance of being able to critically analyze information shared on the Internet, it will, in the immediate future, also be important to be able to critically assess if the author is a human or a machine and furthermore be able to assess the authenticity of generated content. The ability to recognize that you are interacting with an AI is a competency described within AI literacy (Ng et al., 2021; Long and Magerko, 2020). In general, we expect to be seeing a future where work tasks and cognition are distributed between people and AI-based technologies to a high degree which requires people to be able to understand, assess, collaborate, and interact with these technologies.

Many of the aspects or competencies of AI literacy (Ng et al., 2021; Long and Magerko, 2020) overlap with the competencies described in MT literacy (Bowker, 2021b). The recurring themes seem to be the significance of possessing a fundamental understanding of the workings, strengths and weaknesses of the technology, competencies for proper usage and evaluation of AI-generated text, and knowledge of the precautions to take when the technologies are being utilized. Also, different aspects of ethical use of AI are underlined, including transparency of use and privacy issues. Within the context of upper secondary education, together with the ability to recognize AI-generated text, these elements seem to be indispensable, both in terms of language learning, but also in relation to the students' general digital literacy. This should, for example, equip students with an understanding of AI-powered technologies enabling them to critically analyze text content produced by these and to use the technologies to create content themselves, e.g. drawing on post-editing. This includes knowing how and when to use them, being able to determine the risk of using them for a given task and evaluate the content from e.g. linguistic, pragmatic and cultural perspectives.

Apart from what the literature has pointed out, there are some aspects that seem particularly relevant to MT, on the one hand, where the goal is to reproduce the entire content of a source text in another language and, on the other hand, tools such as ChatGPT that can produce new text based on brief prompts. In relation to the former, while the concepts of adequacy and fluency are important, and students should know that machine-translated text may be deceivingly fluent, it is also important that students understand that a good translation is not necessarily just an adequate and fluent rendering of the source text content. Rather, due to culture-specific differences, translations should often be adapted to conform to the norms and conventions of the target culture, "taking into account what target-culture members can be expected to know or feel about the subject in question" (Nord, 1997, p. 46). Thus, a 1:1 fluent machine translation is not necessarily a good translation. In relation to the latter, it seems important that students learn to evaluate the authenticity of the generated text. Scholars (e.g. Dale, 2021) and tool developers themselves have emphasized that language models might generate output that is plausible-sounding, but not consonant with the truth, a phenomenon often referred to as "hallucination" (OpenAI, 2023b). Therefore, users must consider whether they can reasonably assume that generated text is correct and must develop skills in terms of verifying or disproving generated content. Also, such tools are sensitive to the phrasing of the prompts and may, based on one phrasing, claim to not know the answer to a question, but then answer correctly after a slight rephrase. This has given rise to a relatively new phenomenon referred to as "prompt engineering" which involves developing and optimizing prompts in order to use AI models efficiently. In our view, in upper secondary education, students should gain at least a basic awareness of the sensitivity of AI tools to prompt phrasings.

If students are to acquire competencies related to AI-powered language technologies, it is a prerequisite that the teachers' literacy is also developed. In order for them to integrate AI-powered language technologies in FLL, they need not only to gain the same competencies as the students; they also need knowledge on how AI can enhance FLL.

In line with this, the European Commission's (2023) European Digital Education Hub argues that knowledge of how to engage confidently, critically and safely with AI systems and knowledge of technical AI basics are pre-requisites for teaching with AI, i.e. for the application of AI in learning contexts. We argue that teaching with AI in FLL requires didactic knowledge of how AI can support students' use of the technologies as "cognitive partners" in FLL. In this context, Dalsgaard et al. (2023) mention MT systems and AI chatbots as useful tools to explore the foreign language, i.e. to find ways to express oneself in the foreign language, to find solutions to linguistic problems and to produce text. While language teachers' level of literacy as to AI-powered language technologies has not been explored, teachers' general lack of digital competencies such as computational thinking is well documented, and we assume this is also the case for AI literacy in FLL. Initiatives to support teachers in building this literacy would be welcome.

With the current rapid development in AI-powered language technologies and the increasing integration of these technologies into various platforms, it is imperative that research is conducted on students' use of and attitudes towards the tools. Also, we need research exploring different approaches for teaching and learning about the technologies and building the necessary literacies.

With the widespread use of MT by students shown by this study, it seems plausible that students will also, at least gradually, adopt tools such as ChatGPT into their set of digital resources. Thus, although MT tools have already been changing the conditions for upper secondary language teaching and learning for many years, the impact of AI-powered language technologies will probably only increase from here on. In our view, it is of vital importance that we aim to narrow the gap between students' and teachers' practices that has characterized FLL for several years, namely a situation where, roughly speaking, all students are using the tools, and teachers are reacting by prohibiting them. This requires integration of language technologies into FLL and with this integration a legitimisation of a practice that is already widespread. However, at the same time, we should avoid a situation where students simply outsource their work to technology. A possible solution might be to distinguish between "learning contexts" and "performance contexts" (Dalsgaard et al., 2022) to a greater degree. In learning contexts, students can get help and support by entering into a partnership with digital tools, and conversely, in performance contexts, they are expected to have the ability to independently express themselves in the foreign language. Dalsgaard et al. (2022) seem to equate performance contexts with exams where MT tools are currently prohibited, however, there seems to be a potential in also distinguishing between learning and performance contexts in the daily teaching. For instance, Brunø (2023) and colleagues have experimented with letting the students' preparation time at home being a pure learning context without any following assessment by the teacher. During this preparation students have all digital resources at their disposal. This is then followed by tests in the classroom where the students complete assignments directly related to their preparations without having access to digital resources. The aim of Brunø (2023) and colleagues is to give the students the experience that the preparation work pays off in the performance context.

Conclusion and future perspectives

Based on our study, we see a widespread use of AI-powered technologies by students in FLL. However, many educators are sceptical about using these technologies in language teaching and resort to prohibiting them. This is partly due to concerns about whether using the beforementioned technologies will hinder students' language learning and whether the technologies will be abused, e.g. for cheating. This is understandable, however, at the same time, it poses a problem as the students do not develop the necessary literacy and competencies required to use the technologies appropriately, leading to a potential misuse. Furthermore, the students are not prepared for encountering content created by AI-powered technologies which is proliferating with the increase in the use of AI-based solutions in society.

To keep up with the rapid development of new technologies, we recommend teaching students basic digital literacy and competencies to prepare them for AI-powered technologies, and in this paper, based on digital literacy theory, we have suggested elements that should be included. This is not only essential to enable students to use these technologies in effective and appropriate ways, but also to make them able to critically evaluate AI-generated content. Furthermore, given the limited research in this area, there is a need to explore the integration of these technologies into foreign language teaching in ways that scaffold the development of said competencies and that support students' language learning, i.e., in ways that are in line with theories on language acquisition.

Referencer

Bowker, L. (2021a). Promoting Linguistic Diversity and Inclusion: Incorporating Machine Translation Literacy into Information Literacy Instruction for Undergraduate Students. *The International Journal of Information, Diversity, & Inclusion*, 5(3), 127-151. https://doi.org/10.33137/ijidi.v5i3.36159

Bowker, L. (2021b). Machine translation use outside the language industries: a comparison of five delivery formats for machine translation literacy instructions. *Translation and Interpreting Technology Online*, 25-36. https://doi.org/10.26615/978-954-452-071-7_004

Brunø, L. (2023). Sådan bliver rettearbejdet igen meningsfuldt: Evaluering af skriftlig tysk i gymnasiet. *Sproglæreren*, 1/2023, 20-23.

Børne- og Undervisningsministeriet (2023). Hjælpemiddeloversigt. *Accessed* [30 March, 2023] at: https://www.uvm.dk/-/media/filer/uvm/udd/gym/xls22/maj/220506-hjaelpemidler.xlsx

Carré, A., Kenny, D., Rossi, C., Sánchez-Gijón, P. & Torres-Hostench, O. (2022). Machine translation for language learners. In: Kenny, D. (ed.), *Machine translation for everyone: Empowering users in the age of artificial intelligence*, 187-207. Berlin: Language Science Press. https://doi.org/10.5281/zenodo.6760024

Caviglia, F., Dalsgaard, C., Boie, M. A. K. & Thomsen, M. B. (2021). *Undersøgelse af digitale teknologier i fremmedsprogsundervisningen: Sproglæreres anvendelse af digitale teknologier i gymnasiet.* DPU – Danmarks Institut for Pædagogik og Uddannelse, Aarhus Universitet. Accessed [7 April, 2023] at: https://pure.au.dk/portal/files/213331375/FREMDitSPROG_sporgeskemaundersogelse_marts2021.pdf

Chung, E. S. (2020). The effect of L2 proficiency on post-editing machine translated texts. *Journal of Asia TEFL*, *17*(1), 182-193. https://doi.org/10.18823/asiatefl.2020.17.1.11.182

Cronin, M. (2010). The Translation Crowd. *Revista Tradumàtica: traducció i tecnologies de la informació i la comunicació*, 8, 1-7. https://doi.org/10.5565/rev/tradumatica.100

Dale, R. (2021). GPT-3 – What's it good for? *Natural Language Engineering, 27*, 113-118. https://doi.org/10.1017/S1351324920000601

Dalsgaard, C., Boie, M. A. K. & Caviglia, F. (2022). *Frem dit sprog: Digitale værktøjer i sprogfagene.* DPU – Danmarks Institut for Pædagogik og Uddannelse, Aarhus Universitet. Accessed [7 April, 2023] at: https://pure.au.dk/portal/files/279093397/Fremditsprog_final_online.pdf

Dalsgaard, C., Christensen, M. V., Caviglia, F., Andersen, M. S., Kjærgaard, H. W. & Boie, M. A. K. (2023). Sprogteknologier i fremmedsprogsundervisningen: indkredsning af en sprogfaglig teknologiforståelse. *Tidsskriftet Læring og Medier*, 28, 1-23.

Dorst, A. G., Santos Ângelo Salgado Valdez, S., & Bouman, H. M. C. (2022). Machine translation in the multilingual classroom: how, when and why do humanities students at a Dutch university use machine translation? *Translation And Translanguaging In Multilingual Contexts*, *8*(1), 49-66. https://doi.org/10.1075/ttmc.00080.dor

Ducar, C. & Schocket, D. H. (2018). Machine translation and the L2 classroom: Pedagogical perspectives for making peace with Google translate. *Foreign Language Annals, 51*, 779-795. https://doi.org/10.1111/flan.12366
European Commission (2023). *Use Scenarios & Practical Examples of AI Use in Education: Briefing report No. 3 by the European Digital Education Hub's squad on artificial intelligence in education*. Accessed [16 August, 2023] at: https://education.ec.europa.eu/news/the-european-digital-education-hubs-monthly-theme-artificial-intelligence
Gatt, A. & Krahmer, B. (2018). Survey of the State of the Art in Natural Language Generation: Core tasks, applications and evaluation. *Journal of Artificial Intelligence Research*, 61, 65-170. https://doi.org/10.1613/jair.5477
Gretter, S. & Yadav, A. (2016). Computational thinking and media & information literacy: An integrated approach to teaching twenty-first century skills. *TechTrends*, 60, 510-516. https://doi.org/10.1007/s11528-016-0098-4
González-Pérez, L. I., & Ramírez-Montoya, M. S. (2022). Components of Education 4.0 in 21st century skills frameworks: systematic review. *Sustainability*, *14*(3), 1-31.
Griffin, P. & Care, E. (Eds.). (2014). *Assessment and teaching of 21st century skills: Methods and approach*. Springer.
Grizzle, A., Moore, P., Dezuanni, M., Asthana, S., Wilson, C., Banda, F. & Onumah, C. (2014). *Media and information literacy: policy and strategy guidelines*. 1-195, UNESCO.
Jolley, J. R. & Maimone, L. (2022). Thirty Years of Machine Translation in Language Teaching and Learning: A Review of the Literature. *L2 Journal, 14*(1), 26-44. https://doi.org/10.5070/L214151760
Klimova, B., Pikhart, M., Delorme Benites, A., Lehr, C. & Sanchez-Stockhammer, C. (2023). Neural machine translation in foreign language teaching and learning: a systematic review. *Education and Information Technologies, 28*(1), 663-682. https://doi.org/10.1007/s10639-022-11194-2
Koehn, P. (2010). *Statistical Machine Translation*. Cambridge University Press.
Lee, S. M. (2023). The effectiveness of machine translation in foreign language education: a systematic review and meta-analysis. *Computer Assisted Language Learning, 36*(1-2), 103-125. https://doi.org/10.1080/09588221.2021.1901745
Liu, D. & Zhang, X. (2015). Machine translation: general. In: Sin-wai, C. (ed.), *The Routledge Encyclopedia of Translation Technology*, 105-119. Routledge.
Long, D. & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In: *Proceedings of the 2020 CHI conference on human factors in computing systems*, 1-16. https://doi.org/10.1145/3313831.3376727
Martindale, M. J. & Carpuat, M. (2018). Fluency over adequacy: a pilot study in measuring user trust in imperfect MT. *Proceedings of AMTA 2018*, 17-25.
Microsoft (2023). *Introducing Microsft 365 Copilot – your copilot for work*. Accessed [April 2, 2023] at https://blogs.microsoft.com/blog/2023/03/16/introducing-microsoft-365-copilot-your-copilot-for-work/
Ng, D. T. K., Leung, J. K. L., Chu, S. K. W. & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence, 2*. https://doi.org/10.1016/j.caeai.2021.100041

Nord, C. (1997). A Functional Typology of Translations. In: A. Trosborg (ed.), *Text Typology and Translation*, 43-66. John Benjamins Publishing Company. https://doi.org/10.1075/btl.26.05nor

O'Brien, S. (2017). Machine Translation and Cognition. In: Schwieter, J. W. & Ferreira, A. (eds.), *The Handbook of Translation and Cognition*, 313-331. John Wiley & Sons, Inc. https://doi.org/10.1002/9781119241485.ch17

O'Brien, S. (2022). How to deal with errors in machine translation: Post-editing. In: Kenny, D. (ed.), *Machine translation for everyone: Empowering users in the age of artificial intelligence*, 105-120. Berlin: Language Science Press. https://doi.org/10.5281/zenodo.6759982

O'Brien, S. & Ehrensberger-Dow, M. (2020). MT Literacy – A cognitive view. *Translation, Cognition and Behavior, 3*(2), 145-164. https://doi.org/10.1075/tcb.00038.obr

O'Neill, E. M. (2019). Training students to use online translators and dictionaries: The impact on second language writing scores. *International Journal of Research Studies in Language Learning, 8*(2), 47-65. https://doi.org/10.5861/ijrsll.2019.4002

OpenAI (2023a). *Introducing ChatGPT.* Accessed [March 29, 2023] at https://openai.com/blog/chatgpt

OpenAI (2023b). *GPT-4 System Card.* Accessed [April 9, 2023] at https://cdn.openai.com/papers/gpt-4-system-card.pdf

Pym, A. (2019). *How automation through neural machine translation might change the skill sets of translators. Translation in the digital era.* Kuala Lumpur: Universiti Sains Islam Malaysia, 20-22 August 2019. Unpublished manuscript. Accessed [March 11, 2020] at https://www.academia.edu/40200406/How_automation_through_neural_machine_translation_might_change_the_skill_sets_of_translators

Schwab K. (2017). *The fourth industrial revolution.* Currency.

Vold, E. T. (2018). Using Machine-Translated Texts to Generate L3 Learners' Metalinguistic Talk. In: Haukås, A., Bjørke, C. & Dypedahl, M. (eds.), *Metacognition in language learning and teaching,* 67-97. Routledge. https://doi.org/10.4324/9781351049146-5

Way, A. (2013). Traditional and Emerging Use-Cases for Machine Translation. *Proceedings of Translating and the Computer 35,* 1-12.

Wilson, C., Grizzle, A., Tuazon, R., Akyempong, K. & Cheung, C. K. (2014). *Media and information literacy curriculum for teachers* (1-17). UNESCO Publishing.

Abstract

Adaptive learning technology has the potential to tailor learning to suit individual students' needs, desires and competence level. However, qualitative close up studies of students' interaction with adaptive technology are rare. This study explores 4th grade mathematics students' use of adaptive learning technology through screen recordings supplemented with analysis of data generated and tagged by the adaptive engine. The study explores how different types of students, i.e. students with varying mathematical competence levels and motivation towards mathematics, interact with an adaptive learning material, what learning paths emerge for different types of students in their interaction with the learning resource and how student self-efficacy is affected by the interaction. The study shows that students neglect to access and utilize the supportive resources in the learning tool. Rhapsode worked poorly for a student with a combination of low mathematical competence and motivation for mathematics but fairly well for the other three students.

Adaptiv læringsteknologi har potentiale til at skræddersy læring, så den passer til den enkelte elevs behov, ønsker og kompetenceniveau. Kvalitative nærstudier af elevers interaktion med adaptiv teknologi er dog sjældne. Denne undersøgelse udforsker 4. klasses matematikelevers brug af adaptiv læringsteknologi gennem skærmoptagelser suppleret med analyse af data genereret og tagget af den adaptive motor. Undersøgelsen undersøger, hvordan forskellige typer af elever, det vil sige elever med varierende matematiske kompetenceniveauer og motivation i forhold til matematikfaget, interagerer med et adaptivt læremiddel, hvilke læringsveje der opstår for forskellige typer af elever i deres interaktion med den adaptive læringsressource og hvordan elevens self-efficacy påvirkes af interaktionen. Implikationer for både fremtidig design og læreres brug af adaptive læremidler diskuteres.

Between the clicks

Student learning paths when interacting with an adaptive learning resource in 4th grade mathematics

Stig Toke Gissel, UCL Erhvervsakademi og Professionshøjskole, & Rasmus Leth Jørnø, UCL Erhvervsakademi og Professionshøjskole

Adaptive learning, pedagogy and student interaction

A defining characteristic of adaptive learning technology is the dynamic change in system response based on user interaction (Liu et al., 2017). This feature of adaptive learning technology has the potential to personalize learning environments and differentiate challenges and support which, in turn, can lead to strengthening of student competence, motivation and confidence (Pollard & James, 2004; U.S. Department of Education, 2017). The idea that content, learning context, pedagogical approach etc. could be tailored to suit every students' needs and interests holds great appeal and promise.

Paradoxically, research in adaptive technology rarely focuses on observations of individual students, but rather on test data and survey responses (Martin et al., 2020). If we want to understand which groups of students can benefit from using adaptive learning technology and what their interaction can tell us in relation to designing and optimizing adaptive learning designs, a reasonable approach would be to take a closer look at what students actually do when interacting with an adaptive technology.

The overall effects of using ICT in school are ambiguous (Bulman & Fairlie, 2016; OECD, 2015). However, Tamim et al (2011) found multiple studies showing positive effects of ICT-usage and concluded that results depend on contextual factors such as pedagogy, content, and teacher competence. Similarly, Gericke al (2014) concluded that positive and negative effects depend on teaching method and context. One consequence of these findings could be that we need to study use of digital technology for specific educational purposes in rather specific contexts instead of expecting ICT in general to have a positive impact.

However, not all educational technologies are designed to involve teachers in professional judgement and adaptation, and some products may even be designed to offset the role of contextual factors such as heterogeneous student proficiency levels and teacher beliefs. Adaptive technologies are examples of technologies that potentially make teacher intervention superfluous vis-à-vis use of the learning resource (Apoki et al., 2022). Adaptive learning technologies are designed to tailor the learning experience to individual students' learning needs by adapting learning paths based on tracking of students' interaction and input (Somyürek, 2015). Adaptive technologies aspire towards optimal personalized learning experiences by providing immediate and relevant assistance, resources and feedback (Kerr, 2016; Walkington, 2013). Hence, adaptive technologies have the potential to act as a digital *tutor* (Taylor, 1980), i.e. take care of core didactic functions, such as selecting content, presenting content in an optimal fashion, evaluation, differentiation etc.

Research in adaptive learning technologies is typically preoccupied with optimizing the technology to maximize learning gains for the users (e.g. Conejo et al., 2004; Guzmán et al, 2007; Liu et al., 2017; Peng et al., 2019; Tai et al., 2001). Studies about use and outcome of adaptive learning resources in elementary school are limited both in number and robustness (Holmes et al., 2018). Much research in the field is explorative and small scale (Verdu et al., 2008). The vast majority of studies are with students in higher education (Johnson & Samora, 2016; Xie, et al., 2019). It is not surprising, that trials have mainly been conducted with students on higher educational levels, as we could expect the outcome of prolonged individual interaction with a digital adaptive learning resource to depend on levels of self-discipline and meta-cognitive abilities that few primary schools students can fulfill. Overall, these studies indicate positive learning outcomes, but highlight the importance of the facilitation and instruction of educators (Kulik & Fletcher, 2016; Verdu et al., 2008, Wang et al., 2020).

Du Boulay (2019) distinguishes between screen level pedagogy and classroom level pedagogy in relation to adaptive learning. Screen level pedagogy encompasses the interaction between the individual student and the adaptive learning resource which, in turn, depends on the pedagogical approach embedded in the learning resource. Classroom pedagogy is the teacher's realm; the teachers connect and integrate the activities and content from the adaptive learning resource in their broader educational design. The teacher, however, is to some degree left out of the screen level learning situation because the adaptivity depends on interaction between the individual student and the adaptive learning resource.

A recent study shows, that teachers experience that the individualization connected with using adaptive learning resources challenges established classroom norms of collectivity, and that teachers feel a lack of control and insight vis-à-vis their students' learning process (Modén, 2021). Therefore, it would seem relevant to study the interaction between the individual student and the adaptive learning resource to gain insights into what goes on when a student interacts with an adaptive learning resource and to focus research on the student as the key to positive outcomes of using adaptive technologies.

In a literature review, Nakic et al. (2015) investigated the various variables and student characteristics that adaptive technologies should employ to create valid *user models* (also known as learner models), i.e. a set of assumptions about the user inferred from user interaction with the learning resource that guide what is presented for the user. It was found that cognitive ability and personality are the user characteristics that attract attention from most researchers, but that research in the last two decades is increasingly preoccupied with noncognitive characteristics such as emotional, motivational and meta-cognitive factors, in user models.

Learner models are typically constructed based on a combination of system identification of relatively stable user characteristics such as learning style and more variable factors such as learning history and student knowledge (Nguyen, 2015). Data generated by user interaction with the adaptive learning resource and other data need to be interpreted and, in turn, translated into system action. However, the data points used by the system to create and adjust the learner model will in themselves be selected proxies of relevant variables. For example, average user time spent on task items can provide indications of user domain specific ability, motivation, self-efficacy etc. System interpretation of singular user actions is uncertain. Prolonged time spent on a task can be caused by the student's thoroughness, uncertainty, inactivity or distraction, that the student finds the item challenging etc. In essence, the learner model will be a rudimentary image of the student, although perhaps sufficient for the adaptive purposes at hand.

In this study, we compare the data provided by the adaptive learning resource, Rhapsode, to close observations of students interacting with the adaptive learning engine. The aim is to explore what characterizes different types of students' behavior when using the learning resource, by studying what happens between the clicks and other user input that the machine uses for learning analytics, and how the learning resource responds to student actions. This close up investigation of student behavior using an adaptive learning material seems particularly relevant in relation to a technology that intends to personalize learning based on students' characteristics and behavior. Furthermore, the study aims to investigate how students' self-efficacy and motivation is affected by using the adaptive learning resource. Self-efficacy is the perceived capability of an individual to perform given actions (Schunk, 1991). Self-efficacy is interwoven with both academic motivation and academic proficiency. In the seminal theory of Bandura (1977), the individual's expectation of personal efficacy determines if the individual will exhibit coping behavior and invest the necessary and prolonged effort to overcome challenges and adversity. Self-efficacy can be strengthened when a person experiences ability to overcome challenges and coping with threatening situations, which, in turn, can improve academic proficiency, because the behaviors associated with self-efficacy are fundamental to benefiting from learning situations.

Viewing the adaptive learning technology *Rhapsode* through the lens of self-efficacy theory will serve to point out strengths and weaknesses in the learning material towards keeping students' self-efficacy intact or even strengthened and, in turn, keep students motivated and able to learn. Bandura (1977) identifies four sources of self-efficacy: performance accomplishments, vicarious experience, verbal persuasion, and physiological states, with the personal experience of success in performance being the strongest source of self-efficacy.

Hence, the focus of the study is on how students with different combinations of mathematical competence and motivation cope with Rhapsode's way of organizing a learning process and, conversely, how the learning tool copes with students with different prerequisites. However, the mathematics subject is a backdrop for studying student behavior and interaction with an adaptive learning technology. Therefore, we do not refer to previous studies of digitally mediated learning in the mathematics subject, mathematical self-efficacy etc.

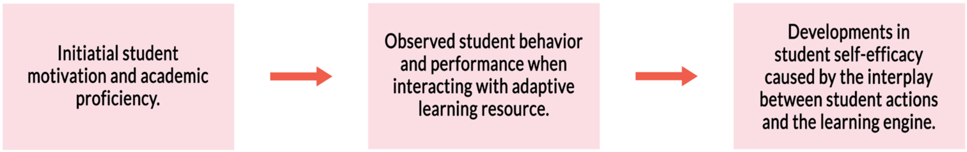
The research question of the study is as follows

How do different types of students interact with an adaptive learning material, what learning paths emerge for different types of students in their interaction with the adaptive learning resource and how is student self-efficacy affected by the interaction?

Method

In this study, we explore how Danish grade 4 students perform and interact with the digital adaptive learning resource, *Rhapsode*, for mathematics. We aim to understand how four students use the adaptive learning resource, more specifically how they utilize the resources made available by the learning material and go about solving the tasks. Furthermore, we aim to understand how students of varying mathematical proficiency and motivation towards the mathematics subject interact with the adaptive learning resource and how student self-efficacy is affected by student interaction with the adaptive learning engine (Figure 1). *Student variables* in this study are baseline and situational academic motivation towards mathematics and academic proficiency towards mathematics. The relevant *situation variables* in the learning situation are student interaction with available resources, i.e. tasks and resources providing explanations, assistance/modeling and feedback, which are delivered by the adaptive learning engine.

Figure 1.
Components in the study.



Data

Because Rhapsode is an adaptive learning resource, potentially students will be supported and challenged appropriately based on their performance and input, as soon as the adaptive engine has determined student proficiency level. Rhapsode monitors student level and progress in knowledge (what the student knows and understands), grit (the ability to persist despite difficulties), as well as meta-learning (students' ability to assess what they understand and are capable of). The teaching aid collects data in relation to the connection between the student's perceived level of certainty on whether they know or can perform specific knowledge bits or tasks and the learning resource offers an explicit, immediate assessment of student performance. How the data is interpreted by the learning resource and transformed into action, we do not know.

The results of the auto generated analysis of student progress and performance is made available in the teachers' dashboard, but in an aggregated and processed format that does not allow to gain insights into individual paths in the learning resource. However, we did gain access from the developer, Area9, to datasheets showing tagging of learning objects and the raw data generated by Rhapsode. Each learning object is tagged with a specific learning objective. Furthermore, each learning object is tagged in a content category, i.e. which type of learning object (explanation, multiple-choice question, tasks that require multiple answers etc.).

The adaptive engine generates data on:

— The exact time the student enters a learning object and time total spent on the learning object.
— Student score (0-100%) and result (correct, wrong or partially correctness in %) in relation to each learning object.
— Student self-assessment of self-confidence in relation to their understanding of explanations and their degree of certainty that they answered a task correctly.

Figure 2.
Rhapsode interactive explanation of parallelogram.

Rhapsode uses user responses and clicks to generate data about the abovementioned elements. However, much user interaction and activity is not tracked in the activity log. Firstly, there are numerous interaction possibilities on each page, which are not recorded by the learning material. Figure 2 shows a screen dump of an explanation of the parallelogram. In the left side of the screen, there is a verbal explanation. The student can choose to have it read aloud and set it to read aloud automatically when entering the page. The student can scroll down and read the written representation in the box. In the center screen the student can check the box below the figure and see the extended lines (blue), which illustrates that the sides are parallel. Furthermore, the explanation consists of three slides, each with similar verbal explanations and illustrations. The student has to click to proceed to the next slide. Nevertheless, the learning material activity log will only record total time spent and student self-assessment of understanding at the end of the slide show. Furthermore, student understanding of the explanation is not tested in conjunction with the explanation object, so there is no appraisal for student effort.

We use screen recordings of student interaction (n=4) with the learning material to observe student behavior. This allows a qualitative task-by-task investigation of student learning trajectories in the digital milieu. The task-by-task analysis is supplemented with a continuous monitoring of students' self-reported self-efficacy level.

Before submitting an answer to each task and after each presented explanation, students must assess their self-efficacy by evaluating how sure they are, that they solved the task correctly or how sure they are, that they understand the explanation put forward by the learning resource.

Screen recording is an unobtrusive method of generating data on user interaction with a digital learning resource. Compared to using for example eye tracking technology to determine what students are fixating upon and for how long, screen recording allows students to work unattended in their natural environment and they do not need to physically wear research equipment. Gaze and mouse movement has been found to interact sufficiently to provide similar information as eye-tracking as a proxy for attention in natural tracing tasks, i.e. when subjects are not instructed to consciously move their eyes (Demšar & Çöltekin, 2017). Eye and mouse movements have been found to be highly coordinated with close to 0 pixel gaze/cursor distance in online search tasks (Rodden et al. 2012). Still, analyzing screen recordings to determine student attention will only provide a rough understanding of student attention. The present study does not claim accuracy in number of seconds spent attending to given objects but rather to present a comprehensive impression of students' behavior when using the learning resource.

An adaptive learning environment such as *Rhapsode* provides the opportunity to study student self-efficacy and motivation through their behavior and performance in the *Rhapsode* learning environment:

- We can observe students' initial coping with mathematical challenges.
- We can track students' development in self-efficacy and motivation as they are making their way through the learning modules based on their actions and self-reported evaluation of their degree of certainty, that they have understood the content or solved the task correctly, which also gives us an indication of students' perceived self-efficacy.

For both points of interest, the interplay between student and learning resource must be taken into account, as the learning resource reacts to student actions. In analyzing the recordings, we first made a qualitative description of students' actions and events in relation to each exposure to a learning object. We also used all the above mentioned data on student performance and action generated by Rhapsode. To contextualize our understanding of student behavior and self-efficacy we use a baseline test and survey measuring student proficiency level in mathematics and academic motivation towards the mathematics subject.

Procedure

We administered a test of mathematics skills using a standardized test battery, *Matematikprofilen [Mathematics profile]* (Gyldendal, n.d.a, n.d.b, n.d.c) to all students in the 4^{th} grade class. In the test, the students' skills, knowledge and competences are measured in relation to the Danish National Mathematics curriculum within the mathematical subject areas: 1) numbers and algebra, 2) geometry and measurement, and 3) statistics and probability. The test consists of both closed tasks (multiple choice) and open tasks where the student must draw and/or write the correct answer. The closed tasks consist of dichotomous items (i.e. true/false), while the open tasks consist of polytomous items, where a number of points are given (e.g. 3-2-1-0) based on examples of what characterizes a task solution on different levels. The test gives a total point score that can be used to divide the students into one of five levels of subject mastery. Students had three hours to complete the test, which was administered by their regular teacher.

To measure the students' self-perceived competence in mathematics, we used a scale consisting of nine items from the international survey Trends in International Mathematics and Science Study (TIMSS) (Martin et al., 2016). The scale is used in TIMSS to compare students' self-perceived competence in mathematics across countries for both 4th and 8th grade students. In the questionnaire, the students were asked to decide on all nine statements, which in different ways are about the students' assessment of their own ability (e.g. "I usually am good at mathematics", "Mathematics is more difficult for me than it is for many of my comrades"). To measure the students' motivation towards mathematics, we used a scale consisting of nine items from TIMSS (e.g. "I like learning mathematics", "I wish I didn't have to learn mathematics") (Martin et al., 2016).

The four students screen recorded were selected at random based on a stratification that divides the students according to two parameters:

1. Whether the students' baseline math test showed a relatively high or low level of math competence.
2. Whether the students' responses to our baseline survey showed relatively high or low motivation in relation to the mathematics subject.

This gives four possible combinations of competence level and motivation (see Table 1), and one student has been selected from each of the four fields.

Table 1.
4th grade students selected for screen recording.

		Competence level		
		Low	Average	High
	1/3 lowest	Student 1		Student 3
Motivation	1/3 average			
	1/3 highest	Student 2		Student 4

Note: The students were selected by simple random selection in each of the four strata.

Hence, the strategy for case selection is maximum variation (Flyvbjerg, 2006) in an attempt to shed light on how student interaction differs with different student characteristics regarding competence level and motivation – and how the learning resource adapts to different types of students.

The study was carried out during the Covid-19 pandemic, more specifically when Danish students were attending school from home. Therefore, students' interaction with Rhapsode was recorded in the students' home, which means that they were not monitored by either researcher or teacher. Therefore, students had no possibility of cooperation and 'cheating' the computer through help from peers, a strategy which has been observed in a strain of the project the present study is part of (Gissel et al., 2020). The students freely choose a learning module for the recording, which they themselves carried out following our instructions. Rhapsode is primarily used for learner self-study, which means that the setup is similar to the usage situation intended by the producer of the learning resource. Students were instructed to use the learning resource for a minimum of 45 minutes, which all four students did.

Analysis

The learning material's sequencing of tasks and explanations

The analyses of the data generated by the screen recordings firstly explore students' learning paths across the four cases when using Rhapsode. As the algorithm behind Rhapsode is a company secret, this endeavor can be viewed as an inductive approach to try to understand how this particular algorithm works.

As a starting point we use the structure and taxonomy of learning modules. Each learning module has a rather narrow subject, e.g. fractions (parts 1-2) or coordinate system. The learning content is divided into multiple learning objectives (some 30 for each module) which in turn are associated with specific content, i.e. interactive questions, multiple choices tasks, math problems or explanations. As these partial learning objectives are structured according to a taxonomy, it is possible to track and understand the order in which Rhapsode presents this micro granulated content as well as which and when content is presented repeatedly.

To illustrate how the algorithm works student 1's interaction with the teaching aid will be analyzed to make general points about Rhapsode's way of presenting content. These points about Rhapsode's way of working thus apply to all four case students. In addition, this case description thoroughly explains which types of tasks student 1 has been exposed to in order to describe the content in the Rhapsode mathematics courses.

Student 1 chose to work with the "Division 2" module for the screen recording session. During the 45 minutes of recording, the student is involved in 59 activities (in two different modules). In Rhapsode activities are grouped into different types which are described in the following along with student 1's exposure to each type:

- 13 of the 59 activities are communicative texts where the student's knowledge is not tested (but the student must evaluate his own understanding).
- 10 are tasks that require a single answer.
- 4 are multiple-choice tasks that check student understanding.
- 11 are assignments that require the student to write two or more answers (e.g. partial calculations of a division piece that is regrouped).
- 8 are sequences where the student is scaffolded by carrying out step-by-step tasks towards a unified, more complex whole, typically a mathematical procedure.

— 13 are so-called *math problems*, which require a little more student independence to be solved (e.g., the student must write an arithmetical problem which has then has to be calculated).

In the Module "Division 2", the student has worked with a mathematical subject area which, from Area9's side, is divided into 25 learning objectives. In Appendix 1 it is explained and shown how the analytical findings are represented. Appendix 1, Table 1 shows that Rhapsode's way of putting together a learning process is atypical compared to digital skill-and-drill learning materials, in which a specific content area and a degree of difficulty is chosen, and then the student solves training tasks of increasing difficulty within the area (Gissel & Skovmand, 2018). In Rhapsode, a larger subject area is, as mentioned, granulated into smaller learning objectives. Furthermore, in Rhapsode's dashboard, you can see that the learning objectives of a given area are arranged in a taxonomy from the simple to the more complex.

However, the students' progress through the learning objectives in Rhapsode is by no means straightforward and linear, as in a typical training learning tool where students work with the same procedures with increasing difficulty level (Gissel & Skovmand, 2018). Rhapsode represents the content area as a network structure, where the student is thrown around between learning goals (points in the network) and up and down in taxonomic level and degree of difficulty.

Thus, we can see that the first task encountered by student 1 relates to learning objective 2.2. Rhapsode tries to find the student's level and the first probing is not in relation to the easiest learning goal, but rather on the top edge of the lowest third. Student 1 answers the first task incorrectly, but the teaching aid does not jump down to the lowest targets for that reason either. Instead, the student gets an explanation in relation to learning objective 2.2. and is then presented with a task in connection with objective 2.3, which the student answers correctly.

Typically, in a session, at the start there is an overrepresentation of tasks from the lower taxonomic levels with detours around the area. Likewise for student 1. If the student answers a few tasks incorrectly, the student typically gets an explanation and then more tasks. However, assignments and explanations never follow one another directly. Typically, for example, it takes 5-6 activities before the student gets a task that tests him in the learning target content that has been explained. And the other way around; in case of an incorrect answer within a learning objective, the student does 5-6 activities in relation to other learning objectives than the one he answered incorrectly within, and then there is either an explanation or a repetition of the mathematical problem (possibly with other numeric values).

In addition, the student is quickly pushed on to new learning goals by Rhapsode. The teaching aid stops challenging the student within a mathematical sub-area as soon as the teaching aid registers or believes that the student has mastered it. The teaching aid will typically consider a learning objective as completed after a single correct answer to a task.

This can be seen, for example, in student 1's meeting with learning objectives from areas 1.1 and 1.2, which the student answers correctly in the first attempt. This is the only time the student meets these two learning objectives. He also gets a few explanations regarding goals in area 2, Division of whole 10s... and answers (apart from a single slip) correctly to tasks from area 2.1-2.3. He gets these done quite quickly (by the 17th activity these are finished). In general, it can also be seen that even if the student answers incorrectly to the first task within a learning goal, it only requires one correct answer from the student before the teaching aid sends the student on to new goals. Sometimes the student has two correct answers before he is sent on.

Another characteristic of the sequencing of activities is that, in the short term, the teaching aid does not let the student hang on to the same learning goal if the student answers incorrectly. In the short term, the student is pushed on - often to learning goals that are taxonomically above what the student has answered incorrectly. This can be seen, for example, in connection with the student's attempt at learning objective 1.3. What happens is that the student answers incorrectly to the first three tasks, but correctly to the fourth, and then the learning tool moves on to other objectives. The student gets his first task within 1.3 "Perform simple division with remainder" (The task is called: 10/3=? Remainder=?) as the 14th activity, the next one after the first wrong answer is presented to the student as the 25th element, and again as the 31st followed and then the last time with a similar interval where the student finally gets it right.

Although the learning path is thus by no means straight but rather unpredictable and must seem varied and perhaps random to the student, since the student cannot know what to expect from the next task, there is a large degree of continuity in the many repetitions of tasks, if the student answers incorrectly. For student 1, there is in relation to learning objective 3.3. five exposures to the same task; generally, students will face the same tasks that they answer incorrectly over and over again until they answer correctly.

Another example of how the learning paths in Rhapsode take shape can be seen in student 1s exposure to learning objective 1.4, which is about using division with remainder in everyday situations. Here, the student first encounters an explanation where the student's knowledge is not tested. In a linear design, such an explanation would be followed by tasks where the student must apply what has been explained or show that it is understood. However, in Rhapsode, the student does not initially get a task where the student has to apply what he has learned in the video. Rather, the student is exposed to three tasks and one explanation related to learning objectives 3.1-3.5, which should be more difficult than objective 1.4. When learning objective 1.4. is revisited the student answers only partially correct. The student then is presented with five tasks from other clusters of objectives before encountering 1.4. again.

Individual student learning paths using Rhapsode

Because Rhapsode is an adaptive learning tool and because case selection aims for maximum variation, we would expect the learning paths to differ depending on student performance and characteristics.

Student 1's learning path and development in self-efficacy
Student 1 placed in the lowest third of the class both in terms of mathematical competence and motivation for the mathematics subject. At the start of the session (the first 14 activities), the student works quite thoroughly and has a high success rate in relation to answering the tasks. He reads the assignment texts, thinks carefully when he is in doubt, and on the fairly simple passages at the beginning of the course, he goes straight to the answer. In other words, the student gets off to a good start.

However, around activity 15, which is a step-by-step, scaffolded task that shows, using graphic representations, how to divide by dividing pieces by 10s, the student is challenged. In the first part of the task, the student revises his answer several times, and ends up writing an incorrect result. In the self-assessment, the student marks "I do not have a clue". Some parts of the task the student gets right, others wrong. When he answers incorrectly, he alternates between ignoring the feedback and examining it quite thoroughly. But in the last part of the task, he spends only a short time on feedback and neither accesses the resource "See more here" nor reads the explanation on the left side of the screen.

From here, the student's progress is characterized by a lack of concentration and thoroughness in reading instructions, explanations and examining feedback. It is a vicious circle, as he approaches the tasks in a skewed way the first time he sees them, and does not become much wiser through his attempts, whereby there is a basis for a series of defeats in relation to the same task. For the multiple-choice question in activity no. 25, for example, he randomly clicks around between the options, at some point sends the answer and does not consult the subsequent feedback. Or when the explanation for learning objective 3.1 is repeated after the student has answered a few tasks incorrectly, he does make the effort to understand the explanation, but quickly clicks his way through all the slides and marks "I know" in the self-assessment. A little later, the teaching aid repeats the explanation, perhaps because the student only spent 6 seconds on the last encounter with the explanation. Normally, an explanation would be followed by a test item. Nevertheless, the student does the same when the explanation is repeated, quickly clicks through it and marks "Got it". When the student is next given a task related to the learning objective, the student must solve a math problem that involves starting by making four piles with 10 in each, and he must decide how many he has distributed. The student hesitates for a while and then writes a lot of numbers and letters in the writing field and deletes it again. Clicks around the page occasionally. Finally writes "40", which is correct, and ticks "I think I know".

Overall, it seems as if the student's self-confidence is declining; his self-assessments are lower and lower as the course progresses: he marks "Not sure", "I think I know" or "I do not have a clue". Even if the teaching aid prompts him to try to revise his answer based on a hint in the feedback section, the student quickly clicks on. Occasionally he invests energy in answering correctly, but rarely succeeds. On several occasions, he clicks around on the screen's action points instead of being on task. This is where the student's recording ends. The student's motivation thus quickly seems to fade in the event of adversity, and he does not seem motivated to complete the tasks, as he rarely invests the necessary effort.

Another characteristic feature of this student's progress in the teaching aid is that the student does not prioritize familiarizing himself with the explanations that are available. The student typically doesn't spend time looking at feedback or revising his wrong answers, nor does he draw on the resources that are available. The student does not seem to expect or be interested in learning anything he does not already know, and this means that the student is actually practicing very basic mathematical operations (e.g. dividing a simple piece) without understanding what he will use them for, for example, that he can use regrouping to calculate more difficult calculations.

As can be seen in Table 2, the student does not finish the module in the 45 minutes. Several of the learning objectives are still active due to wrong answers and five of the learning objectives have not been touched at all yet when the recording ends. Therefore, there is still a long way to go for student 1. Nevertheless, we can also see from some of the previous learning objectives that the student actually manages to answer correctly, for example with learning objectives 3.1 and 3.2. Here it should be noted that the student is exposed to several repetitions of the same tasks and has access to see the correct answers and often also a recommendation for procedure. Whether the student actually understands how to transfer what he is doing to other situation, or whether he can remember the result/procedure, remains uncertain.

In summary, the teaching aid does not seem to be able to maintain the motivation of the student with a relatively low level of mathematical competence and motivation for the mathematics subject. Furthermore, the student does not develop mastery during the session, but rather loses self-confidence in relation to his mathematical self-efficacy (Mozahem et al., 2021).

Student 2's learning path and development in self-efficacy

Student 2 scored in the lowest third on the mathematics test, but is high in self-reported motivation in relation to the mathematics subject. The student chose to work with the module "Multiplication 2". The student's characteristics help explain his behavior in Rhapsode. The student fairly consistently overestimates his own abilities, has fairly high self-confidence and a performance-oriented (but not mastery-oriented) motivation, but lacks basic math skills/understanding.

In the module, the student worked with a total of 23 learning objectives (see Appendix 2, Table 1). The student has a high cadence. He manages to be in contact with 114 learning objects in the 45 recorded minutes. In comparison, student 1 managed 59. He answers many tasks and is not afraid to try to answer, and to answer quickly – even without familiarizing himself with the task formulations. He also goes through the explanations very quickly, very often without familiarizing himself with them. The student thus shows great motivation in relation to answering the tasks. Also, student 2 marks a high degree of self-efficacy in the student's self-assessments in Rhapsode. Thus, the student quite consistently marks that he "knew it" in advance during explanations, and that he is sure that he answers the tasks correctly.

Student 2 gets off to a relatively good start. At the start of the session, the student takes plenty of time to read explanations and solve the tasks. The student also looks at part of the feedback for incorrect answers. This indicates motivation to learn. However, right from the start he gets a lot of incorrect answers when he is challenged beyond target area 1, Commutative law. By task 14, however, it seems that the student begins to doubt his own abilities. He is faced with a passage that is slightly higher in the learning target taxonomy than the previous tasks. The student fills in the wrong answer and for once marks "I think I know it".

Around activity 33-51, Rhapsode stays within learning objective areas 2-3. This is probably because the student answers the majority of the tasks incorrectly. After a quarter of an hour, the teaching aid begins to push the student up the taxonomic ladder anyway, even if the student's poor performance does not justify an increase in the degree of difficulty. This means that the student is challenged beyond his abilities in long passages. Around activity 61, the student also atypically marks "Not sure" in the self-evaluation and answers incorrectly. This could be a sign that he is losing his confidence, which was otherwise at its peak in the first half of the recording.

Student 2 is clearly challenged. He gets a total of 9 exposures for the same task, albeit with different numerical values: The tasks are variations on the piece 800x400, i.e. multiplication of whole 100s. In addition, the student receives the same explanation a total of three times. Nevertheless, the first eight times he encounters the task, he answers incorrectly.

The student thus makes an effort to answer, but not to a great extent to familiarize himself with the explanations and consult feedback. It's amazing how long he can maintain the belief that he can answer the same tasks correctly, despite repeated feedback to the contrary. Gradually the student begins to doubt himself. This can also be seen by the fact that he changes a correct to an incorrect answer twice before submitting his answer.

The student melts down at the end of the 45 minute session. He has been given many repetitions of the same tasks to which he has answered incorrectly. In the end, he apparently doesn't care anymore. In the process, however, he shows a striking persistence despite being constantly told that he is answering incorrectly. He has high belief in his own abilities until the last third of the recording. He overestimates himself. He does not look carefully at explanations and usually not at feedback either, but answers the next tasks without hesitation and is usually sure that he answers correctly. This student probably leaves the session with lower self-efficacy than when he started.

Student 3's learning path and development in self-efficacy

Student 3 is characterized by good math skills but low motivation for the math subject. The student chooses to work with the module "Coordinate system". However, the student manages to complete the module in around 18 minutes, which is why he also begins the module "Perimeter and area". In the analysis, we will concentrate on his progress through the module "Coordinate system" as it shows how the learning tool challenges the student, who has either already learned or can easily understand the content and answer the tasks. In the "Coordinate system" module, the student has worked with 25 learning objectives (see Appendix 3, Table 1).

The student completes the 25 learning objectives through only 45 interactions with learning objects. As can be seen in Appendix 3, he only gets a single exposure to learning objects in relation to 13 of the 25 learning objectives. The value of the learning tool not wasting the skilled student's time; the student is quickly sent on when the teaching aid finds that he has mastered an objective.

Given the observation that students are typically sent on quickly when they can answer a single task correctly within a learning objective, there is a rather surprising course around the learning objective "Plot a point given its coordinates" (4.3). The student gets off to a good start. He gets an explanation of how to plot points into a coordinate system. At the end of the explanation is a task that he answers correctly. 5 tasks later comes a task that tests whether he can apply what he was told in the explanation – he answers this correctly. He is then given an identical task (with the same values), which he also answers correctly. Right after, the explanation is repeated with the task for the same learning objective, where the student simply clicks on instead of solving the task (which is understandable given the history). Then the student is exposed to the same task as the previous two.

The student displays a high level of metacognition. He does not have many mistakes and is also self-confident in his assessment of his own achievements, as he marks "Know it" or "Knew it" a total of 39 times. Typically, he shows uncertainty about the tasks in which he makes mistakes. All in all, the student gets through the module fairly quickly. It seems like he knew the material well and just needed a refresher. Student 3 gets refreshed or quickly learned the few things that he has forgotten or could not do beforehand. As the student is not challenged largely by the module and as he can accurately assess his own competencies and uncertainties his self-efficacy would seem to be consolidated or heightened through the session.

Student 4's learning path and development in self-efficacy
According to the test result and survey, student 4 is characterized by a relatively high mathematical competence level and high motivation for the mathematics subject. The student chose to work with the module "Division 2" where she worked with 26 learning objectives. The student managed to be in contact with a total of 56 learning objects in the 45 minutes she recorded (Appendix 4, Table 1).

Student 4 is quite thorough in her approach to answering the tasks, as she thinks carefully before answering. She also spends quite a lot of time familiarizing herself with the various resources – including explanations. The student generally makes a great effort to answer correctly and understand. However, at an early stage, the student clicks away the explanation on the left side (Figure 2) and it remains gone for the rest of the session.

The student gets some wrong answers at the start of the session, and the teaching aid presents the student with quite a few explanations. The student also answers incorrectly to some of the tasks that test the student's understanding in connection with explanations. The student thus does not have an easy time. It is thus the student's rather high degree of self-efficacy that pulls her through, and her thoroughness could indicate that she is mastery-oriented.

The student exhibits a high degree of metacognition. She marks relevant and precise in the self-assessments of how certain she is that she has answered correctly or understood an explanation when she is not sure. Only three times does she mark "Know that" for an explanation, and a total of 34 times she assesses that she "Thinks it is right", what she has answered in an assignment. "Completely blank" is marked six times. This humility or caution may be the reason why Rhapsode, with this particular student, in several cases provides additional tasks within a learning objective, even if the student has managed to answer correctly once. If the student does not feel confident, the student is given tasks to consolidate her knowledge and skills. However, it cannot be seen that the student increases her self-efficacy or belief in herself through the session. But in this student's case, motivation and self-efficacy translate into thoroughness and care in relation to familiarizing herself with the resources, whereby the learning tool works well for the student, who slowly but surely works her way through the learning objectives.

Concluding discussion

The study has shown how assignments and explanations are presented to the student in Rhapsode. The learning tool probes the student from the start to determine competence level. Thus, the learning tool does not start with the easiest learning objectives. Moreover, the student is constantly sent around between different learning objectives, which can be both high and low in the producer's taxonomic arrangement of learning objectives.

The student rarely encounters tasks related to the same learning objectives in succession. In the case of an incorrect answer, the student is met with an explanation or a repetition of the task (most often with new numerical values), but these only come after the student has dealt with a variety of tasks and explanations related to other learning goals. Typically, approximately six activities will pass before the learning objective is taken up again.

Students are quickly sent on to new learning goals. A single correct answer is enough for a learning objective to be considered completed in Rhapsode – unless the student himself expresses doubts in his self-assessment. This may cause student failure to experience mastery; for some students, they have answered several tasks incorrectly, but when they finally get a correct answer within a learning objective, Rhapsode moves on to other objectives rather than allowing the student to consolidate. The need to consolidate and experience mastery should be considered in future designs.

In general, students neglect to access and utilize the supportive resources offered by the learning tool. Everyone ignores the explanatory box on the left side of the screen, the opportunity to revise answers after feedback is rarely used, and it is certainly not consistent that students examine the feedback they receive.

Since uniform tasks are repeated in the event of incorrect answers from the student, and potentially repeated many times, it can be difficult to know whether the student fundamentally understands the mathematical connections that the task requires, or whether the student can remember previous answers and use this to solve the task without substantial understanding.

Rhapsode worked differently for the four case students, and there were substantial differences in learning paths for students who have relatively high or low level of mathematical competence combined with resp. relatively high or low motivation in relation to the mathematics subject.

Rhapsode did not work optimally for student 1, who in the test and survey was in the lowest third in terms of both mathematical competence and motivation. The student lost motivation and self-confidence and was unable to get help from the support measures of the teaching aid and probably did not learn much new. Teachers must be aware that a student with relatively poor mathematical competence and motivation may need to be kept on track in relation to familiarizing himself with the teaching aid's explanations, consulting the teaching aid's feedback in case of wrong answers and investing the time necessary to answer the tasks.

Student 2 was characterized by a combination of low mathematical skills but high motivation for the subject. This was expressed by the fact that the student was very persistent despite receiving feedback from the teaching aid that he had answered many tasks incorrectly. He did not act very strategically in relation to the supporting resources in the teaching aid, but he still managed to move forward through the learning objectives until he lost heart at the end. Therefore, the learning tool worked well for this student as long as his motivation was intact.

For Student 3, who had good initial mathematics skills but low motivation for the mathematics subject, the teaching aid would be suitable as preparation for a course on a given mathematical subject area. The student has potentially become more aware of what he can do, and this will be a good starting point to build on. Even if the student scores low in terms of motivation for the mathematics subject, it seems that the student can maintain attention with Rhapsode's pace and flow.

Student 4, who scored high on both mathematical competence and motivation for the subject, the teaching aid also worked well. The module the student had chosen caused her challenges, but her thoroughness, mastery-oriented motivation and self-efficacy led her steadily and calmly towards fulfilling the learning goals.

Thus, high self-efficacy appears to be crucial for students to get a good result from working with Rhapsode. If the student gives up too easily or fails to invest effort, it does not appear that Rhapsode can create a positive and productive learning experience for the student.

The teaching tool can be used to uncover discrepancies between a student's self-efficacy and actual ability. This discrepancy can result, for example, in the student being performance-oriented in relation to answering many tasks but is not inclined to learn new things from the teaching material. In this situation, the teacher should show the student how, through the explanations and feedback of the teaching aid, he can be better equipped to solve the tasks correctly.

This study confirms the findings of previous studies as presented in the state-of-the art of this paper, that adaptive technology has not yet matured to be able to replace the teacher and make classroom level pedagogy (Du Boulay, 2019) superfluous. This is particularly relevant regarding showing students the need of using the resources in the learning material but also to broaden the scope vis-à-vis the mathematics subject. For example, students may need the teacher to present alternative ways of approaching tasks, as the teaching aid only presents one approach in explanations.

References

Apoki, U.C., Hussein, A.M.A., Al-Chalabi, H.K.M., Badica, C. & Mocanu, M.L. (2022). The Role of Pedagogical Agents in Personalised Adaptive Learning: A Review. *Sustainability*, *14*(11), 6442. https://doi.org/10.3390/su14116442

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review, 84*(2), 191-215.

Bulman, G. & Fairlie, R. W. (2016). Technology and Education: Computers, Software, and the Internet. In: Hanushek, E.A., Machin, S. & Woessmann, L. (Eds.) Handbook of the Economics of Education, vol 5, 239-280. Elsevier. https://doi.org/10.1016/B978-0-444-63459-7.00005-1

Conejo, R., Guzmán, E., Millán, E., Trella, M., Pérez-De-La-Cruz, J. L. & Ríos, A. (2004). SIETTE: A Web-Based Tool for Adaptive Testing. I*nternational Journal of Artificial Intelligence in Education, 14*(1), 29-61.

Cuban, L. (2015). The Lack of Evidence-Based Practice: The Case of Classroom Technology. Retrieved from: https://www.larrycuban.wordpress.com/2015/02/05/the-lack-of-evidence-based-practice-the-case-of-classroom-technology-part-1/

Demšar, U. & Çöltekin, A. (2017). Quantifying gaze and mouse interactions on spatial visual interfaces with a new movement analytics methodology. *PloS one, 12*(8), e0181818. https://doi.org/10.1371/journal.pone.0181818

Du Boulay, B. (2016). Artificial intelligence as an effective classroom assistant. *IEEE Intelligent Systems, 31*(6), 76-81.

Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative inquiry, 12*(2), 219-245.

Gerick, J., Eickelmann, B., Vennemann, M. (2014). Zum Wirkungsbereich digitaler Medien in Schule und Unterricht. *Jahrbuch der Schulentwicklung*, 18, 206-238.

Gissel, S. T., Gynther, K., Hansen, T.I., Højgaard, T., Jørnø, R.V.L., Nortvig, A.-M. & Pettersson, M. (2020). *Rhapsode – design, brug og virkning*. Rapport, Læremiddel.dk. https://laeremiddel.dk/wp-content/uploads/2021/06/Rhapsode_Rapport_L%C3%A6remiddel.dk_080621.pdf

Gissel, S. T. & Skovmand, K. (2018). *Kategorisering af digitale læremidler: En undersøgelse af didaktiske, digitale læremidlers karakteristika*. (2 udg.) Læremiddel.dk. https://laeremiddel.dk/wp-content/uploads/2018/05/Kategorisering-af-digitale-l%C3%A6remidler.pdf

Guzmán, E., Conejo, R. & Pérez-de-la-Cruz, J. L. (2007). Improving Student Performance using Self-Assessment Tests. *IEEE Intelligent Systems,* 22(4), 46-52. DOI: 10.1109/MIS.2007.71.

Gyldendal (n.d.a). **Om Matematikprofilen**. Retrieved December 11, 2020, from https://matematikprofilen.gyldendal.dk/

Gyldendal (n.d.b). *Om Matematikprofilen – Uddybende beskrivelse*. Retrieved December 11, 2020, from https://portalmotor.gyldendal.dk//-/media/Home/matematik/matematikprofilen/dokumenter/Kategoribeskrivelser/Kategoribeskrivelser---Matematikprofilen-3.ashx

Gyldendal (n.d.c). *Om Matematikprofilen – Resultater*. Retrived December 11, 2020, from https://matematikprofilen.gyldendal.dk/3-klasse/resultater

Holmes, W., Anastopoulou, S., Schaumburg, H. & Mavrikis, M. (2018). *Technology-enhanced personalised learning: untangling the evidence*. Robert Bosch Stiftung. Available at: http://oro.open.ac.uk/56692/

Johnson, D., & Samora, D. (2016). The potential transformation of higher education through computer-based adaptive learning systems. *Global Education Journal, 2016*(1), 1-17.

Kerr, P. (2015). Adaptive learning. *ELT Journal, 70*(1), 88-93. https://doi.org/10.1093/elt/ccv055

Kulik, J. & Fletcher, J. D. (2015). Effectiveness of Intelligent Tutoring Systems: A Meta-Analytic Review. *Review of Educational Research, 86*(1), 42-78. https://doi.org/10.3102/0034654315581420.

Liu, M., McKelroy, E., Corliss, S. B. & Carrigan, J. (2017). Investigating the effect of an adaptive learning intervention on students' learning. *Education Tech Research and Development, 65*, 1605-1625. DOI:10.1007/s11423-017-9542-1

Martin, F., Chen, Y., Moore, R. L. & Westine, C. D. (2020). Systematic review of adaptive learning research designs, context, strategies, and technologies from 2009 to 2018. *Education Technology Research and Development*, 68, 1903-1929. https://doi.org/10.1007/s11423-020-09793-2

Martin, M. O., Mullis, I. V. S., Hooper, M., Yin L., Foy, P. & Palazzo, L. (2016). Creating and Interpreting the TIMSS 2015 Context Questionnaire Scales. In: Martin, M.O, Mullis, I.V.S, Hooper, M. (Ed.), *Methods and Procedures in TIMSS 2015*. TIMSS & PIRLS International Study Center, Boston College.

Modén, M. U. (2021). *Teaching with digital mathematics textbooks. Activity theoretical studies of data-driven technology in classroom practices*. Doctoral Dissertation. Department of Applied Information Technology, University of Gothenburg. http://hdl.handle.net/2077/69472

Mozahem, N. A., Boulad, F. M. & Ghanem, C. M. (2021). Secondary school students and self-efficacy in mathematics: Gender and age differences. *International Journal of School and Educational Psychology, 9*(1), 142-152.

Nakic, J., Granić, A. & Glavinić, V. (2015). Anatomy of student models in adaptive learning systems: A systematic literature review of individual differences from 2001 to 2013. *Journal of Educational Computing Research, 51*(4), 459-489. doi:10.2190/EC.51.4.e.

Nguyen, L. (2015). A User Modeling System for Adaptive Learning. *The 17th International Conference on Interactive Computer aided Learning (ICL2014)*. Doi: 10.1109/ICL.2014.7017887

OECD (2015). *Students, Computers and Learning: Making the Connection*. Retrieved from https://www.oecd.org/publications/students-computers-and-learning-9789264239555-en.htm

Peng, H., Ma, S. & Spector, J. M. (2019). Personalized Adaptive Learning: An Emer- ging Pedagogical Approach Enabled by a Smart Learning Environment. In: M. Chang, E. Popescu, Kinshuk, N-S. Chen, M. Jemni, R. Huang, J. M. Spector & D. G. Sampson (red.), *Foundations and Trends in Smart Learning. Lecture Notes in Educational Technology* (p. 171-176). Springer. DOI:10.1007/978-981-13-6908-7_24

Pollard, A. & James, M. (eds) (2004). *Personalised Learning: A Commentary by the Teaching and Learning Research Programme*, TLRP/ESRC, Swindon, UK. doi: 10.1080/0305764X.2017. 1375458.

Rodden, K., Fu, X., Aula, A. & Spiro, I. (2008). Eye-mouse coordination patterns on web search pages. *Conference on Human Factors in Computing Systems – Proceedings*. 2997-3002. 10.1145/1358628.1358797.

Schunk, D. H. (1991). Self-Efficacy and Academic Motivation, *Educational Psychologist*, (26)3-4, 207-231. DOI: 10.1080/00461520.1991.9653133

Somyürek, S. (2015). The new trends in adaptive educational hypermedia systems. *International Review of Research in Open and Distributed Learning, 16*(1), 221-241.

Tai, D. W-S., Tsai, T-A. & Chen, F. M-C. (2001). Performance Study on Learning Chinese Keyboarding Skills Using the Adaptive Learning System. *Global Journal of Engineering Education, 5*(2), 153-161.

Tamim, R., Bernard, R. M., Borokhovski, E. & Abrami, P. C. (2011): What Forty Years of Research Says About the Impact of Technology on Learning. *Review of Educational Research, 81*(1), 4-28.

Taylor, R. P. (1980). Introduction. In: R. P. Taylor (Ed.), *The computer in school: Tutor, tool, tutee*, 1-10. Teachers College Press.

U.S. Department of Education, Office of Educational Technology. (2017). *Reimagining the Role of Technology in Education: 2017 National Education Technology Plan Update*. https://tech.ed.gov/files/2017/01/NETP17.pdf

Verdu, E., Regueras, L. M., Verdu, M. J., De Castro, J. P. & Pérez, M. A. (2008). Is adaptive learning effective? A review of the research. WSEAS International Conference. *Mathematics and Computers in Science and Engineering*, 7.

Walkington, C. (2013). Using Adaptive Learning Technologies to Personalize Instruction to Student Interests: The Impact of Relevant Contexts on Performance and Learning Outcomes. *Journal of Educational Psychology, 105*(4),932-945. https://doi.org/10.1037/a0031882.

Wang, S., Christensen, C., Cui, W., Tong, R., Yarnall, L., Shear, L. & Feng, M. (2020). When adaptive learning is effective learning: comparison of an adaptive learning system to teacher-led instruction. *Interactive Learning Environments*. https://doi.org/10.1080/10494820.2020.1808794.

Xie, H., Chu, H.-C., Hwang, G.-J. & Wang, C.-C. (2019). Trends and development in technology-enhanced adaptive/personalized learning: A systematic review of journal publications from 2007 to 2017. *Computers & Education*, 140, 1-16.

Appendix 1. Student 1's progress through a learning module in Rhapsode and explanation and display of analytical findings

In the two rows on the left in Table 2, the 25 learning objectives are listed in column 2. The learning objectives are in the order that they have in Area9's taxonomic order. Each learning objective within a subarea is provided by us with a number for identification in the analyzes (column 1, from 1.1. to 5.1).

Table 1.
Student 1's progress through the "Division 2" module.

#	Learning objective	E1	E2	E3	E4	E5	E6	Class avg.
0	Introduction							
	Remainder							
1.1	Remember that "remainder" is what is left after dividing equally	(5)						92%
1.2	Determine the remainder in a division	(6)						100%
1.3	Perform simple division with remainder	(14)	(25)	(31)	(38)			57%
1.4	Use simple division with remainder	F (19)	50 (26)	F G (32)	(39)			58%
	Division of whole 10s, 100s and 1000s with 10s and 100s							
2.1	Divide numbers ending in 0 by 10	(2)	(8)					85%

2.2	Divide numbers ending in 00 by 10	(1)	G (7)					92%
2.3	Divide numbers ending in 000 by 10	(3)						92%
2.4	Divide numbers ending in 00 by 100	F (4)	(10)					50%
2.5	Divide numbers ending in 000 by 100	(11)	G (17)					91%
Divide numbers up to 99 by eight-digit numbers								
3.1	Divide numbers up to 99 step by step with graphical repre-sentation	F (9)	40 (15)	80 G (22)	F (28)	F G (34)	G2 (41)	77%
3.2	Divide numbers up to 99 step by step in everyday situations	44 (16)	69 G (23)	50 G3 (30)	84 G3 (35)	F G (34)	G4 (43)	73%
3.3	Divide numbers up to 99 using regrouping	F (21)	(27)	G (37)	G2 (45)	25 G3 (34)	G4 (57)	53%
3.4	Divide numbers up to 99 by 3-5 without a remainder	F (36)	(44)	G (52)	G2 (58)			41%
3.5	Divide numbers up to 99 by 6-9 without a remainder	(20)	90 G (33)	G1 (40)	G2 (47)	F (49)	G3 (55)	42%
3.6	Divide numbers up to 99 by the remainder							80%
3.7	Divide numbers up to 99 step by step with graphical repre-sentation	(53)						69%
3.8	Describe an everyday problem with a division expression							100%

Quotients of products of 10								
4.1	Use tables to divide numbers up to 1000 ending in 0 by eight-digit numbers	F (12)	(18)	G (24)				60%
4.2	Use tables to divide numbers up to 1000 that end in 0 by whole tens	(13)						35%
4.3	Divide numbers up to 1000 that end in 0 by even-digit numbers							80%
4.4	Divide numbers up to 1000 that end in 0 by whole tens							57%
4.5	Divide numbers up to 999 by a single digit number							0%
4.6	Divide numbers up to 999 using regrouping	(29)	G (42)	G2 8 (48)	F (54)			68%
4.7	Divide numbers up to 999 by 3 - 5 without a remainder	(46)	G (50)	G2 (56)				52%
4.8	Divide numbers up to 999 by 6 - 9 without a remainder							83%
Divide numbers up to 999 by the remainder								
5.1	Use division of numbers up to 999 in everyday situations	(59)						45%

In Table 1a, the rows with the 20 learning objectives that the student managed to work on in the 45 minutes he recorded are marked with colored boxes. The table also shows how many times the student has encountered a mediating sequence or task that targets each learning goal (E1, E2, etc.), as well as the student's progression and success in each activity. For each exposure, it is marked whether the student answered correctly (green), incorrectly (red), partially correct/incorrect (yellow). Numbers at partially true/false show percent correctness.

F marks that the object is an explanation, blue box with F that it is an explanation that does not test the student's understanding or skills. Numbers at partially true/false show percent correctness. F marks that the object is an explanation, and a blue box with F signals that it is only an explanation that does not test the student's understanding or skills. Numbers in parentheses show the chronology or progression between learning objectives and objects; (1) is the first learning object the student has encountered after the introduction to the module. G marks that the task is a repetition of a previous one (possibly with other values); in the case of multiple repetitions, the number of repetitions is marked by prefixing the number with G, e.g. "G2" (second repetition, i.e. third occurrence). The column on the far right shows the class's average percentage of correctness on tasks within the learning objective.

Appendix 2. Student 2's path through a module in Rhapsode.

Table 1.
Student 2's path through a module in Rhapsode.

#	Learning objective	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	Class avg.
0	Introduction														
		Commutative law													
1.1	Define the commutative law of multiplication	F (5)	(11)												71%

1.2	Apply commutative law	(6)													64%
		Multiply by 10, 100 and 1000													
2.1	Multiply a whole number by 10	F (3)	(9)	G (15)	G2 (21)	F G (29)	G3 (35)	G4 (41)							75%
2.2	Remember the rule for multiplication by 10	(42)	G (56)	F (62)	G2 (70)										46%
2.3	Multiply a whole number by 100	(1)													75%
2.4	Remember the rule for multiplying by 100	(2)	G (8)	F (14)	G2 (20)	G3 (22)	G4 (28)	G5 (36)							54%
2.5	Multiply a whole number by 1000	(4)	F (12)	G (23)	G2 (30)										70%
2.6	Remember the rule for multiplication by 1000	F (31)	G (38)	G2 (44)	F G2 (46)	G3 (52)	G4 (58)	G5 (65)	F G3 (71)	G6 (80)					55%
		Multiply numbers ending in 0													
3.1	Multiply a one-digit number and a whole 10's	(10)	G (16)												58%
3.2	Multiply two whole 10s	(27)	G (37)	G2 (47)	G3 (54)	G4 (63)	F (66)	G5 (73)	G6 (78)	G7 (84)	F G (88)	G8 (94)	G9 (102)		39%
3.3	Multiply a one-digit number and a whole 100's	F (7)	(13)	G (19)	G2 (25)	G3 (33)									50%

3.4	Multiply whole 10s and whole 100s	(17)	F (24)	G (34)											44%
3.5	Multiply whole 100s	(18)	G (26)	G1 (32)	F (39)	G2 (45)	F G (31)	G3 (57)	G4 (64)	G5 (68)	F G2 (73)	G6 (79)	G7 (85)	G8 (92)	28%
		Estimated calculation													
4.1	Identify correct rounding of two-digit factors in the context of estimates	F (99)	(105)	(109)											67%
4.2	Determine the approximate product of two two-digit numbers	(91)													67%
4.3	Determine the approximate product in everyday situations	F (7)	G (106)												67%
		Multiply one-digit and two-digit numbers													
5.1	Determine the product of a one-digit and a two-digit number up to 19 step by step using the area method	F (90)	75 (96)	75G (103)	F G (108)										91%
5.2	Multiply one-digit and two-digit numbers up to 19 with division	F (59)	F G (69)	75 (75)	75 G (81)	F G2 (86)	G2 (93)	G3 (100)							83%
5.3	Multiply one-digit and two-digit numbers up to 99 with division	(40)	75 G (49)	F (55)	(61)	G (67)	F G (74)	50 G2 (83)	G3 (89)	F G2 (95)	G3 (101)	G4 (107)	F G3 (112)	G5 (114)	38%
5.4	Multiply a one-digit and a two-digit number vertically	F (43)	F G (48)	G (53)	G2 (60)	F G2 (76)	G3 (82)	G4 (87)							50%
5.5	Calculate the product of a one-digit and a two-digit number in everyday situations	(98)	G (104)	G2 (110)											77%

		Multiply one-digit and three-digit numbers													
6.1	Multiply one-digit and three-digit numbers vertically	F (50)	(72)	G (77)											57%
6.2	Multiply one-digit and three-digit numbers	(111)													64%

Note: Learning objectives are shown in the two rows on the left of the table. Each learning objective within a sub-area is provided with a number for identification in the analyses. The table shows the number of exposures within a learning objective (E1, E2, etc.). For each exposure, it is marked whether the student answered correctly (green), incorrectly (red), partially correct/incorrect (yellow). Numbers at partially true/false show percent correctness. Blue marks that the object is an explanation, blue box with F that signifies an explanation that does not test the student's understanding or skills. Numbers in parentheses show the chronology or progression between learning objectives and objects; (1) is the first learning object the student has encountered after the introduction to the module. G marks that the task is a repetition of a previous one (possibly with other values); in the case of multiple repetitions, the number of repetitions is marked by placing the number after G, e.g. "G2" (second repetition, i.e. third occurrence). The column on the far right shows the class's average correct percentage on tasks within the learning objective.

Appendix 3. Student 3's path through a module in Rhapsode.

Table 1.
Student 3's path through a module in Rhapsode.

Ny#	Learning objective	E1	E2	E3	E4	E5	Class avg.
0	Module introduction						
	Definitions						
1.1	Identify the x-axis of a coordinate system						100%

1.2	Identify the y-axis of a coordinate system						88%
1.3	Identify the origin of a coordinate system	F					86%
1.4	Define the origin						65%
1.5	Define coordinates						52%
1.6	Describe the notation convention						60%
1.7	Identify the x-coordinate of a point given by its coordinates	F					57%
1.8	Identify the y-coordinate of a point given by its coordinates		G				65%
	Identify a point given its coordinates						
2.1	Select points with the same x-coordinate						82%
2.2	Select points with the same y-coordinate						56%
2.3	Identify a point given its coordinates		G				69%
	Write the coordinates of a given point						
3.1	Describe how to find the x-coordinate of a given point						90%
3.2	Describe how to find the y-coordinate of a given point						72%
3.3	Write the x-coordinate of a given a point						100%
3.4	Write the y-coordinate of a given a point						90%
3.5	Describe the meaning of the x-coordinate of a point						
3.6	Describe the meaning of the y-coordinate of a point	F					
3.7	Write the coordinates of a given point		G	G2			89%
	Plot a point						
4.1	Plot a point with a given x-coordinate						100%
4.2	Plot a point with a given y-coordinate						75%

4.3	Plot a point given its coordinates	F		G	F	G2	96%
4.4	Given the point (x,y), plot the point (y,x)						47%
	Paths and shapes						
5.1	Identify a paths/shape given by points	F		75	F		81%
5.2	Write the coordinates of the corners of a path/ shape						52%
5.3	Plot the corner points of a path/shape						36%

Note: Learning objectives are shown in the two rows on the left of the table. Each learning objective within a sub-area is provided with a number for identification in the analyses. The table shows the number of exposures within a learning objective (E1, E2, etc.). For each exposure, it is marked whether the student answered correctly (green), incorrectly (red), partially correct/incorrect (yellow). Numbers at partially true/false show percent correctness. Blue marks that the object is an explanation, blue box with F that signifies an explanation that does not test the student's understanding or skills. Numbers in parentheses show the chronology or progression between learning objectives and objects; (1) is the first learning object the student has encountered after the introduction to the module. G marks that the task is a repetition of a previous one (possibly with other values); in the case of multiple repetitions, the number of repetitions is marked by placing the number after G, e.g. "G2" (second repetition, i.e. third occurrence). The column on the far right shows the class's average correct percentage on tasks within the learning objective.

Appendix 4. Student 5's path through a module in Rhapsode.

Table 1.
Student 4's path through a module in Rhapsode.

#	Learning objective	E1	E2	E3	E4	E5	E6	E7	Class avg.
0	Introduction								
	Remainder								

1.1	Remember that "remainder" is what is left after dividing equally	F (3)	(10)						92%
1.2	Determine the remainder in a division	(4)							100%
1.3	Perform simple division with remainder	(26)							57%
1.4	Use simple division with remainder in everyday situations	(28)							58%
Division of whole 10s, 100s and 1000s with 10s and 100s									
2.1	Divide numbers ending in 0 by 10								85%
2.2	Divide numbers ending in 00 by 10								92%
2.3	Divide numbers ending in 000 by 10	(1)							92%
2.4	Divide numbers ending in 00 by 100	(2)	F (9)	G (16)	G2 (22)				50%
2.5	Divide numbers ending in 000 by 100	(8)							91%
Divider tal op til 99 med etcifrede tal									
3.1	Divide numbers up to 99 step by step with graphical representation	F (5)	40 (12)	60 G (18)	F G (24)	F G2 (30)	80 G2 (39)	100 G3 (46)	77%

3.2	Divide numbers up to 99 step by step in everyday situations	25 (14)	75 G (20)	75 G2 (27)	F (31)	75 G3 (37)	G4 (43)		73%
3.3	Divide numbers up to 99 using regrouping	(6)	F (15)	G (21)					53%
3.4	Divide numbers up to 99 by 3 - 5 without a remainder	(32)	G (34)	G2 (38)					41%
3.5	Divide numbers up to 99 by 6 - 9 without a remainder	(25)	G (45)						42%
3.6	Divide numbers up to 99 by the remainder	F (44)							80%
3.7	Use division of numbers up to 99 in everyday situations	F (33)	F (40)						69%
3.8	Describe an everyday problem with a division expression	(51)							100%
Quotients of products of 10									
4.1	Use tables to divide numbers up to 1000 ending in 0 by eight-digit numbers	F (7)	(13)	G (19)	(52)	G (54)			60%
4.2	Use tables to divide numbers up to 1000 that end in 0 by whole tens	(17)	G (23)	F (29)	(35)	(42)	G (49)		35%
4.3	Divide numbers up to 1000 that end in 0 by even-digit numbers	F (48)	F G (55)						80%
4.4	Divide numbers up to 1000 that end in 0 by whole tens	F (11)	(50)						57%

4.5	Divide numbers up to 999 by a single digit number								0%
4.6	Divide numbers up to 999 using regrouping	(36)							68%
4.7	Divide numbers up to 999 by 3 - 5 without a remainder	(34)							52%
4.8	Divide numbers up to 999 by 6 - 9 without a remainder								83%
Divide numbers up to 999 by the remainder									
5.1	Use division of numbers up to 999 in everyday situations	F (41)	(47)	70 (53)	G (56)				45%

Note: Learning objectives are shown in the two rows on the left of the table. Each learning objective within a sub-area is provided with a number for identification in the analyses. The table shows the number of exposures within a learning objective (E1, E2, etc.). For each exposure, it is marked whether the student answered correctly (green), incorrectly (red), partially correct/incorrect (yellow). Numbers at partially true/false show percent correctness. Blue marks that the object is an explanation, blue box with F that signifies an explanation that does not test the student's understanding or skills. Numbers in parentheses show the chronology or progression between learning objectives and objects; (1) is the first learning object the student has encountered after the introduction to the module. G marks that the task is a repetition of a previous one (possibly with other values); in the case of multiple repetitions, the number of repetitions is marked by placing the number after G, e.g. "G2" (second repetition, i.e. third occurrence). The column on the far right shows the class's average correct percentage on tasks within the learning objective.

Abstract

Emerging discussions highlight the role of chatbots in education, which centres around fundamental perspectives concerning technology as either beneficial or an unnecessary disruption to the learning and assessment processes. However, it is crucial to recognise that no technology can be deemed value-neutral, and technological euphoria often blinds us to the unintended consequences of its use. Hence, there is a need for a practical understanding informed by value-based perspectives to comprehend how the encoded experiences created by ChatGPT transform the teaching of students. This article makes a significant contribution to the field of educational research, as only limited attention has been given to higher education students' reflections on the pedagogical use of artificial intelligence (AI) chatbots for learning and assessment.

The article discusses, through empirical findings, the constructive alignment between learning objectives and assessment when employing AI chatbots. The data collection process revolves around students' reflective experiences while utilising the potential of ChatGPT, which are subsequently analysed through focused group interviews. The paper, informed by these observational data, presents a conceptual framework for teachers to employ and assess ChatGPT as a means to scaffold learning among students.

Chatbots rolle i undervisningen er ofte betragtet som enten gavnlig eller en unødvendig forstyrrelse i lærings- og vurderingsprocessen. Ingen teknologier er værdineutrale, og teknologisk eufori gør os ofte blinde for et utal utilsigtede konsekvenser af deres brug. Der er behov for en mere praktisk forståelse og værdibaserede perspektiver på, hvordan de kodede oplevelser skabt af ChatGPT ændrer undervisningen og studerendes læringsforløb. Denne artikel bidrager til dette forskningsfelt, da der indtil nu kun har været begrænset opmærksomhed på studerendes refleksioner over brugen af AI Chatbots på de videregående uddannelser.

Artiklen analyserer etiske og pædagogiske perspektiver på videregående uddannelsers AI anvendelse i den pædagogiske praksis gennem casestudiet. Empirien er baseret på fokusgruppeinterviews og centrerer sig om de studerendes reflekterende oplevelser, tanker og følelser i relation til brug af ChatGPT. Artiklen konceptualiserer fundene i en begrebsramme for undervisernes til brug for vurdering af ChatGPT til at stilladsere læring hos de studerende.

The role of AI chatbots in scaffolding: Linking learning outcomes with assessment

Camilla Gyldendahl Jensen, Professionshøjskolen UCN, Susanne Dau, Professionshøjskolen UCN, & Peter Gade, Professionshøjskolen UCN

Introduction

In 2022, the release of Open AI's chatbot ChatGPT marked a significant transformation in the integration of artificial intelligence (AI) within educational contexts, ushering in a new era of AI-supported education. Shortly thereafter, scholarly discourse began to unfold, examining the potential advantages and dilemmas associated with using this chatbot for educational purposes (e.g., Haleem et al., 2022; Sharma & Yadav, 2022). The new AI-assisted era has been recognised as a period of exploration and uncertainty, where the full extent of the latest technological advancements remains to be fully understood. Central to this discourse is the role of chatbots in educational settings, which is often debated in terms of their fundamental value: whether they are perceived as beneficial tools that enhance learning or as unnecessary disruptions to both educational processes and assessment practices.

The integration of AI in student learning has caused concerns within educational institutions, particularly around the evaluation of student output (Cotton et al., 2023). A primary concern is that reliance on AI tools may result in a superficial understanding of academic material, with students potentially depending on AI for task completion rather than engaging deeply with the subject matter. Additionally, if it is not evident that students have independently completed their work, this ambiguity could compromise the integrity of the assessment process, raising questions about the depth of students' academic understanding. Delegating tasks to AI tools without comprehending the underlying processes may lead to shallow learning experiences.

Concerns also arise regarding academic dishonesty, where students might utilise AI to plagiarise or replicate work without proper citation, thereby compromising academic integrity. Furthermore, there is apprehension that AI might replace the development of specific skills crucial for students' personal and intellectual growth. However, recent reviews of AI teaching and learning underscore the importance of appropriate tools to support pedagogical approaches that scaffold advanced cognitive progression in students (Ng et al., 2023).

Identifying and implementing pedagogies that promote deep understanding and enable the evaluation of students' independent contributions to completion is essential to address these concerns and ensure an equitable assessment reflective of students' true capabilities. Therefore, educators face the challenge of balancing the effective utilisation of AI's strengths while simultaneously ensuring an authentic evaluation of students' effort and learning. The foundation for the current article is rooted in an epistemological approach inspired by constructivist theory focusing on the intersection of AI with pedagogy. Its objective is to explore how pedagogical methods support the development of effective and meaningful learning activities, outcomes (Hassan et al., 2022), and assessments (Biggs & Tang, 2007).

While AI-assisted chatbots present both challenges (Tlili et al., 2023) and advantages (Kasneci et al., 2023; Zhai, 2022) in educational contexts, their introduction significantly transforms students' learning processes and skill development. With the increasing availability of advanced technology, chatbots are consistently being adopted as tools to foster student learning and engagement. However, it is important to acknowledge that no technology is inherently value-neutral, and technological euphoria often blinds us to the myriad of unexpected and unintended consequences of its use (Selwyn, 2011).

As a result, many educators hold reservations about integrating AI tools, stemming from a lack of critical reflection and the necessary design work to align these technologies with the core principles of education (Williamson, 2019). Unfortunately, while it is widely recognised that specific technologies shape the practices and potential of engaged learners, critical reflections on the pedagogical and ethical implications of implementing learning technologies are relatively scarce (Williamson, 2019; Zawacki-Richter et al., 2019).

Through a meticulous analysis of the benefits and drawbacks associated with the employment of AI chatbots from a student-centric perspective, this article delves into the application of the ChatGPT chatbot within educational environments. It presents a pedagogical framework that has been developed from the analysis of empirical data and discusses strategies for the effective integration of ChatGPT to positively reinforce the role of AI in educational settings. Consequently, the research question guiding this article is as follows:

Research question: How do students perceive the role of educators in facilitating their use of ChatGPT to enhance learning and the evaluation of learning outcomes?

Addressing this question requires a deep understanding of students' viewpoints regarding teachers' instructional methods, student learning, and the evaluation of learning outcomes. This approach is necessitated by our adoption of a constructivist perspective on learning. Therefore, an examination of the construction of instructional practices using ChatGPT cannot be conducted without incorporating the viewpoints of students, as they are integral to the learning process and actively participate in shaping instructional activities. We thus acknowledge that the learning process is neither solely an individual task nor solely determined by instruction. Students collaboratively shape and construct their learning environment alongside educators, employing ChatGPT as a tool in this process. Given that our research question explores the notion of scaffolding, it is also relevant to clarify our understanding of this concept. Instructional scaffolding refers to the support provided to learners, enabling them to accomplish tasks that would be unattainable without such assistance. Scaffolding enhances learners' independent capabilities and is employed to assist them as required (Driscoll, 2000).

In answering the research question, we will employ suitable methodologies, which will be elaborated in the subsequent description of the methods used.

Methodology

The study aims to address the research question using a specific case study where ChatGPT has been implemented to examine its alignment with learning objectives. The primary objective is to develop conceptual frameworks that integrate ChatGPT as a vital component of the learning experience. The case study focuses explicitly on a teaching course that allows students to consciously shape their academic path by immersing themselves in a selected profession-relevant topic and gaining expertise in that area. The ultimate goal is to enable students to independently contribute to the professional and methodological advancement of their selected profession through research-informed knowledge pertaining to one or more relevant topics.

The case chosen for this study is the Architectural Technology and Construction Management (ATCM) programme at University College of Northern Denmark (UCN), deemed paradigmatic for its potential to set a precedent in utilising ChatGPT to assist students' academic writing processes. The case focuses on architectural and construction management education classes during their third, fourth, and fifth semesters. It has been selected as particularly illustrative for demonstrating the impact of introducing a technology such as ChatGPT, given the emphasis on diverse writing formats within the students' educational and professional journey as construction architects. Furthermore, these students are accustomed to engaging with new technology in various forms, ranging from digital systems to technical drafting. This familiarity with technological innovation enables them to articulate their experiences adequately with such technologies, thereby enhancing the potential for effective communication of their experiences to the research team. The selection of students for the study was based on criteria ensuring a broad representation from the classes and concrete experience with using ChatGPT in their coursework.

According to Flyvbjerg (2006), the selection of case studies should enable researchers to explore events and phenomena in holistic and real-life contexts with minimal control. Thus, case studies are an effective means to capture social processes and evolutions, making them a suitable methodology for mapping complex scenarios where individuals and artefacts interact. They are, therefore, particularly suitable for investigating the students' interaction with and use of chatbots, providing various data collection methods and opportunities for methodological triangulation. Yin (2013) proposed a structured approach to case study research, encompassing the design, data collection, data analysis, and reporting of findings. However, it is important to note that not all data may be recorded in writing, and the final analysis is subject to professional judgement based on available resources. Case stories thus act as empirical evidence, constructing a history centred around a pivotal plot or primary categories analysed through the content structure (Yin, 2013).

This case study adheres to the following steps:

- Case selection: Six cohorts, each comprising approximately 30 students from the ATCM programme during their third, fourth, and fifth semesters, have been selected.
- Data collection: Data were gathered through two focus group interviews, all of whom had extensively utilised the chatbot in their assignments. The data should align with the research question and offer a rich and detailed case description.
- Data analysis: The collected data were methodically analysed to identify patterns and themes. This process involved coding the data, pinpointing crucial concepts or themes, and examining the interrelations between case elements.
- Writing up: The case study findings are evaluated, articulated concisely, and related directly to the research question through a writing process. This includes providing a comprehensive description of the case, an in-depth analysis of the data, and a discussion concerning the implications of the findings.

Data collection predominantly focuses on the students' reflective experiences, perceptions, and emotional responses regarding ChatGPT usage. These reflective experiences are discussed and verbalised through conversations within focus group interviews. The data collection process is primarily exploratory, with an emphasis on comprehending and explaining the phenomenon. The analytical goal is to build theoretical constructs based on empirical evidence, achieved through an ongoing thematic coding process that combines various components.

Thematic coding was used to identify and categorise patterns, themes, and concepts in the qualitative data analysis (see Table 1). This method comprised a thorough examination and review of the data to identify recurrent patterns, topics, and notions, followed by the creation of codes or labels to systematically categorise and structure these identified themes. These codes may be descriptive or interpretive, depending on the desired level of abstraction. The thematic coding procedure is instrumental in identifying similarities and variances within the data, thereby facilitating the generation of insights, the development of theoretical frameworks, and the formulation of conclusions (Boyatzis, 1998; Saldaña & Omasta, 2016).

Table 1.
Overview of the thematic code categories retrieved from the empirical data.

Code categories	Number of statements	Code description	Axial coding	Description of axial codes
Critical approach	14	Students demonstrate awareness of AI's limitations and articulate a critical approach to knowledge acquisition and interaction with the chatbot.	The students' statements suggest a need for cognitive support in fostering transparency, ethical considerations, and a critical attitude in their learning.	**Cognitive perspective** This involves considering issues related to transparency, ethics, and a critical attitude.
Lack of knowledge	4	Students highlight the chatbot's limited knowledge in their professional area, necessitating a critical approach.		
Source validation	3	In this category, the importance of validating the sources provided by the chatbot is emphasised.		
Transparency and ethical filters	5	Students express the importance of transparency in the chatbot's knowledge sources and the application of ethical filters affecting its responses.		
Conversation partner	8	AI is used as a conversational partner, described as a professional partner in the learning process that enhances autonomy.	The students' statements indicate a requirement for pedagogical support, emphasising the importance of diverse learning activities that guide their educational journey.	**Pedagogical perspective** This includes a focus on aspects such as engagement and training, along with the provision of necessary resources and support.
Writer's block	2	Students discuss using AI to overcome writer's block.		
Context comprehension	7	This code requires students to comprehend the context when using AI.		

Brainstorm	4	In these statements, students discuss how AI can be used as a tool or partner for brainstorming.	The students' statements indicate a requirement for pedagogical support, emphasising the importance of diverse learning activities that guide their educational journey. This approach is notably practical, designed to foster engagement and enrich learning experiences while also providing essential resources and support.	**Pedagogical perspective** This includes a focus on aspects such as engagement and training, along with the provision of necessary resources and support.
AI as Google Translate	3	Students discuss their use of AI for translating articles, comparable to their utilisation of Google Translate.		
AI as a text rewriter	10	This category includes descriptions of how students use AI to reformulate existing text, aiming to improve their written presentation.		
AI as a search engine	10	Students describe using AI as a tool to search for literature based on theories.		
Analytical exploration	1	They describe using AI to conduct an analysis based on collected data or literature.		
Introduction to AI	10	These statements emphasise the importance of an introduction to AI from the teacher, which is crucial for developing a critical approach.	The students' responses suggest a necessity for an enhanced educational approach where a well-defined framework and criteria for engaging with AI are established, enabling effective assessment of student learning.	**Educational perspective** This involves considering issues related to the broader educational context in which chatbots are being used, including the specific goals and objectives of the course or programme, the students' needs and abilities, and the resources and support available.
AI as a time-saving tool	4	The statements illustrate AI's role as an efficiency enhancer in learning activities, notably in literature research, awaiting teacher guidance, and data management tasks.		
AI as a personal supervisor	10	Students describe how AI assumes the role of a supervisor in their learning process, where they can seek specific clarifications on theoretical concepts and methodologies.		

Problem-solving engagement	4	The students' accounts reveal how engaging with complex problems fosters the development of distinct and personalised work, even when integrating the GPT chatbot.		

Table 1 illustrates the frequency with which these thematic codes are addressed in the empirical data and includes an in-depth description of the axial coding process.

Analysis of the empirical data

In the following section, we present an analysis of selected empirical data, utilising a thematic coding framework as outlined in Table 1, alongside the number of statements. The analysis is based on interview data sourced from students in their fourth and fifth semesters. The focus is on code categories that are represented by 10 or more statements in the empirical data. This selection criterion is necessitated by the page limitations of this article, which prevent a comprehensive analysis of all themes. Additionally, the analysis occasionally encompasses multiple code categories due to their interrelated nature.

AI for rewriting and feedback: The imperative of a critical approach

The initial statement evaluates the utility of the AI-powered language model, ChatGPT, in enhancing students' writing abilities and increasing the professionalism of their expressions. The students acknowledge its proficiency in offering alternatives for vocabulary and phrasing, which aids in refining their writing skills towards a more professional and academic style. In the following statement, a student elaborates on their propensity to use colloquial language and the desire to refine their writing to a more formal academic level:

> ” [...] Then you say, 'Would you be kind to rewrite this in neat academic language?' Then it actually does that, and the more times it does that for you, the more naturally you read it through and begin to reflect on the text that you have written yourself. You consider what words you use, sentences, and other details. In this way, one's own language becomes nicer, as it's not just spoken language.

This reflection indicates that requesting ChatGPT to provide more academic rephrasing aids in developing writing skills. Engaging with examples of academic language prompts the student to critically evaluate their style, choice of words, and sentence construction, thus refining their written expression to be more sophisticated and distinct from spoken language. Additionally, another quote explains how ChatGPT helps diversify their writing to avoid repetition, thereby enhancing their conceptual understanding:

> ” Well, can you rephrase this so that you might have the opportunity to do it instead of just continuing to repeat yourself? So, it develops – what should one say – the scope of what you can write about, I think.

By using ChatGPT for formal rephrasing, students encounter new vocabulary and sentence structures and learn writing conventions not previously familiar to them. This exposure is crucial for authentic learning relevant to real-world applications. Moreover, the feedback and guidance from ChatGPT facilitate a dynamic process of writing and revision, which is essential for deep learning and mastering the complexities of effective writing. Collectively, these insights demonstrate the significant potential of AI-powered language models in not only refining writing skills but also in equipping individuals with communication skills vital for real-world scenarios.

AI as a search engine: Streamlining source selection and providing guidance in academic research

In the following excerpt, the students discuss their utilisation of the ChatGPT language model to refine their research focus and identify pertinent sources. They characterise the model as a valuable guide that aids them in identifying a range of possibilities related to specific topics, providing them with an overview of different theories and methods for further exploration:

> ” [...] So, we've mostly used it to narrow down instead of scouring the internet empty for theories and methods. You can ask, 'Okay, what would be suitable here?' and then it comes up with 7–8 different options. At the same time, you can get pros and cons for each if you request it. It's like having someone with insight advise, 'We should use this option because of this reason, or choose that one because it fits better.' So, it saves a lot of time when you need to investigate which sources or methods to use for different things.

One significant benefit of utilising ChatGPT is its ability to streamline the research process for students by efficiently identifying relevant sources and offering a range of potential pathways for exploration.

> ” Yes, but I think it’s kind of the same thing, meaning it can guide you onto the right track. In the direction that… it’s quite good at that.

Additionally, the students articulate that ChatGPT helps them focus their research and select pertinent information sources. While they do not rely on ChatGPT as a direct information source, they value it as a tool to navigate towards relevant materials for their studies. Nevertheless, it is crucial for students to critically assess the reliability and credibility of the resources suggested by ChatGPT. As a language model based on machine learning algorithms, ChatGPT generates responses informed by the input it receives.

> ” […] And we’ve also found out that it’s quite good at providing sources for various things, but you do need to be critical of what it comes up with because it’s not certain that you can find that source. Or sometimes it might be a bit more wiki-like.

> ” […] Yes, I think it’s important to teach the students that they can’t just blindly go along with what it says, but that they can use it if they’ve hit a wall and that they themselves need to be able to find information about the topic they’re writing about.

> ” […] So we’re constantly critical of what we read because, well, we at least experienced an example where it said something that was incorrect. So, everything it writes, in one way or another, we need to be critical of, and we are.

The students thus emphasise the significance of adopting a critical approach, which is a crucial aspect of authentic learning. Their reflections include instances where they critically evaluated and identified inaccuracies in the content provided by ChatGPT. Despite acknowledging the utility of ChatGPT in providing useful information, the students nevertheless recognise the imperative of critical assessment and source verification. Furthermore, they emphasise that learners should proactively pursue knowledge rather than relying solely on ChatGPT’s insights. The students advocate for educational approaches that instil critical thinking towards sources and foster independent information-seeking skills, cautioning against the use of ChatGPT as merely a “copying machine” in the learning process.

> [...] Well, I would say that it brings a bit more responsibility for learning because, if you don't care about getting the right or wrong answer, you can just take what the bot says and use it for any purpose. But it's important to be critical of it and also investigate your studies yourself, so it doesn't just become a copying machine because we know we won't learn anything that way. So, I think by using it, it's also important to be very aware that you need to understand what it's saying and be critical of it, not just copy it because you'd also like to learn something.

The statement underscores the necessity for students to adopt a critical stance and independently seek knowledge. These competencies are essential to authentic learning, which prioritises the practical application of knowledge and skills while fostering the development of students as lifelong learners. The student's analysis of ChatGPT's limitations, coupled with the emphasis on analytical evaluation of sources, aligns with authentic learning and assessment principles.

AI as a personal supervisor: Integrating AI support with an analytical mindset

In the following quotes, the students reflect on the use of ChatGPT in comparison to a human advisor or mentor, acknowledging that the chatbot serves effectively as a guide in understanding topics or concepts, as well as its efficacy in delivering specific responses to inquiries:

> [...] It's, you know, easier and faster to ask ChatGPT than to write to a supervisor on Teams, as that supervisor has to go through 10 other groups spread over two semesters. So, it's much faster because then you can move on.

> [...] Personally, I also feel that you get more out of it by asking the bot. Because then you've sort of found the way there yourself. If you figure out how to use it as we've done, well, then you have the right keywords to get there. But if you simply ask a supervisor without trying and they give you all the answers, it doesn't help you in the way you should be researching because the issue of not knowing can come up again. So, it's helpful that you can use the bot to find it yourself instead of having to ask a supervisor every time there's a problem.

However, the students observed that responses from ChatGPT might not invariably be entirely accurate, which necessitates the essential task of double-checking the information provided. Conversely, human advisors are highly regarded for their capacity to provide comprehensive clarification of concepts and offer personalised guidance. Although ChatGPT can provide direct answers, the students highlighted that human advisors are more adept at delivering detailed explanations and guidance for a deeper understanding of the subject matter. They can aid learners in understanding the reasoning behind their recommendations and offer additional resources. Human advisors are also trusted sources of information due to their educational qualifications and professional experience. However, their availability can be limited, potentially leading to delays for learners seeking immediate answers to their questions.

> ” […] Well, I would say it can't be used exclusively as a supervisor because it doesn't have that experience-based aspect that a supervisor who has been working with it for many years has. I mean, there are clearly some personal experiences, whereas this one just goes in and provides an answer.

> ” […] You trust a supervisor more, whether it's one-to-one or in a group setting. Partly, it's because they have, well, a solid educational background. You could say you trust them almost blindly with what they say.

Another aspect the students highlighted was the utilisation of ChatGPT to generate ideas, examples, and information related to their topics. The statement indicates that ChatGPT was found advantageous for brainstorming, as it offered a diverse range of ideas and theoretical perspectives that could be used in their subsequent academic endeavours. Nevertheless, it is advisable to integrate ChatGPT with traditional sources such as books and articles to foster a more comprehensive understanding of the subject matter.

> ” Yes, I also think that when we did some research and asked it to generate different problem statements for us, it could give us an insight into some things we hadn't thought about, like coming up with different ideas for it. And yes, it also helped a bit with finding some sources, though not all were useful, but it could tell us where and in which books we could read certain things and such.

> ” Sometimes you feel like you're hitting your head against a wall where you don't really know how to move forward, right? And it's good for just asking, for example, 'How do I move on from here? What am I missing?' And then it spits something out, and maybe from all the things it spits out, you find out, 'Aha, that was actually a pretty smart way to go.' It's a bit faster in that sense.

Overall, the students' experiences suggest that ChatGPT was instrumental in boosting creativity, generating diverse ideas, and expanding their comprehension of various topics. However, the effectiveness of interaction with ChatGPT depends significantly on the ability to construct the relevant context. Authentic learning, which encompasses engaging students with real-world challenges and complexities beyond the traditional classroom, is fundamental in this process. In the context of interacting with ChatGPT, the statement implies that students cannot merely input indiscriminate questions and anticipate accurate responses. It is imperative for them to learn the skill of formulating pertinent questions and constructing a relevant context for their inquiries. Proficiency in this context-building process is vital for ensuring that ChatGPT delivers precise and reliable information.

> ” Yes, that's also what we've found out, figuring out how to write to it. You can't just write anything. It's about finding out how to ask questions. Yeah, it's about figuring out how to ask it in order to get the right information.

The effectiveness of ChatGPT, as previously discussed, depends largely on the user's capacity to provide context and formulate appropriate questions. These parallels bring to the forefront a significant educational challenge: aligning technological tools with educational objectives to foster meaningful and measurable outcomes. Similar to how students must master the art of questioning and context creation with ChatGPT, educators are also tasked with refining their methodologies to successfully integrate AI tools within their teaching practices. As we further explore the data, the concept of "constructive alignment" emerges as a key framework for steering the research. This concept will be explored in greater detail in the subsequent section.

The role of AI chatbots in the constructive alignment of teaching and assessment

The fusion of ChatGPT and AI in educational settings is an emerging trend. Educators should, therefore, contemplate the optimal and ethical application of these tools (Jensen et al., 2022; Tromp et al., 2011). One particular challenge in this domain is the achievement of constructive alignment between student learning objectives and assessment practices. The concept of "constructive alignment", formulated by Biggs (1996), explains the connection between the established learning objectives and their corresponding assessments. As ChatGPT and various AI technologies increasingly influence how students acquire and refine knowledge and skills, it becomes essential to foster constructive alignment in reevaluating the underlying assumptions. This entails a critical review of (1) the educational objectives and (2) the methodologies employed to achieve these goals (Lent, 2017).

Incorporating AI within educational frameworks necessitates addressing both the challenges and opportunities it presents. Designing and scaffolding a real-world context for students is crucial in establishing a constructive alignment between assessment activities and learning objectives, even when chatbots are utilised in the learning process. For instance, programming ChatGPT to simulate interactive discussions or problem-solving exercises could enable students to demonstrate their acquired knowledge and skills in a more dynamic and authentic manner. It is thus essential to thoughtfully consider the learning objectives and ethical implications associated with the use of ChatGPT and other AI technologies. This requires ensuring that the use of chatbots not only supports student learning and engagement but also upholds ethical standards and respects student autonomy. Additionally, the assessment of students' engagement with ChatGPT must be conducted with reliability and fairness. Such critical reflections highlight the necessity of addressing ethical, logistical, and practical considerations when utilising ChatGPT in educational contexts.

Traditionally, assessment in learning has been a tool for regular grading, where teachers' evaluations are predominantly informed by tests or examinations based on standardised criteria (Harlen, 2005). However, with the advent of sophisticated technologies like ChatGPT and other AI systems, the conventional concept of "assessment of learning" becomes insufficient, as a summative form of assessment requires clear expectations and a sole focus on content (Carver-Thomas & Darling-Hammond, 2017; Gasior, 2013; Harlen, 2005; Hughes, 2014). The potential of AI to generate educational content poses a challenge to the traditional summative approach of fairly evaluating students' knowledge and skills. Consequently, there is an emerging need to explore and integrate contemporary assessment paradigms such as "authentic", "performance", "alternative", "ipsative", and "sustainable" assessments. The purpose of these approaches is to complement traditional methods, with the goal of more effectively capturing learning scenarios in an AI-enhanced educational environment.

These novel assessment concepts share a common characteristic: they emphasise assessment *for/as* learning, which provides a more formative character that aims to promote reflection by mapping the development of learning processes. Essentially, assessment for/as learning is based on the teacher's feedback, focusing on improving the students' learning potential and their capacity for *"self-awareness of learning needs, self-regulation of the learning process, and lifelong autonomous learning and self-determination in decision-making from an ecological and socially responsible perspective"* (Rodríguez-Gómez & Ibarra-Sáiz, 2015, p. 5). The theoretical debate surrounding assessment in education has conventionally centred on two distinct paradigms: *"assessment of learning"* and *"assessment for/as learning"*. This delineation is essentially a choice between formative and summative assessment methodologies (William & Thompson, 2017). According to William (2018), the distinctions between "assessment of", "for", and "as" learning can be succinctly summarised as follows:

Assessment for/as learning is fundamentally supportive in nature.

Assessment of learning, conversely, focuses on measurement.

While assessment for/as learning involves descriptive processes, assessment of learning typically relies on scoring.

Utilising *assessment for/as learning* approach equips students with comprehensive support in navigating both the formal and informal aspects of their learning process. This approach enables deeper analysis and reflection on the connection between their values, experiences, and the development of practical understanding and beliefs. It is imperative that assessment activities accurately reflect the influence of students' behaviours and normative understandings on their ability to develop professional judgement and establish credibility. Within this framework, "authentic assessment" emerges as a particularly effective strategy. This form of assessment involves engaging students in real-world tasks that demonstrate their knowledge and skills, aimed at evaluating students' capacity to apply their learning in genuine, significant settings rather than just memorising facts or completing tasks. Authentic assessment often requires students to produce a product or undertake a project, often incorporating elements of collaboration, problem-solving, and communication skills. However, simply acknowledging the benefits of authentic assessment is insufficient. There is a pressing need to integrate theoretical knowledge with practical strategies, particularly in utilising emerging technologies like chatbots within educational frameworks. In this context, the principle of constructive alignment becomes essential, serving as a foundation for optimising the use of these technologies in education.

By using *assessment for/as learning,* students gain robust support in addressing both their formal and informal considerations. This approach aids them in analysing and reflecting on how their practical understanding and beliefs are influenced by their values and experiences. Specifically, assessment activities should mirror how a student's behaviour and normative understandings impact their capability to forge professional judgment and credibility. Authentic assessment stands out in this respect. It tasks students with real-world challenges to showcase their knowledge and skills. Rather than mere recollection of facts or task completion, it demands application in genuine, impactful contexts, encompassing facets like collaboration, problem-solving, and communication.

A framework for scaffolding student learning and engagement using AI chatbots

In the integration of AI chatbots within an authentic learning environment, the concept of constructive alignment becomes paramount. Achieving this alignment necessitates the meticulous definition of learning objectives and the thoughtful design of both learning activities and assessment tasks. This process involves navigating pedagogical challenges and framing learning and assessment tasks based on three distinct perspectives. These perspectives, as revealed by empirical data and analysis, include cognitive considerations, a pedagogical lens, and the specific educational context (see Table 2).

Table 2.
An overview of code categories and associated learning activities.

Code categories	Learning activities
Cognitive perspective: This involves considering issues related to transparency, ethics, and a critical attitude	
Critical approach	Engage in critical thinking Reflect on personal biases and assumptions
Lack of knowledge	Promote transparency in learning Provide guidelines Provide comprehensive information Ensure ongoing feedback
Source validation	
Transparency and ethical filters	
Pedagogical perspective: This involves considering issues related to engagement and training, resources, and support	
Conversation partner	Facilitate human interaction
Context comprehension	
Writer's block	Encourage asking questions Ensure the availability of additional resources Provide feedback on progress
Brainstorm	
AI as Google Translate	
AI as a text rewriter	
Analytical exploration	Engage in project-based learning activities Develop and enhance problem-solving skills
Educational perspective: This involves considering issues related to the broader educational context in which chatbots are being used, including the goals and objectives of the course or programme, the needs and abilities of the students, and the resources and support available.	
Introduction to AI	Provide student support and guidance Address technical considerations

AI as a time-saving tool	Monitor students' use of the chatbot
AI as a time-saving tool	
Problem-solving engagement	Integrate AI into the curriculum Facilitate students in setting their learning goals

Table 2 presents an overview of code categories and associated learning activities from cognitive, pedagogical, and educational perspectives.

From the perspective of authentic learning, using a chatbot enables learners to engage with authentic tasks and receive real-time feedback. The students' positive experiences with the chatbot indicate that such tools can effectively familiarise learners with technologies increasingly prevalent in contemporary professional environments. Regarding assessment, the chatbot emerges as a tool for formative evaluation, offering learners ongoing feedback on their progress.

The following section introduces a pedagogical framework (see Table 3) developed through the examination of empirical data and theoretical discourse surrounding ChatGPT. This framework aims to constructively scaffold the integration of AI within higher education settings.

Table 3.
A pedagogical framework for scaffolding the use of AI in higher education.

Code categories	Code categories	Code categories
Cognitive perspective: This involves considering issues related to transparency, ethics, and a critical attitude		
Provide guidelines	Provide students with guidance on effective ChatGPT usage, set boundaries to prevent over-reliance, and encourage the development of critical thinking and problem-solving skills.	It is essential to evaluate whether students' usage of ChatGPT avoids creating underlying value conflicts and sustains ethical engagement.

Provide comprehensive information	Inform students about ChatGPT's capabilities and limitations, ensuring transparency about its role in the learning process.	Educators must evaluate students' understanding to mitigate technological euphoria and enhance awareness of the potential unintended consequences of using ChatGPT.
Promote transparency in learning	Encourage students to log all interactions with ChatGPT, promoting transparency in data or knowledge generation and collection.	Clarity in the assessment process can be achieved by offering transparent insights into students' strategies for interacting with ChatGPT to achieve learning objectives.
Engage in critical thinking	Engage students in problem-solving tasks that foster critical evaluation of the information and guidance from ChatGPT.	There is a need to support students in independently evaluating their learning progress, particularly through problem-solving activities that enhance critical thinking abilities.
Ensure ongoing feedback	Students should proactively address potential issues through continuous feedback and corrective actions, while educators support students in understanding how ChatGPT impacts their learning.	Fostering an environment of trust and fairness is crucial, ensuring that ChatGPT is used as an educational tool rather than a mere solution provider.
Reflect on personal biases and assumptions	Encourage students to consider their biases and assumptions in light of the information provided by ChatGPT.	Students should be encouraged to develop an open-minded approach and a willingness to consider diverse perspectives, enhancing critical understanding of explored topics.
Pedagogical perspective: This involves considering issues related to accessibility, motivation, engagement and training, resources, and support		
Engage in project-based learning activities	Guide students to use ChatGPT in project-based learning, encouraging the creation of projects or products rather than direct solution-seeking from ChatGPT.	Ensuring that students harness ChatGPT to deepen understanding and assist in problem-solving rather than merely obtaining solutions is crucial for optimising the educational experience.
Develop and enhance problem-solving skills	Encourage students to employ ChatGPT in solving problems.	There is a need to empower students to showcase their capacity to effectively utilise ChatGPT, demonstrating the application of their learning in a practical, problem-solving context.

Encourage asking questions	Encourage students to critically evaluate the information provided by ChatGPT by formulating and asking pertinent questions. They should consider the accuracy, relevance, and context of ChatGPT's responses through critical inquiry and cross-reference with other sources.	Asking questions is a powerful tool for promoting authentic assessment and learning. Encouraging learners to ask questions aids them in developing critical thinking skills, promoting active learning, supporting self-assessment, providing feedback, and fostering a sense of curiosity and exploration.
Provide feedback on progress	Collaborate with students to establish clear objectives for utilising ChatGPT and deliver consistent feedback on their progress.	Feedback and guidance on a student's progress support authentic learning by clarifying expectations, identifying areas of strength and weakness, encouraging reflection, supporting goal setting, and contributing to formative assessment processes. Feedback thus helps students develop a deeper understanding of the material and achieve their learning goals.
Facilitate human interaction	Ensure that ChatGPT usage complements rather than replaces human engagement by creating opportunities for students to consult and collaborate with teachers, mentors, or peers.	When using a chatbot, human interaction is crucial for providing emotional support, addressing misunderstandings, promoting critical thinking, offering feedback, and supporting differentiation. By incorporating human interaction into a ChatGPT learning experience, teachers and instructors can create an engaging, supportive, and effective learning environment for all students.
Ensure the availability of additional resources	Offer comprehensive support to students, guiding them towards additional resources and information when ChatGPT's assistance is limited or insufficient.	Providing a variety of resources when using ChatGPT underpins authentic assessment and learning by promoting exploration and encouraging critical thinking. Teachers and instructors can help students develop a deeper understanding of the content by providing access to a wide range of resources and information.

Educational context: This involves considering issues related to the broader academic context in which chatbots are being used, including the goals and objectives of the course or programme, the needs and abilities of the students, and the resources and support available.		
Integrate AI into the curriculum	Align the integration of ChatGPT with the curriculum and course objectives, ensuring its inclusion in lesson plans and assessments to allow students to practise and demonstrate competencies.	Curriculum alignment enhances authentic learning with ChatGPT by focusing on learning outcomes and fostering coherence. Aligning ChatGPT usage with the curriculum enables educators to foster a learning environment that is both effective and engaging.
Monitor students' use of the chatbot	Monitor student engagement with ChatGPT to confirm it complements classroom participation and other interactive forms of learning rather than replacing them.	Monitoring students' use of ChatGPT offers critical insights into their learning process. It serves as a valuable tool for educators, enabling them to identify students' challenges, offer specific feedback, and promote self-reflection, thereby supporting overall student learning and achievement.
Facilitate students in setting their learning goals	Encourage students to set their learning objectives by leveraging the resources provided by ChatGPT, promoting self-directed and goal-oriented education.	Setting learning goals enables students to enhance their sense of independence.
Provide student support and guidance	Provide essential support and resources for chatbot usage, which includes access to training, documentation, and troubleshooting assistance, as well as online guides and dedicated technical support.	Providing students with adequate resources and guidance for using ChatGPT can support authentic assessment and learning by promoting access, engagement, exploration, differentiation, and reflection. By ensuring students receive the necessary support for engaging with ChatGPT and achieving their learning goals, educators can establish a learning environment that is effective and equitable, fostering engagement for all students.
Address technical considerations	Address technical considerations related to the use of ChatGPT, focusing on data privacy, security, and bias mitigation, and ensure its inclusivity and accessibility for all students.	When using ChatGPT, technical considerations play a critical role in promoting an authentic learning environment by ensuring that ChatGPT is reliable, accessible, secure, integrated, and personalised.

Human interaction	The use of ChatGPT should not replace human interaction but supplement and enhance it, so it is important to consider opportunities for students to interact with actual human experts or peers.	When using a chatbot, human interaction is important to provide emotional support, address misunderstandings, promote critical thinking, provide feedback, and support differentiation. By incorporating human interaction into a ChatGPT learning experience, teachers and instructors can create an engaging, supportive, and effective learning environment for all students.

Table 3 provides a comprehensive pedagogical framework designed to scaffold the effective integration of AI tools, such as ChatGPT, into higher education.

Conclusion

The conclusion on establishing constructive alignment between learning objectives and assessment when integrating AI chatbots encompasses three distinct dimensions: cognitive, pedagogical, and educational, as outlined in Table 3. Analysis and development of the pedagogical framework indicate that using ChatGPT promotes transparency and autonomy within an ethical framework. Transparency is a key feature, as ChatGPT provides explanations for its suggestions, alerting students to the necessity of a cautious approach when relying on identified sources. This empowers students with various informed options tailored to their preferences and needs. The autonomous experience is highlighted by ChatGPT's role in enhancing writing and research skills, offering guidance and feedback, and thus empowering students to remain focused on their learning goals. However, it is pivotal to note the importance of critical evaluation in assessing the reliability and credibility of sources suggested by ChatGPT, given its basis in machine learning algorithms and user input.

On pedagogical and educational levels, the utilisation of ChatGPT underscores the need for teaching strategies and methods that align with authentic learning and assessment environments. Regarding motivation and engagement, a pedagogical approach emphasising active learning and student-centred activities is beneficial in maintaining student engagement and motivation. Activities that foster student choice and autonomy, such as project-based learning or collaborative tasks, significantly enhance student engagement when using ChatGPT. By focusing on pedagogy that considers accessibility, motivation, engagement, training, resources, and support, educators can leverage ChatGPT to foster a more inclusive and effective learning environment for all students. Chatbots offer tailored learning experiences by adapting to individual learning styles and preferences. They also provide support to students facing learning barriers, keeping them engaged with timely feedback, practice opportunities, and interactive learning environments. Chatbots play a crucial role in assisting learners to guide their individual educational progression.

References

Biggs, J. (1996). *Enhancing teaching through constructive alignment.* Higher education, 32(3), 347-364.

Biggs, J. & Tang, C. T. (2007). *Teaching for quality learning at university: What the student does* (3rd ed.). McGraw-Hill.

Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development*. Sage Publications.

Carver-Thomas, D. & Darling-Hammond, L. (2017). *Teacher turnover: Why it matters and what we can do about it*. Learning Policy Institute. Teacher Turnover: Why It Matters and What We Can Do About It | Learning Policy Institute

Cotton, D. R., Cotton, P. A. & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 1-12.

Driscoll, M. P. (2000). *Psychology of learning for instruction*. Allyn & Bacon.

Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative inquiry*, *12*(2), 219-245.

Gasior, K. (2013). Education at a glance 2012: OECD indicators. *Sociologický časopis/Czech Sociological Review*, *49*(6), 994-997.

Haleem, A., Javaid, M. & Singh, R. P. (2022). An era of ChatGPT as a significant futuristic support tool: A study on features, abilities, and challenges. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations*, 2(4), Article 100089.

Harlen, W. (2005). Teachers' summative practices and assessment for learning – tensions and synergies. *Curriculum Journal*, *16*(2), 207-223.

Hassan, R., Ali, A., Howe, C. W. & Zin, A. M. (2022, October). Constructive alignment by implementing design thinking approach in artificial intelligence course: Learners' experience. In *AIP Conference Proceedings* (Vol. 2433, No. 1). AIP Publishing.

Hughes, G. (2014). *Ipsative assessment: Motivation through marking progress*. Springer.

Jensen, C. G., Dau, S., Clausen, N. R., Ryberg, T. & Bertel, L. B. (2022). Dilemmaer i translation og transformation af erfaringsbaserede kollaborative lærings-former til digitale læringsplatforme. *Læring og Medier*, *15*(25).

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F. & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences*, 103, 102274.

Lent, J. R. (2017). *The patterning instinct: A cultural history of humanity's search for meaning*. Prometheus Books.

Ng, D. T. K., Lee, M., Tan, R. J. Y., Hu, X., Downie, J. S. & Chu, S. K. W. (2023). A review of AI teaching and learning from 2000 to 2020. *Education and Information Technologies*, *28*(7), 8445-8501.

Rodríguez-Gómez, G. & Ibarra-Sáiz, M. S. (2015). Assessment as learning and empowerment: Towards sustainable learning in higher education. In *Sustainable learning in higher education: Developing competencies for the global marketplace* (pp. 1-20). Springer.

Saldaña, J., & Omasta, M. (2016). *Qualitative research: Analysing life*. Sage Publications.

Selwyn, N. (2011). *Education and Technology: Key Issues and Debates.* Bloomsbury Publishing. http://ebookcentral.proquest.com/lib/aalborguniv-ebooks/detail.action?docID=661054

Sharma, S. & Yadav, R. (2022). Chat GPT – A technological remedy or challenge for education system. *Global Journal of Enterprise Information System*, *14*(4), 46-51.

Tlili, A., Shehata, B., Aderkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R. & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments, 10*(1), Article 15.

Tromp, N., Hekkert, P. & Verbeek, P. P. (2011). Design for socially responsible behavior: A classification of influence based on intended user experience. *Design Issues*, *27*(3), 3-19.

Wiliam, D. & Thompson, M. (2017). Integrating assessment with learning:
What will it take to make it work?. In *The future of assessment* (pp. 53-82). Routledge.

Williamson, B. (2019). New power networks in educational technology. *Learning, Media and Technology*, *44*(4), 395–398. https://doi.org/10.1080/17439884.2019.1672724

Yin, R. K. (2013). *Case study research: Design and methods*. Sage Publications.

Zawacki-Richter, O., Marín, V. I., Bond, M. & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, *16*(1), 1-27.

Zhai, X. (2022). *ChatGPT user experience: Implications for education*. *Available at SSRN 4312418*. ChatGPT User Experience: Implications for Education by Xiaoming Zhai :: SSRN

Abstract

Eye-tracking-teknologien er blevet tilgængelig for måling af elevers læsning i skolen. I en skolehverdag skal eye-tracking-data kunne omsættes til pålidelige mål med praktisk relevans, fx til undervisningsevaluering. Dette studie præsenterer øjenbevægelsesmålet fokusord, der er udviklet til at opmærke nedbrud i elevers flydende læsning og dermed et forsøg på at oversætte eye-tracking-data til læsefaglig indsigt. Studiet er designet til at validere målets pålidelighed som læse-/stave-indikator og endvidere målemetodens robusthed. 84 elever fra fem 4.-klasser deltog, hvoraf data fra 68 elever indgår. Udover elevernes eksisterende resultater i læse-/stavetest, optog vi deres øjenbevægelser under selvstændig højtlæsning. Vi finder stærke korrelationer mellem øjenbevægelser og læse-/stavetestresultater, som indikerer målemetodens robusthed. Videre korrelerer fokusord med klassiske øjenbevægelsesmål, hvilket er tegn på samstemmende validitet. Endelig forklarer fokusordene unik variation i elevernes læseforståelseseffektivitet og stavefærdighed ud over, hvad det stærkest korrelerede klassiske mål, fikseringstid, forklarer. Resultaterne understøtter, at fokusordsprocent er robust og tæt forbundet med læse- og stavefærdigheder.

Availability of eye-tracking technology has increased, also for measuring pupils' reading at school. In a school day, eye-tracking data must be able to be translated into measurements that are both reliable and have practical relevance, e.g. for evaluating teaching. This study presents the eye movement measure focus-words designed to mark breakdowns in students' reading fluency and hereby translate eye-tracking data into reading insights. The study is designed to validate the reliability of the measure as an indicator of reading and spelling proficiency and the robustness of the measurement method in practice. A total of 84 pupils from five fourth-grade classes participated, of which data from 68 pupils was included. Together with the students' existing test scores in spelling and reading comprehension, we collected recordings of their eye movements while they read aloud independently. We find strong correlations between eye movements and reading/spelling test results, which indicates that the measurement method is robust. Further, focus-words correlate with classic eye movement measures, which is a sign of co-validity. Finally, focus-words explain unique variation in students' reading comprehension and spelling proficiency beyond what the most strongly correlated classical measure, fixation time, explains. The results support that focus-word-percentage is a robust metric closely related to literacy.

Måling af flydende læsning med øjenbevægelser i skolen

Eye-tracking kan under selvstændig læsning i 4. kl. måle variation i læse- og stavefærdigheder.

Sigrid Klerke, Det Kongelige Akademi,
& Stine Fuglsang Engmose, Professionshøjskolen Absalon

Indledning

Det er veletableret, at der er en sammenhæng mellem læsefærdigheder og muligheder for at deltage i samfundet. Den internationale læseundersøgelse PIAAC (Rosdahl, Fridberg, Jakobsen & Jørgensen, 2013) peger på, at gruppen af unge og voksne danskere, som ikke er funktionelle læsere, har større sandsynlighed for at nå kortere i uddannelsessystemet, for at stå uden for arbejdsmarkedet, for ikke at tage en efteruddannelse og højere risiko for ringe helbred. Disse resultater bliver bekræftet i den nationale undersøgelse af voksnes basale færdigheder (Larsen, Jakobsen & Rosdahl, 2022). Også en undersøgelse, der følger op på danske unges uddannelsesstatus fire år efter, de som 16-årige har deltaget i PISA-undersøgelsen, finder en stærk sammenhæng mellem læsefærdighed og uddannelsesstatus (Andersen, 2005). Andersen (2005) finder, at unge, som ikke er i uddannelse og heller ikke har gennemført en ungdomsuddannelse i løbet af de fire år, har lavere læsefærdighed. Samtidig er gruppen af unge, som ikke er funktionelle læsere, uændret fra PISA-undersøgelsen i 2009 og frem til 2018. Andelen ligger stabilt omkring 15 % (Christensen, 2019).

Det er i grundskolen, eleverne lærer at læse. Det afspejler sig i kompetence-, færdigheds- og vidensmålene for danskfaget (Børne- og undervisningsministeriet, 2019). Efter 2.kl. er kompetencemålet for læsning, at eleverne kan læse enkle tekster sikkert og for færdigheds- og vidensområdet afkodning er målet, at eleverne ved slutningen af 2.kl. kan læse ord i tekster til klassetrinnet sikkert. Der er altså fra grundskolens start et fokus på at udvikle sikker afkodningsfærdighed, som kan anvendes til at læse tekster. Senere i grundskolen udvides fokus i målene for færdigheds- og vidensområdet afkodning mod, at eleverne også udvikler en hurtig og automatiseret afkodningsfærdighed med umiddelbar ordgenkendelse. I sine bøger "Læsning og læseundervisning" (2016) samt "Læsevanskeligheder" (2021), gennemgår professor Carsten Elbro årtiers international forskning i læseudvikling og -undervisning. Elbro (2016 og 2021) peger på, at direkte og systematisk undervisning i skriftens lydprincip er afgørende for udviklingen af sikker afkodnings- og stavefærdighed, herunder hører omfattende øvelse med at læse tekster af passende sværhedsgrad. Dette gælder i særdeleshed for elever med ordblindhed, som har svært ved at tilegne sig sikker og automatiseret afkodnings- og stavefærdighed (Elbro, 2021). Læsning er dog mere end afkodningsfærdighed. Læsning rummer også det at forstå de tekster, man læser. For at kunne tilpasse læseundervisningen til den enkelte elev og identificere elever, som ikke udvikler deres læsefærdighed, og som måske har brug for mere intensiv læseundervisning, har folkeskolen haft Folkeskolens Nationale Test i Dansk for 2., 4., 6., og 8.kl. Testene har været kritiseret for at give ikke-gyldige informationer om elevernes færdigheder (Ravn, 2015, Kousholt, 2015), og de er ikke længere i brug, men nye nationale test i dansk er på vej (Børne- og undervisningsministeriet, 2022). Ved siden af de obligatoriske test anvender mange skoler supplerende standardiserede pædagogiske og diagnostiske test. Disse kan fx følge elevernes læsefærdighed over tid, pege på særlige nuancer i fx afkodningen, stavningen eller læseforståelsen den enkelte elev er usikker i, og sammenligne den enkelte elevs færdigheder med en normgruppe. Standardiserede test er på den måde et væsentligt redskab i den løbende tilpasning af læseundervisningen til den enkelte elev.

Der er dog også ulemper ved traditionelle, standardiserede test. Man kan kritisere formatet i disse test for at være "langt" fra det arbejde, eleverne normalt laver, når de læser tekster i skolens fag. For at kunne sammenligne elevernes færdigheder med en normgruppe, er de standardiserede test designet med en fastlagt ramme, for hvornår og hvordan de kan gennemføres. Testsituationen er på den måde ikke designet til at ligne almindelig undervisning. Test af denne type er fx født med den begrænsning, at man ved gentestning inden for kort tid ofte vil se, at eleverne scorer højere end første gang, fordi de kender testen[1]. En anden bekymring er, at testsituationen kan være ubeha-

1 Standardiserede test kommer dog ofte med information om test-retest reliabilitet, så man har en fornemmelse af denne øvelseseffekt.

gelig for nogle elever. Dette kan fx ses i en national spørgeskemaundersøgelse med besvarelser fra 768 lærere om brugen af Folkeskolens Nationale Test (Bundsgaard og Puck 2016). Svarene fra lærerne i denne undersøgelse tyder på, at mange lærere oplever, at der er elever, der bliver kede af det under testen, og at testen ikke opleves meningsfuld for de svageste elever. Endvidere er denne type af test ofte opbygget, så de kan indfange variationen i en klasse fx mellem den svageste og stærkeste læsere. Det betyder, at en stor del af eleverne som gennemfører testen vil opleve, at de skal prøve at løse opgaver, som er for svære for dem, eller at de kun når halvdelen af testens opgaver. Det er ikke sikkert, at de elever, som scorer lavt på testen, oplever arbejdet med testens opgaver som et nederlag, men hvis de gør, så kan netop gentagne erfaringer med nederlag påvirke elevernes selvbillede i faget negativt. I en gennemgang af litteraturen om selvbillede og elever med læsevanskeligheder konkluderer Lena Swalander (2012), at der er en vedvarende sammenhæng mellem problemer med læsning og selvbilledet og en øget risiko for, at denne gruppe af elever får negative følelser i forbindelse med præstationer (Swalander, 2012). I en sammenligning af 78 skoleelever med ordblindhed og 77 jævnaldrende uden, finder Alexander-Passe (2008) signifikante forskelle mellem de to grupper, hvor eleverne med ordblindhed oplever sig mere pressede. Det kommer særligt til udtryk i interaktionen med lærere, prøvesituationer og færdighedstest. Eleverne fortæller om angst, generthed, ensomhed, kvalme, rystende hænder og bankende hjerte. Flere studier finder lignende sammenhænge mellem studerende med ordblindhed og angst ved prøvesituationer (Caroll & Iles, 2006, Nelson et. al., 2015). Disse undersøgelser peger på, at der for gruppen af elever i læsevanskeligheder er grund til at tage risikoen for at skabe negative erfaringer med læsning i betragtning, når de skal gennemføre test, der stiller høje krav til læsefærdigheden.

I forskning i læsning og læseudvikling er øjenbevægelser veletablerede indikatorer for kognitiv processering (Rayner, 2012, Holmqvist, et al., 2011). I takt med at eye-tracking er modnet som teknologi, og der er udviklet pålidelige ikke-begrænsende og brugervenlige instrumenter til at optage øjenbevægelser, er interessen også steget for at anvende øjenbevægelser i læseevaluering (Gran Ekstrand et al., 2021, Benfatto et al., 2016, Klerke et al., 2018, Bingel et al. 2018). Forventningen er, at eye-tracking-teknologien kan give nye muligheder i evalueringen af læsning. Særligt kan metoden imødegå nogle af ulemperne ved de standardiserede test, fordi øjenbevægelser kan optages under uformel læsning af egne tekster. Formodningen er, at sådanne mindre formelle optagelser kan bruges til at evaluere læsefærdighed løbende, så læreren ad den vej kan få indsigter, som kan danne grundlag for den løbende tilpasning af læseundervisningen til den enkelte elev.

Traditionelt er eye-tracking-mål i læseforskning blandt andet blevet brugt til at udpege aspekter af skriftsproget som systematisk kan påvirke læseres øjenbevægelser, fx ved at føre til længere eller flere fikseringer på de ord som kræver ekstra processering. Derfor forventer vi i dag at se effekter af ord der er sjældne, lange eller har en overraskende grammatisk rolle. Effekterne ses konsistent på de fikseringsbaserede mål som gennemsnitlig fikseringstid og varigheden af den første fiksering per ord (Rayner, 2012). Et andet klassisk fokus har været at lede efter øjenbevægelsesmål som pålideligt skelner aspekter ved læsere som fx ordblindhed (Benfatto et al., 2016, Franzen et al., 2021). Eksempler på mål der kan bidrage til at skelne forskelle i læsestrategier og -færdigheder, er andelen af fikseringer brugt på at genlæse teksten, andelen af oversprungne ord og hvilke retninger øjnene typisk flytter sig i. Mål, der aggregerer elevens sammenhængende læsning, er især anvendt til dette formål, fordi læseprocesser relateret især til forståelsen fører til en wrap-up-effekt, hvor øjnene ofte dvæler længere ved ord, der står ved sætningsgrænser (Meister, et al., 2022, Rayner, 2012, Franzen et al., 2021). Selvom både klassiske og moderne mål er blevet brugt til at skelne læseres færdigheder og teksters sværhedsgrad, er forskelle i øjenbevægelser en effekt af og ikke en årsag til forskelle i læsning. Det betyder også at det ikke er didaktisk meningsfuldt hverken at forhindre genlæsning eller at forsøge bevidst at ændre på fikseringstider for at styrke selvstændig læsning (Holmqvist et al., 2011, Rayner, 2012)[2]. Ligeledes er målene ikke designet til at tolke noget om det enkelte individ, hvilket har den konsekvens, at det ikke er oplagt, hvordan praktikere kan omsætte varigheden af første fiksering per ord til undervisningstiltag.

Der mangler altså fortsat en praktisk og metodisk modenhed; dels skal det afklares, i hvilket omfang øjenbevægelser, også afspejler læseres læsefærdigheder under hverdagslæsning på egen hånd, og dels er der brug for mål, der kan tolkes uden specialoplæring i brug af eye-tracking. Ideelt set skal målene kunne tolkes og anvendes med den specialviden, læsevejlederen allerede har.

Dette studie tager hul på at adressere begge udfordringer. Først korrelerer vi øjenbevægelsesmål fra forskningslitteraturen målt i en realistisk skolekontekst med standardiserede testresultater for

2 Holmqvist et al., 2011 indeholder en grundig gennemgang af et meget bredt udvalg af klassiske eye-tracking mål, deres ophav, varianter og hvordan de typisk anvendes og tolkes. Rayner, 2012 opsummerer specifikt læseforskning med eye-tracking.

læseforståelse og stavning, som er, set ud fra forlagenes testudbud, standardiserede test, der ofte anvendes, som redskaber til klasse-screening på mellemtrinnet. Videre præsenterer vi en kvantificeret udgave af målet *fokusord*, der automatisk opmærker nedbrud i flydende læsning. Opmærkningen er udviklet, så de udpegede ord tillader læsevejlederen at fortolke disse fokusord i analogi til fejllæsninger i konventionel læseevaluering som fx. *running record* (Wangebo, 2011, D´Agostino et al., 2021). Ligesom fejllæsninger, skaber fokusord et udgangspunkt for evalueringen af en elevs læsning, men i stedet for manuelt opmærkede fejl, udpeger fokusord automatisk de *ord, der bremsede elevens flydende oplæsning* målt med øjenbevægelser. Det er didaktisk meningsfuldt at kunne udpege og afspille de konkrete ord, der kostede eleven ekstra opmærksomhed, for at planlægge og evaluere læseundervisningen. Her undersøger vi om procenten af fokusord, i lighed med fejlprocenten (cf. Nielsen et al., 1992), afspejler læse- og stavefærdigheder.

Vi har operationaliseret projektet i følgende tre forskningsspørgsmål og hypoteser:

1. Hvor robust er den kendte forbindelse mellem læse- og stavefærdigheder og eye-tracking-mål, når elever læser sammenhængende tekster med eye-tracker i et skolevenligt-setup?

Vores overordnede hypotese er, at de klassiske eye-tracking-mål ikke vil være lige robuste i dette setup, hvor der må forventes mere støj sammenlignet med kontrollerede laboratorie-designs, hvor målene ellers normalt korrelerer med læsefærdigheder.

2. Korrelerer eye-tracking-målet fokusord med læse- og staveresultater i samme grad som de klassiske eye-tracking-mål?

Hypotesen er, at det nye eye-tracking-mål, fokusord, afspejler mængden af ord, som eleven har svært ved at afkode, og derigennem afspejler både læse- og stavefærdigheder. Der er flere måder at kvantificere målet fokusord. Vi bruger korrelationsanalysen til at udpege den variant som er numerisk stærkest korreleret med læse-/stavemålene som den mest robuste. Vi vælger altså at afgøre dette empirisk da vi ikke har tilstrækkeligt teoretisk belæg for at favorisere en enkelt variant på forhånd. Dette betyder at vi har ekstra mange sammenligninger som gør det sværere at opnå pålideligt signifikante resultater.

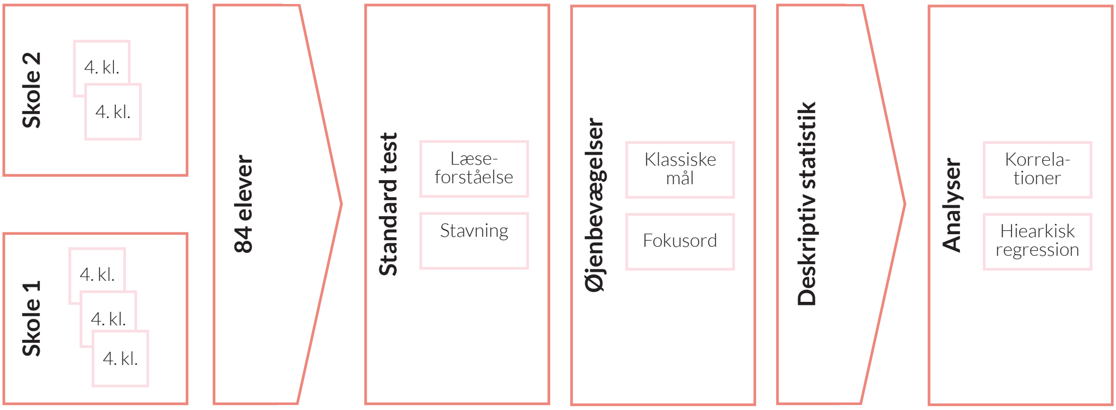
3. Forklarer fokusord unik variation i elevernes læseforståelse og stavefærdighed ud over den variation, der forklares af det traditionelle eye-tracking-mål, der er numerisk stærkest forbundet med læse- og stavemålene?

Her er hypotesen, at fokusordenes særlige følsomhed for afbrud i flydende læsning kan afspejle unikke aspekter af læseforståelse og stavefærdighed, som ikke allerede afspejles i det klassiske eye-tracking-mål, der er numerisk stærkest korreleret med læse- og stavefærdigheder. Hvis det er tilfældet understøtter dette studie, at fokusord er mere og andet end det klassiske eye-tackingmål, og samtidig er en pålidelig læse-/stave-indikator.

Metode

Undersøgelsens design er illustreret i Figur 1. Vi gennemfører korrelationsanalyser og hierarkiske regressionsanalyser med data fra to typer af måleredskaber: Dels standardiserede test af læseforståelse og stavning, og dels optagelser af øjenbevægelser under tekstlæsning (se figur 1).

Figur 1.
Undersøgelsens design



Deltagere

To skoler fra to kommuner deltog med i alt 84 elever fra fem forskellige klasser på fjerde årgang. Vi har inkluderet alle, der deltog både i mindst en af de to standardiserede tests og i læsning med eye-tracker. Læsninger fra tre elever, der har brugt oplæsningsstøtte under læse- eller stavetesten blev ekskluderet. Fire læsninger er ekskluderet, hvor elever har angivet, at de kendte bogen fra tidligere. På baggrund af tre foruddefinerede kriterier for datakvalitet (se bilag 1) blev ni læsninger ekskluderet. I fire af tilfældene var eye-trackingen for upræcis, i fire tilfælde var der udfald i data fra eye-tracking, undervejs og i et tilfælde havde en elev stillelæst. Tabel 1 viser antallet af elever før og efter eksklusion.

Tabel 1.
Antal deltagere før og efter eksklusion (se tekst).

	Læsetest	Stavetest	Eye-tracking
Antal elever deltaget	84	81	84
Antal elever bibeholdt	78	77	68

Gennemførelse
For at minimere negative oplevelser, har vi tilstræbt at indsamle så lidt testdata som muligt. Derfor er de standardiserede test i undersøgelsen netop de test af læseforståelse og stavning, som skolerne alligevel gennemførte, så disse blev gennemført som en del af kommunernes generelle læseindsats 1-3 måneder før læsning med eye-tracker blev gennemført. Øjnbevægelsesdata er indsamlet i maj måned 2021.

Under måling af øjenbevægelser læste eleverne tre tekster højt ved en computer med eye-tracker. Under indsamlingen af øjenbevægelsesdata har vi søgt at minimere elevernes tid væk fra klassen ved at planlægge deres deltagelse sammen med skolens læsevejleder. Eleverne blev derfor i klassen introduceret til at læse e-bøger med eye-tracker via en mundtlig præsentation og en video der viste login, kalibrering og læsning med eye-tracker (intro-video). Eleverne fik ved gennemførelsen af deres læsning hjælp ved tekniske vanskeligheder, men læste ellers højt for sig selv på egen hånd. Hver læsning sluttede efter et fast tidsinterval, hvilket undlod at udstille nogen som hurtigere eller langsommere læsere. Seks bøger var udvalgt til forsøget. Alle elever fik læst de seks bogtitler op, og blev bedt om at udfylde et ark med et billede af hver bog, hvor de skulle sætte kryds ud for de bøger, de tidligere havde læst i. Læsningerne blev gennemført over en dobbeltlektion, hvor 2-4 elever ad gangen læste samtidig i et separat lokale under opsyn. Hver elev gennemførte tre læsninger, en demo-læsning på 1 minut (lix 6), en læsning i en forudbestemt bog i 3 minutter (lix 18) og en læsning i en selvvalgt bog i 2 minutter (lix 11-25). Vi analyserer den forudbestemte læsning på 3 minutter. Det gør vi for at give sammenligningen med de standardiserede test bedst mulige vilkår i denne første undersøgelse af fokusord. Bogen til den forudbestemte læsning var derfor valgt, så teksten matchede den standardiserede læseforståelsestest på ord- og sætningslængde. Dermed er data i denne undersøgelse fra den del af læsesituationen, hvor alle elever læser den samme tekst. Tidtagningen med eye-tracker følger øjenbevægelserne og stopper midlertidigt, hver gang eleven ser væk fra teksten i mere end et halvt sekund. Når den planlagte læsetid er gået, bliver det umuligt at bladre videre i bogen, men eleven vælger selv hvornår bogen lukkes. Det betyder, at hver optagelse varer lidt længere end hhv. 1, 3 og 2 minutter, hvis eleven enten har kigget væk fra teksten, har brugt ekstra tid på at læse sidste side færdig, eller hvis eye-trackerens kalibrering er meget upræcis. Første side i hver e-bog er et billede af forsiden af bogens originale omslag, og første side regnes derfor altid kun med som et billede, der højst kan tælle for et halvt sekund af den samlede læsetid. Eleverne beskrev overvejende opgaven som let.

Det er denne læsesituation vi omtaler, som skolevenligt-setup. Dette skal ses i lyset af det klassiske eksperimentelle setup, hvor læserens hoved fx støttes til at være i samme position under hele læsningen eller næsten identiske stimuli gentages mange gange.

Forud for deltagelse blev forældrene informeret i et kort skriv, og der blev indhentet informeret samtykke fra forældre/værge til deltagelse.

Anvendte skriftsproglige mål

Tekstlæseprøve 5 (Møller, 2013) er en læseforståelsesprøve med fokus på elevens evne til at drage inferenser. Teksten er bygget op af almindelige ord for at mindske ordforrådets betydning. Testen bruger to spørgsmålsformater 1) cloze i teksten, hvor eleven skal vælge det rigtige ord og 2) ja/nej spørgsmål for hver 3. til 4. side. Testen har 44 items. Det er muligt at sammenligne klassen og eleverne med normer og kriteriebaserede kategorier (se bilag 1 for flere detaljer).

Vi bruger tre mål fra læseforståelsesprøven, da vi ønsker at belyse sammenhængen mellem fokusord og mål af læseforståelse der er mere eller mindre følsomme for afkodningsfærdighed. Det første mål er besvarelsestiden, som er det tætteste vi kommer et rent mål af læsehastighed i testen. Det andet mål er antal rigtigt bevarede forståelsesspørgsmål per minut. Vi bruger dette som effektivitetsmål, hvor scoren både påvirkes af, hvor sikker forståelsen er, og også hvor hurtig afkodning og forståelse foregår. Det sidste mål er testens præcisionsmål, der angiver procent rigtigt besvarede forståelsesspørgsmål. Dette mål afspejler sikkerheden i elevens forståelse. Denne sikkerhed kan dog også være påvirket af usikker afkodning, fx ved at langsom afkodning gør det svært for eleven at huske information fra teksten, eller fordi upræcis afkodning giver eleven et usikkert grundlag at forstå teksten på.

Staveprøve 3 (Juul, 2019) er en indsætningsdiktat med 36 items. De er udvalgt til at være almindelige ord med almindelige grammatiske suffikser. Det er muligt at sammenligne klassen og eleverne med normer og kriteriebaserede kategorier (se bilag 1 for flere detaljer). Vi bruger målet andel korrekt stavede ord, som teoretisk set er mest sammenligneligt med elevens afkodningspræcision. Ordafkodning og stavning er stærkt korreleret. Fx finder Caravolas og kollegaer i et longitudinelt studie, hvor de følger engelsksprogede elever fra midten af børnehaveklassen til slutningen af 2.kl., at korrelationen allerede fra slutningen af børnehaveklassen er stærke og at den ved slutningen af 2.kl. er meget stærk med en Pearsons korrelation på 0,81 (Caravolas m.fl., 2001). Færdighederne er altså tæt forbundne, men ikke den samme færdighed.

I begge test rapporterer forfatterne om høj intern konsistens for skalaen. Der er ligeledes foretaget Rasch-analyse, hvor items uden Rasch homogeneity er sorteret fra. Det betyder, at vi for begge test kan regne med, at vi kan sortere personerne på baggrund af deres faktiske dygtighed inden for testens område.

Anvendte øjenbevægelsesmål

Det grundlæggende mål af øjenbevægelser under læsning er en fiksering. Kun et lille område midt i synsfeltet står skarpt nok til at se bogstaver og ord tydeligt. For at afkode en hel tekst, må det centrale synsfelt derfor flyttes rykvis over teksten i et tempo og mønster, der tillader læserens afkodning og sprogforståelse at følge med. Fikseringer under oplæsning varer typisk 100-500 ms, og den nedre grænse for at afkodning af bogstaver kan nå at ske, er ca. 50 ms. Et ryk er en saccade og baglæns saccader kaldes regressioner. Regressioner udgør 10-15% af trænede læseres læsning (Rayner, 1998).

Varigheden af den første fiksering på et ord bliver kodet ind i øjenmusklernes bevægelsesprogram, inden ordet er set. Den første fiksering afspejler derfor den tid, læseren *forventer* at bruge på at genkende det fikserede ord. Når processeringen kræver længere tid end forventet, opsøger læserens øjne opklarende information automatisk og ubevidst. Da signaler til øjenmuskulaturen tager tid at udføre, kan en planlagt saccade nå at blive udført, selvom et ord viste sig at kræve mere opmærksomhed end forventet. Det giver et mønster, hvor udfordrende ord ofte genbesøges indenfor kort tid, for at tillade afkodningen og sprogforståelsen at indhente øjnene (von der Malsburg & Vasishth, 2011). Fokusord er designet til at udpege ord, der kostede læseren så meget mere opmærksomhed end ventet, at det kan have afbrudt den flydende læsning. Dels fordi der kan være et mønster i de ord læseren snubler over, og dels fordi læserens brug af afkodnings- og forståelsesstrategier kan vise sig netop her.

— Vi beregner følgende øjenbevægelsesmål: Læsehastighed
— To klassiske mål defineret af fikseringstider
— To klassiske mål der forudsætter sammenhængende tekstlæsning
— Seks varianter af målet fokusord der detekterer sekvenser af fikseringer, som bryder læserens flydende progression.

Læsehastighed er et udbredt læsemål, og eye-tracking måler læsehastigheden med høj præcision, men målet kræver ikke eye-tracking. Læsehastighed er almindeligt brugt til at evaluere elevers læsefærdighed (Elbro, 2021). Hastigheden måles i ord pr. minut, og vi beregner den ud fra antallet af viste ord, fra første tekstside vises og frem til bogen lukkes.

Det er en veletableret indikator for læsning, og indgår ofte som en delkomponent i standardiserede test af generelle læsefærdigheder. I denne undersøgelse fungerer læsehastigheden under eye-tracking som en pålidelighedslæseindikator.: Hvis målets sammenhæng med læse- og stavemålene forsvinder i vores skolevenlige naturlige skole-setup, så er det et tegn på, at der er sket noget andet end læsning under eye-trackingen, og datasættet vil ikke kunne belyse forskningsspørgsmålene. Vi forventer derfor, at målet læsehastighed korrelerer stærkt med både læse- og stavemålene, hvis det er tilfældet.

Gennemsnitlig varighed af ord-fikseringer er varigheden af hver fiksering i gennemsnit. I forsøg med matchede kontroller er der fundet signifikant forlængede ord-fikseringer for både ordblinde voksne (Franzen et al., 2021) og for begynderlæsere (Benfatto et al., 2016). Gennemsnittet kan forvrænges af meget lange fikseringer, hvilket dog modvirkes, når læsetiden øges.

Gennemsnitlig varighed af første fiksering pr. ord (first fixation duration) er gennemsnitstiden af første fiksering på hvert ord. Målet afspejler særligt afkodningsprocesser og læserens forventning til tekstens sværhedsgrad (Holmqvist et al. 2011). Målet har blandt andet været anvendt til at undersøge hvor tidligt ord yderst i det centrale synsfelt kan genkendes – en evne der først udvikles efter tilegnelsen af basale læsefærdigheder (Sperlich et al., 2015). Når læsere udvikler bedre perifer afkodning, giver det højere læseeffektivitet, som bl.a. giver kortere første fiksering pr. ord.

Begge ovenstående mål er klassiske øjenbevægelsesmål, der ændrer sig med generelle kognitive færdigheder, specifikke læsefærdigheder og aspekter af teksten. Det gælder både under læsning af enkeltord og tekstlæsning (Rayner, 1998, Holmqvist et al. 2011). Vi forventer på den baggrund en signifikant sammenhæng mellem disse mål og med mål som afspejler afkodningsfærdighed. Det er en praktisk svaghed ved begge mål, at bidraget fra fx læsefærdigheder eller ordenes sværhedsgrad er et svagt signal som let drukner i støj fra andre variationer i hhv. læserne og ordene, hvis de ikke kontrolleres på gruppeniveau i et studie. Dette gør målene svære at bruge uden for laboratoriet til evaluering og undervisning af elever.

First pass dwell time og *andel overskimmede ord* (ratio of words skipped during first pass reading) er begge mål, der kan bruges til at måle læsning af sammenhængende tekster. De er følsomme for ordblindhed hos voksne funktionelle læsere (Franzen et al., 2021). Da ordblinde læseres øjenbevægelser ofte er sammenlignelige med yngre læsematchede kontroller (Benfatto et al., 2016), *forventer vi*, at målene også er følsomme for væsentlige forskelle i læse- og stavefærdighed i vores elevgruppe. First pass dwell time er den samlede tid i sekunder, som læseren brugte på at læse fremad i tekstens læseretning. Målet inkluderer al den tid, læseren har set på det senest læste ord, inden et nyt ord blev fikseret, men ikke nogen genlæsning. Andelen af overskimmede ord er andelen af tekstens ord, som slet ikke blev fikseret i løbet af first pass dwell time. Begge mål kan afspejle variation i læsestrategier. Tidsmålet, first pass dwell time, vil ligge højt for en læser, der bruger tid på bogstavafkodning, og falde, i takt med at afkodningen automatiseres, og med at læserens strategi baserer sig mere på forudsigelse af teksten, fx med perifer ordgenkendelse. Andel overskimmede ord vil omvendt stige, hvis læseren er god til at forudsige teksten og udnytte mere perifer information. Målet ligger lavere for voksne ordblinde end for en kontrolgruppe (Franzen et al., 2021). Begge mål har den svaghed, at de ikke kan skelne almindelige sikre læsere fra usikre læsere, der bruger en strategi, som ofte søger forud i teksten efter ekstra information.

Fokusord er ord, der blev besøgt i en række af fikseringer med en sammenlagt fikseringstid på mindst et sekund. Der kan være fikseringer uden for ordet undervejs, men hver afbrydelse kan højst vare et halvt sekund. Hvis et ord bliver fikseret flere gange og med længere afbrydelser, registreres det som fokusord, hver gang en sekvens af fikseringer opfylder kriteriet for fokusord. Ved at tillade korte afbrydelser mellem fikseringer på det samme ord, bliver målet robust overfor forskellige afkodningsstrategier, som fx. informationssøgning udenfor ordet, hovedbevægelser og grimasser med øjnene samt måleusikkerhed, som fx ved blink. Et helt sekund brugt på et enkelt ord indikerer, at læseren har brugt ekstra opmærksomhed på ordet, og denne opbremsning i øjenbevægelserne kan være et tegn på, at den flydende læsning er blevet afbrudt, fx af en afkodnings- eller forståelsesudfordring. Et målt fokusord svarer dermed til en sammenhængende sekvens i oplæsningen, der er meningsfuld at afspille og lytte til for nærmere evaluering.

Vi skelner mellem andelen af *læsetiden*, der bidrager til fokusord og andelen af *læste ord,* der blev fokusord. Andelen af læsetid vægter varigheden af afbrydelser højt, mens andelen af ord, der blev fokusord, vægter hyppigheden af afbrydelser højt.

I udviklingsdata fra tidligere optagede læsninger under lignende forhold har vi observeret, at fokusord har en overrepræsentation af ord nær sætningsgrænser. Her kan opbremsningen også skyldes wrap-up-effekten, der er forbundet med forståelsesorienterede stadier i form af syntaktisk og semantisk processering af hele sætningen (Meister, et al., 2022). Her er det derfor mindre sandsynligt at opbremsningen skyldes afkodning og forståelse af det konkrete fokuserede ord. For at isolere de opbremsninger der primært skyldes ordprocessering fra de opbremsninger der også kunne skyldes sætningsprocessering, inddeler vi de identificerede fokusord i hhv. *sætningsmediale fokusord* og *fokusord ved sætningsgrænser*. I alt sammenligner vi seks varianter af fokusordsmålet, for at afgøre empirisk hvilken der bedst afspejler variation i elevernes læse og stave-færdigheder.

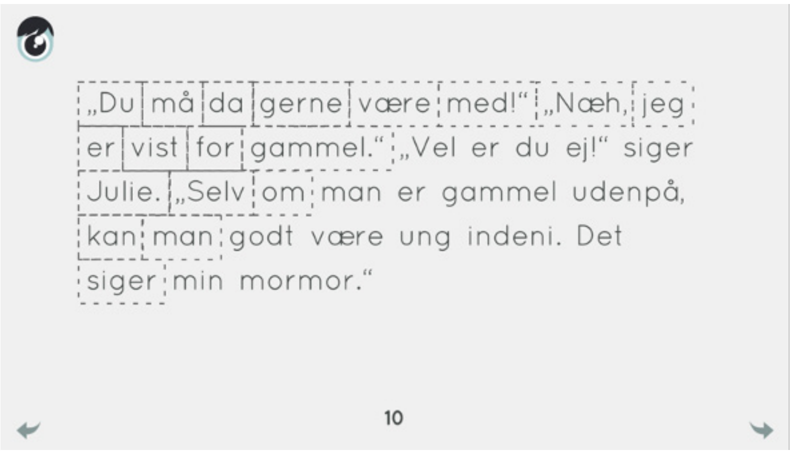
Da der er risiko for, at læsninger med nul fokusord dækker over særligt støjfyldt data, ekskluderes de fra at indgå i denne analyse (se bilag 1).

Optagelse af øjenbevægelser

Elevernes øjenbevægelser blev optaget med en Tobii 4C tracker ved en samplingfrekvens på 90Hz på skolernes egne bærbare computere. Teksten blev præsenteret i fonten Quicksand med op til 7 linjer ad gangen, og den blev vist i et centralt område på maksimalt 80% af skærmens bredde og 60% af højden. Den store margin undgår tracking i yderpositioner hvor præcisionen er lavest. Når der er billeder i teksten, deles det centrale areal mellem billedet til højre og teksten til venstre. I hjørnerne af skærmen vises nederst knapper til at bladre i bogen. Knapperne aktiveres med museklik, piletaster eller uafbrudt øjenfiksering i 2 sekunder. I skærmens øverste hjørner findes til venstre en menu-knap i form af et logo. En timer, der viser rest-læsetiden, vises, når læseren kigger eller peger i øverste højre hjørne.

Fikseringer regnes for at lande på et ord, hvis de lander i et rektangulært areal der dækker ordet plus en ikke-overlappende margin på hhv. 1/2 linjeafstand vertikalt og 1/2 ordmellemrum horisontalt (se Figur 2). Det betyder at alle fikseringer i det centrale tekstområde altid knyttes til et af de viste ord.

Figur 2.
En side i demo-læsningen.



De stiplede kasser rundt om ordene viser, hvordan de optagede fikseringer fordeles mellem tekstens ord. Stiplingerne vises ikke under læsningen.

Analysemetoder
I korrelationsanalyserne bruger vi Spearmans rangkorrelation til at sammenligne målene af læseforståelse og stavefærdighed med klassiske og nye eye-tracking-mål. Det tillader os at detektere, om der er monoton sammenhæng mellem to mål, også hvis et mål nærmer sig et plateau. Det er fx tilfældet, hvis der er et stort fald i læsehastighed mellem elever med hhv. to og fire stavefejl, men kun et lille fald i læsehastighed fra 10 til 12 stavefejl. Vi undersøger altså, om en ændring i et øjenbevægelsesmål afspejler en ændring i et læse- og stavemål, uanset om ændringer i fx. læsehastighed kan oversættes til en proportional ændring i stavefejl.

En tommelfingerregel for styrken af sammenhænge er, at Pearsons korrelationer over 0,5 og under -0,5 er moderate, over 0,7 og under -0,7 er stærke, samt over 0,9 og under -0,9 er meget stærke (Hinkle, Wiersma og Jurs, 2003). To variable med en Pearsons korrelationskoefficient på over 0,7 eller under -0,7 deler ca. 50 % af variansen med hinanden. Når to mål deler meget varians, er det et tegn på, at de i høj grad måler noget ens. Derfor kan stærke sammenhænge tolkes som et tegn på samstemmende validitet mellem to mål.

For Spearmans rho findes der ikke samme retningslinjer for sammenhængen mellem koefficientens styrke og specifikke værdier af koefficienten. Værdien har dog tendens til at være meget sammenlignelig med, og lidt lavere end Pearsons *r* (Laerd Statistics, 2018). Som for

Pearsons korrelationskoefficient veksler værdien af Spearmans rho mellem -1 og 1. Jo tættere koefficienten er på 1 eller -1, desto stærkere er sammenhængen mellem rangordningen i de to mål, og jo tættere på 0, desto svagere er denne sammenhæng (Laerd Statistics, 2018). Så selvom der ikke findes faste regler for, hvor høj Spearmans rho skal være, for at en sammenhæng er meget stærk, så minder værdierne om Pearsons korrelationskoefficient. Derfor vælger vi at trække grænsen for stærke sammenhænge ved Spearmans rho værdier over 0,6.

Når vi laver signifikanstest af mange korrelationer, øges risikoen for at finde tilfældigt signifikante forskelle. Dette tager vi højde for ved dels at Bonferroni-justere signifikansniveauet og dels rapportere 95%-konfidensintervaller for korrelationerne. Intervallets bredde giver et indtryk af hvor præcist den observerede korrelation kan forventes at afspejle populationen og grænseværdiens afstand fra 0 giver et indtryk af korrelationens pålidelighed. Hvis 0 er indenfor intervallet, er korrelationen ikke pålidelig.

I hierarkiske regressionsanalyser undersøger vi om det nye eye-tracking-mål, som er designet til at kunne anvendes i praksis med et evaluerings- og didaktisk sigte, forklarer andet og mere af målene af læseforståelse og stavning end de klassiske eye-tracking-mål, hvilket er et udtryk for, at det nye og de klassiske eye-tracking-mål ikke er identiske mål. I de hierarkiske regressionsanalyser kan vi introducere klassiske mål i analysen før det nye eye-tracking-mål. Hvis det nye mål, på dette trin, bidrager signifikant til at forklare variation i målene af læseforståelse og stavning, så understøtter det, at det nye mål afspejler variation i elevernes læse- og stavefærdighed, som ikke afspejles af klassiske eye-tracking-mål. Det understøtter, at selvom det nye mål er stærkt forbundet til klassiske mål, er de ikke identiske, og at det nye mål samtidig bidrager til at forklare mere variation i målene af læseforståelse[3] og stavning end det klassiske eye-tracking-mål.

Vi anvender det ene af de to overordnede typer af fokusmål. Vi vælger den variant af fokusordsmål, der korrelerer numerisk stærkest med læse- og stavemålene i korrelationsanalyserne. Da vores deltagergruppe er begrænset, kan vi ikke have alle de klassiske eye-tracking-mål med i regressionsanalyserne uden at det går ud over analysens mulighed for at vise signifikant bidrag fra de enkelte mål, så vi udvælger også her det mål, der korrelerer stærkest med læse- og stavemålene.

3 I de hierarkiske regressionsanalyser anvender vi læseforståelseseffektiviteten, altså antal korrekt besvarede spørgsmål per minut. Det skyldes, at vi ønsker at forklare den samlede læsefærdighed ud fra eye-tracking-målene.

Resultater

I dette afsnit belyses det, hvordan forbindelsen mellem læse- og stavemål henholdsvis eye-tracking-mål kommer til udtryk, når eleverne i dette studie læser sammenhængende tekster med eye-tracker.

Som baggrund for korrelations- og hierarkiske analyser viser Tabel 2 deskriptiv statistik for de inkluderede læse-, stave- og eye-tracking-mål.

Tabel 2.
Deskriptiv statistik for alle mål.

	M	median	SD	Min	Max
Læse- og skrivemål					
Rigtige per minut[a]	3,04	2,90	1,21	0,80	6,50
Samlet besvarelsestid[a]	14,73	13,70	5,93	6,80	43,60
Procent i korrekt besvaret[a]	88,62	91,00	9,92	57,0	100,0
Andel korrekt stavede ord[b]	57,57	61,00	23,03	3,00	100,00
Eye-tracking-mål					
Læsehastighed (ord per min.)	112,32	117,82	30,68	39,67	178,53
Fokusord (procent af læste ord)	5,37	2,98	5,95	0,45	26,23

Mediale fokusord (procent af læste ord)	3,14	1,51	3,95	0,18	17,68
Wrap-up fokusord (procent af læste ord)	2,32	1,50	2,13	0,23	9,29
Fokusord (procent af læsetid)	14,62	10,49	11,42	2,49	51,62
Mediale fokusord (procent af læste ord)	8,04	5,27	7,31	0,78	32,56
Wrap-up fokusord (procent af læsetid)	6,87	5,65	4,64	0,85	21,88
Gns. varighed af ord-fikseringer (sekunder)	0,27	0,27	0,07	0,11	0,53
Gns. varighed af første fiksering pr. ord (sekunder)	0,28	0,28	0,08	0,11	0,59
"First pass dwell time" (sekunder)	67,14	67,98	21,31	14,59	111,00
Overskimmede ord i "first pass" (procent af læste ord)	0,46	0,48	0,11	0,17	0,69

[a] Læseforståelse, Hogrefe,
[b] Stavning, Hogrefe
Vi rapporterer gennemsnittet, medianen, standardafvigelsen og spredningen for hvert mål.

Vi vurderer styrken af sammenhængen mellem læse- og stavemål og de klassiske eye-tracking-mål med Spearmans rangkorrelationer (se Tabel 3). Vi leder efter sammenhænge mellem de to typer af mål, der minder om, dem vi kender fra eksperimentelle setup.

Korrelationerne i Tabel 3 viser, at læse- og stavemålene er moderat til stærkt korreleret med læsehastighed, med signifikante korrelationskoefficienter på mellem 0,420 og 0,655 og konfidensintervaller mellem 0,051 og 0,886. Disse sammenhænge bekræfter, at eleverne under optagelsen af øjenbevægelser, som forventet, har brugt tiden på almindelig oplæsning. Der er dermed ikke grund til at tro, at eleverne oplevede situationen som noget andet end normal læsning, og vi kan tolke resultaterne i dette lys. Det er værd at bemærke at styrken af korrelationen mellem læsehastighed og læse- og stavemålene minder om den, vi ser mellem de forskellige varianter af eye-tracking-mål og læse- og stavemålene. Der er således ingen eye-tracking-mål, der er stærkere forbundet til læse- og stavemålene end det simple mål af læsehastighed.

Af Tabel 3 fremgår det, at sammenhængene mellem læseforståelses- og stavemålene og den gennemsnitlige varighed af ord-fikseringer, samt den gennemsnitlige varighed af første fiksering per ord, er moderate, med signifikante negative korrelationskoefficienter mellem -0,469 og -0,565. Sammenhængen mellem læse- og stavemålene og de to tekstbaserede mål, first pass dwell time og overskimmede ord i first pass, er svage, og alle korrelationer er for disse to eye-tracking-mål ikke-signifikante. I eksperimentelle studier har alle fire klassiske mål vist sig at have en sammenhæng med læsefærdigheder. Det er dog ikke det mønster, som vi finder, når vi måler øjenbevægelser i en skole-hverdag. Her er det kun to af fire eye-tracking-mål (gennemsnitlig varighed af ord-fikseringer og gennemsnitlig varighed af første fiksering per ord), der er robuste for den ekstra målestøj, der er forbundet med indsamling af øjenbevælgelsesdata i et skolevenligt-setup.

Tabel 3.
Spearmans rho (rs) korrelationskoefficienter for læse-, stave- og eye-tracking-mål.

	Læsning			Stavning
	Rigtige per minut	Samlet besvarelsestid	Procent korrekt besvaret	Andel korrekt stavede ord
Eye-tracking-mål				
Læsehastighed (ord per min.)	0,615** (0,325; 0,807)	-0,592** (-0,781; -0,298)	0,420** (0,051; 0,737)	0,655** (0,320; 0,886)
Fokusord (procent af læste ord) A	-0,565** (-0,789; -0,261)	0,547** (0,217; 0,759)	-0,385** (-0,717; -0,031)	-0,607** (-0,818; -0,271)
Mediale fokusord (procent af læste ord) B	-0,526** (-0,766; -0,210)	0,493** (0,161; 0,741)	-0,385** (-0,725; -0,017)	-0,631** (-0,864; -0,309)
Wrap-up fokusord (procent af læste ord) B	-0,491** (-0,763; -0,143)	0,498** (0,148; 0,752)	-0,273 (-0,630; 0,113)	-0,487** (-0,735; -0,166)
Fokusord (procent af læsetid) A	-0,632** (-0,812; -0,362)	0,608** (0,305; 0,801)	-0,422** (-0,714; -0,053)	-0,629** (-0,805; -0,281)

Mediale fokusord (procent af læsetid) B	-0,562** (-0,777; -0,255)	0,520** (0,200; 0,756)	-0,440** (-0,743; -0,084)	-0,666** (-0,853; -0,332)
Wrap-up fokusord (procent af læsetid) B	-0,553** (-0,755; -0,211)	0,565** (0,239; 0,758)	-0,271 (-0,594; 0,134)	-0,469** (-0,748; -0,085)
Gns. varighed af ord-fikseringer	-0,469** (-0,770; -0,066)	0,465** (0,078; 0,754)	-0,319 (-0,703; 0,025)	-0,562** (-0,826; -0,267)
Gns. varighed af første fiksering pr. ord	-0,480** (-0,807; -0,102)	0,481** (0,121; 0,793)	-0,318 (-0,722; 0,006)	-0,565** (-0,817; -0,281)
”First pass dwell time”	-0,051 (-0,435; 0,412)	0,019 (-0,457; 0,430)	-0,132 (-0,496; 0,255)	-0,092 (-0,412; 0,283)
Overskimmede ord i ”first pass” (procent af læste ord)	-0,155 (-0,484; 0,256)	0,085 (-0,357; 0,436)	-0,266 (-0,641; 0,164)	-0,216 (-0,543; 0,198)

n=68; A: n=65, B: n=64, ** angiver at korrelationen er signifikant p<,05 (Bonferroni-justeret to-halet). Tal i parentes er grænserne for 95%-konfidensintervallet (bootstrap-samplet med replacement, n=1000).

For at validere, at det nye eye-tracking-mål fokusord både er forbundet til læse- og stavefærdigheder og klassiske eye-tracking-mål, bruger vi Spearmans rangkorrelationer (se Tabel 3 og 4) til at vurdere styrken af sammenhængen mellem:

1. Klassiske eye-tracking-mål og det nye eye-tracking-mål fokusord (Tabel 4)
2. Læse- og stavemålene og det nye eye-tracking-mål fokusord (Tabel 3).

Spearmans rho korrelationerne i Tabel 4 viser, at der er stærke sammenhænge mellem fokusord og både den gennemsnitlige varighed af ord-fikseringer og den gennemsnitlige varighed af første fiksering per ord (r_s mellem 0,688 og 0,816 og konfidensintervaller der ligger mellem 0,369 og 0,936, se Tabel 4). Dette tolkes som et tegn på samstemmende validitet mellem fokusord og de to fikseringsmål. Omvendt er sammenhængen mellem fokusord og målene “First pass dwell time” og overskimmede ord i “first pass” ikke stærke (r_s mellem 0,145 og 0,425 og konfidensintervaller mellem –0,349 og 0,650 hvoraf kun et enkelt netop ikke inkluderer 0, se Tabel 4).

Spearmans rho-korrelationerne i Tabel 3 viser, at styrken af sammenhængen mellem læse- og stavemål og fokusord varierer med hver type af mål.

Sammenhængen mellem andel korrekt stavede ord er generelt stærkt forbundet med det samlede fokusordsmål og med mediale fokusord (r_s mellem -0,666 og -0,599, og konfidensintervaller mellem –0,864 og –0,271, se Tabel 3), og moderat forbundet til wrap-up fokusord (r_s på hhv -0,487 og -0,469, og konfidensintervaller mellem –0,748 og –0,085, se Tabel 3). De stærke til moderate sammenhænge tolkes, som et tegn på at fokusord, ligesom de klassiske eye-tracking-mål, er forbundet med stavefærdighed.

Sammenhængen mellem fokusordsmålene og læsemålene er generelt moderate. Kun for læsemålet procent korrekt besvarede læsesforståelsesspørgsmål finder vi svage og ikke signifikante sammenhænge, nemlig med wrap-up-varianten af fokusordsmålet. Det tolker vi som et udtryk for, at fokusord særligt er forbundet med samlet læsefærdighed, og i mindre grad med læseforståelsespræcision. Derfor anvender vi det bredere læseforståelseseffektivitetsmål i de hierarkiske regressionsanalyser (se Tabel 6).

I de hierarkiske regressionsanalyser (se Tabel 5 og 6) vil vi gerne anvende det samme fokusordsmål, uanset om det er stavning eller læseforståelseseffektivitet, vi ønsker at forklare variation i. Fokusord opgjort som procent af læsetid opnår med en enkelt undtagelse konsistent numerisk højere signifikante Spearmans rho-korrelationer (se Tabel 3) med læse-/stavemålene end fokusord opgjort som procent af læste ord. Vi vælger mediale fokusord, da det mål har den højeste observerede korrelation med stave- og læsemålene (r_s =-0,666 og konfidensintervallet *-0,853 til -0,332)* og samtidig udviser stærk sammenhæng på tværs af stave- og læsemål.

Tabel 4.
Spearmans rho korrelationskoefficienter for klassiske eye-tracking-mål og fokusord.

	Klassiske eye-tracking-mål			
	Gns. varighed af ord-fikseringer	Gns. varighed af første fiksering pr. ord	"First pass dwell time"	Overskimmede ord i "first pass" (procent af læste ord)
Fokusord				
Fokusord (procent af læste ord) A	0,811** (0,577; 0,935)	0,816** (0,587; 0,936)	0,264 (-0,206; 0,633)	0,380 (-0,050; 0,706)
Mediale fokusord (procent af læste ord) B	0,770** (0,502; 0,907)	0,784** (0,523; 0,922)	0,220 (-0,261; 0,560)	0,344 (-0,073; 0,628)
Wrap-up fokusord (procent af læste ord) B	0,816** (0,543; 0,931)	0,814** (0,558; 0,930)	0,331 (-0,105; 0,650)	0,425** (0,008; 0,752)
Fokusord (procent af læsetid) A	0,753** (0,474; 0,908)	0,752** (0,480; 0,915)	0,190 (-0,318; 0,594)	0,312 (-0,155; 0,630)
Mediale fokusord (procent af læsetid) B	0,707** (0,413; 0,896)	0,715** (0,402; 0,902)	0,145 (-0,349; 0,541)	0,278 (-0,147; 0,601)
Wrap-up fokusord (procent af læsetid) B	0,701** (0,392; 0,885)	0,688** (0,369; 0,883)	0,194 (-0,259; 0,573)	0,299 (-0.184; 0,631)

n=68, A. n=65, B. n=64, ** angiver at korrelationen er signifikant p<,05 (Bonferroni-justeret to-halet). Tal i parentes er grænserne for 95%-konfidensintervallet (bootstrap-samplet med replacement, n=1000).

Vi laver to multiple hierarkiske regressionsanalyser i to trin for at forklare variation i henholdsvis læseforståelseseffektivitet og andel korrekt stavede ord. I første trin inddrager vi det klassiske eye-tracking-mål gennemsnitlig varighed af første fiksering pr. ord., da dette mål er det af de fire klassiske eye-tracking-mål, der har den numerisk stærkeste sammenhæng med de to læse- og stavemål (Tabel 3). Det klassiske mål inddrages i trin 1. I trin 2 inddrages målet mediale fokusord (procent af læsetid), som er det af de seks fokusordsmål vi vurderer, er bedst til at forklare variation i stavning og læsning på tværs af målene andel korrekt stavede ord og læseforståelseseffektivitet (Tabel 3).

Tabel 5 viser modellerne med detaljer for andel korrekt stavede ord og Tabel 6 for læseforståelseseffektivitet. Antagelserne for den multiple hierarkiske regressionsanalyse er i vidt omfang overholdt (Se Bilag 2).

Tabel 5.
Multipel hierarkisk regression.

	Andel korrekt stavede ord			
	Model 1		Model 2	
Variabel	B	β	B	β
Konstant	100,553**		70,475**	
Gns. varighed af første fiksering pr. ord	-152,703**	-,491	18,328	,059
Mediale fokusord (procent af læsetid)	-	-	-2,431**	-0,745
R^2	,241		,478	
F	19,335**		29,347**	
ΔR^2	,241**		,254**	
ΔF	19,335**		30,127**	

n=63, **p<,001. Note. Niveauet for signifikans er bonferroni justeret til p=,025 for to outcomevariable i samme datasæt.
Forklaret variation i andel korrekt stavede ord fra gns. varighed af første fiksering pr. ord og mediale fokusord.

I modellen med antal af korrekt stavede ord som outcome-mål gjaldt det, at den fulde model, som bruger gennemsnitlig varighed af første fiksering pr. ord og mediale fokusord til at forudsige andel korrekt stavede ord, var signifikant (tilpasset R^2 =,478, $F(2, 60)$ =29,347, p <,001, se Tabel 5).

Tilføjelsen af mediale fokusord som procent af læsetid til forklaringen af andel korrekt stavede ord (Model 2 i Tabel 5) ledte til en signifikant forøgelse af R^2 på ΔR^2 =,254, F(1, 60) = 30,127, p <,001. For deltagerne i dette studie forklarer det nye fokusordsmål altså en unik andel af

Tabel 6.
Multipel hierarkisk regression .

	Læseforståelseseffektivitet			
	Model 1		Model 2	
Variabel	B	β	B	β
Konstant	4,816**		3,524**	
Gns. varighed af første fiksering pr. ord	-6,301*	-,392	,987	,061
Mediale fokusord (procent af læsetid)	-	-	-,101**	-,609
			,320	
R^2	,154			
F	11,269*		14,855**	
ΔR^2	,154*		,166**	
ΔF	11,269*		14,855**	

variationen i andel korrekt stavede ord.
n=64, *p=,001, **p<,001. Note. Niveauet for signifikans er bonferroni justeret til p=,025 for to outcomevariable i samme dataseæt.
Forklaret variation i læseforståelseseffektivitet fra gns. varighed af første fiksering pr. ord og mediale fokusord.

I modellen med læseforståelseseffektivitet som outcome-mål gjaldt det, at den fulde model, som bruger gennemsnitlig varighed af første fiksering pr. ord og mediale fokusord til at forudsige læseforståelseseffektivitet, var signifikant (tilpasset R^2 =,320, $F(2, 61)$ =14,855, p <,001 se Tabel 6). Tilføjelsen af mediale fokusord som procent af læsetid til at forklare læseforståelseseffektivitet (Model 2 i Tabel 6) ledte til en signifikant forøgelse af R^2 på ΔR^2 =,166, $F(2, 62)$ = 14,855, p <,001.

I de fulde modeller for læseeffektivitet (Tabel 6) og andel korrekt stavede ord (Tabel 5) er det kun den standardiserede beta-koefficient for mediale fokusord, der er signifikant forskellige fra nul, hvilket indikerer, at denne variabel er lineært forbundet til læseforståelseseffektivitet og stavning. Desto mere tid eleverne bruger på fokusord i medial position i sætningen, desto dårligere læseforståelse har de, og desto dårligere staver de. De standardiserede beta-koefficienter viser, at én standardafvigelses længere tid brugt på fokusord i medial position, hænger sammen med en lavere score i andel korrekt stavede ord svarende til -0,745 standardafvigelser og en lavere læseforståelseseffektivitet svarende til -0,609 standardafvigelser. I afrundede tal betyder det at hver yderligere 4-5 sekunder ud af et minuts læsning brugt på mediale fokusord svarer til et forventet fald på næsten to færre korrekt stavede ord og trekvart færre korrekt besvarede læseforståelsesspørgsmål per minut i de standardiserede test.

Vi har eksplorativt lavet endnu et sæt hierarkiske regressionsanalyser, da vi gerne vil belyse om mediale fokusord forklarer variation i henholdsvis læseforståelseseffektivitet og stavepræcision ud over simpel læsehastighed (se bilag 3). Vi kan se af de analyser, at mediale fokusord ikke forklarer unik variation i læseforståelseseffektivitet (tabel 8, bilag 3) ud over et simpel mål af læsehastighed, men at tilføjelsen af mediale fokusord til læsehastighed til at forklare præcision i stavning (Model 2 i Tabel 7) ledte til en signifikant forøgelse af R^2 på ΔR^2 =,093, $F(2, 65)$ = 9,801, p <,01.

Diskussion og konklusion

Vi har undersøgt, hvorvidt eye-tracking-mål i et skolevenligt-setup afspejler elevers læse- og stavefærdigheder, ved at optage 68 elevers øjenbevægelser under læsning og sammenholde dem statistisk med elevernes resultater på standardiserede læse- og stavetest.
Som svar på det første forskningsspørgsmål, så understøtter undersøgelsens resultater, at begge de klassiske eye-tracking-mål, der er

defineret af fikseringstider, er forbundet med læse- og stavemålene, når elever læser sammenhængende tekster med eye-tracker i et skolevenligt setup. Det betyder, at disse klassiske eye-tracking-mål er robuste nok til at blive indsamlet i et skolevenligt setup og stadig afspejle elevernes læse-og stavefærdigheder. Det har begrænset direkte praktisk relevans da disse mål er svære at tolke, men det tillader os at anvende dem som grundlag for valideringen af det nye fokusordsmål i dette studie.

Fra eksperimentelle studier ved vi, at den samlede læseprocess bidrager til fikseringstiden, som derfor vil være påvirket af først afkodnings- og siden læseforståelsesprocesser. I det skolevenlige-setup finder vi også, at målet har sammenhæng med både stave- og læseforståelsesmålene.

Varigheden af første fiksering pr. ord sammenkædes i litteraturens mere eksperimentelle forsøgsdesign tættest med afkodningshastighed end forståelsesprocesser. Afkodning og stavning er, som tidligere beskrevet, tæt forbundne færdigheder. Og selvom stavning stiller endnu højere krav til præcisionen i de lagrede ortografiske repræsentationer, end afkodning gør, er stavemålet undersøgelsens bedste korrelat til stærkeste mål for elevernes afkodningsfærdighed. Vi finder også, som forventet, at første fiksering pr. ord er stærkest korreleret med stavepræcisionsmålet, men både distributionen (Tabel 2) og korrelationerne (Tabel 3) for første fikserings varighed er stort set identiske med målet for den gennemsnitlige fikseringstid. I gennemsnit blev hvert ord fikseret 1,75 gange, altså under to, hvilket betyder, at de fleste ord kun fikseres én gang, og indikerer at de første fikseringer er dominerende i datasættet. Dette afspejler også, at teksten har relativt mange korte, almindelige ord, og at den er let at læse for denne elevgruppe.

Vi ser altså, at elever der staver sikkert, er tilbøjelige til at have kortere varighed af første fiksering pr. ord og kortere gennemsnitlig fikseringstid, mens elever, som staver mindre præcist, er tilbøjelige til at have længere varighed af første fiksering pr. ord og længere gennemsnitlig fikseringstid. Vi tolker dette som et udtryk for, at øjenbevægelserne påvirkes systematisk af den underliggende variation i deltagernes afkodnings- og stavefærdighed.

De klassiske eye-tracking-mål udviser en robust sammenhæng med læse- og stavemål i undersøgelsens skolevenlige-setup. Den mest oplagte forklaring er, at vi i denne undersøgelse, ligesom for standardiserede læse- og stavemål, sammenligner alle elevers læsning af samme tekst. Det er sandsynligt at de klassiske eye-tracking-mål ville have en svagere sammenhæng med læse- og stavemål, hvis eleverne havde læst forskellige tekster. I et design med forskellige tekster er det vores hypotese, at variationen i øjenbevægelser, der skyldes stave- og læsefærdigheder, og som dermed driver sammenhængen med de standardiserede læse- og stavemål, kan blive overskygget af den variation i øjenbevægelser, som skyldes teksternes forskellighed, fx. i ord- og sætningslængde. Denne undersøgelses design med én fælles tekst kan derfor let overvurdere de klassiske eye-tracking-måls sammenhæng med læse- og stavemål i et skolevenligt-setup.

De tekstorienterede eye-tracking-mål, first pass dwell time og andel overskimmede ord var mindst følsomme for forskelle i læse- og stavemålene i vores studie. Dette er et overraskende resultat, da vi forventede, at bedre læsere ville score lavere på det første mål, og højere på det sidste mål, fordi læsere med sikker afkodning og god forståelse bedre kan udnytte skimmestrategier og visuel perifer information. Målenes ringe evne til at afspejle elevernes læse- og stavefærdigheder kan skyldes, at måden teksten blev præsenteret på, kan have forurenet begge mål. Det kan være sket, fordi læserens blik ofte slutter med at fokusere på bladreknappen nederst til højre og derefter søger diagonalt op over den nye side for at læse videre. Da begge mål forudsætter, at den første fiksering på et ord på siden er et forsøg på at læse ordet i kontekst, regnes alle foranstående ord på samme side, der endnu ikke har været fikseret, som oversprungne. Derfor kan korte orienterende blik, uden reel betydning for læsningen, have undergravet disse to måls validitet i et skolevenligt setup. Den forklaring kan testes ved at annotere øjenbevægelsesdata manuelt, så eventuelle orienteringsblik skæres fra på hver læst side, men vi har valgt ikke at gå videre med at undersøge muligheden, da den type manuel annotering ikke er tilgængelig i skolen i praksis. Ekstra blanke sider med fikseringskryds mellem hver bladring kunne eliminere den risiko, men de påvirker læserens arbejdshukommelse og flow uhensigtsmæssigt, og gør samtidig læsning med eye-tracker påfaldende anderledes og mindre attraktivt didaktisk.

Vi finder, som svar på det andet forskningsspørgsmål, at det nye fokusordsmål korrelerer med læse- og stavemål mindst i samme grad som de klassiske eye-tracking-mål, det teoretisk er sammenligneligt med, og som vi derfor har udvalgt i denne undersøgelse. Vi har undersøgt fokusord både målt som andelen af læste ord og andelen af læsetid brugt på fokusordene. Det generelle mønster i korrelationsanalysen er, at andelen af tid brugt på fokusord er stærkest korreleret med undersøgelsens læse- og stavemål. Det peger på, at længere afbrydelser af den flydende læsning kan være et lidt tydeligere tegn på læse- eller stavevanskeligheder end flere afbrydelser.

Vi har isoleret fokusord der står ved en sætningsgrænse, og fokusord der ikke gør, fordi opbremsninger ved sætningsgrænser er almindelige, også for stærke læsere, og ofte lyder upåfaldende under oplæsning. Korrelationsanalysen afspejler dette, og vi finder, at de mediale fokusord er konsistent stærkere associeret med stavemålet og det læseforståelsesmål som vægter præcision frem for hastighed, sammenlignet med wrapup-varianten. Det indikerer, at opbremsninger midt i sætninger er et tydeligere tegn på vanskeligheder med præcision i læsning og stavning end opbremsninger ved sætningsgrænser.

I korrelationsanalyserne i denne undersøgelse er fokusordsmålet stærkest forbundet til læse- og stavefærdighed, når det defineres som andelen af tid brugt på mediale fokusord, men spredningen og overlappene i konfidensintervallerne betyder at en gentagelse af studiet kunne falde ud til fordel for en anden variant af fokusordsmålet.

Som svar på tredje forskningsspørgsmål, viser de hierarkiske regressionsanalyser, at fokusord forklarer unik variation både i elevernes læse- og stavefærdighed ud over den variation, der forklares af det traditionelle eye-tracking-mål, der er stærkest korreleret med læse- og stavemålene. Resultaterne underbygger, at målet fokusord, der er udformet med praksis for øje, på samme måde som klassiske eye-tracking-mål, kan afspejle elevernes læse- og stavefærdigheder, og endda forklarer variation i disse ud over den, som de klassiske mål kan indfange. Det kan skyldes, at fokusord er designet til at detektere fluency-nedbrud, hvilket betyder, at det er følsomt for nedbrud i både afkodning og forståelse. Det er netop en væsentlig del af, hvad standardiserede stave-, afkodnings- og læseforståelsestest måler, hvorimod klassiske eye-tracking-mål især er brugt i læseforskning til at skelne variation i trænede læseres flydende læsning. Vi kan dog ikke ud fra denne undersøgelse forklare hvad der præcis gør, at fokusord bidrager til at forklare unik variation i læse- og stavefærdighed ud over den et klassisk eye-tracking-mål forklarer.

Vi finder samtidig at simpel læsehastighed er bedre til at forklare variation i læseforståelseseffektivitet end fokusord. Det resultat underbygger, at fokusords bidrag til at forklare forskelle i læseforståelseseffektivitet i høj grad knytter sig til læsehastigheden. Den variation fokusord forklarer er indeholdt i den variation læsehastighed forklarer. Vi bliver på den måde ikke bedre til at forklare forskellene i elevernes læseforståelseseffektivitet gennem måling af fokusord end gennem måling af simple læsehastighed. Dette er dog ikke tilfældet for stavning, hvor fokusord forklarer unik variation ud over den som forklares af læsehastigheden. Fokusord bidrager ud over læsehastighed, når vi skal forklare stavning. Det resultat kan tale for, at det særligt er følsomheden for fluency-nedbrud, der knytter sig til afkodning, der vægter i denne undersøgelse af fokusord. I det lys kan det særligt væsentlige bidrag fra fokusord være at udpege ord, som har givet anledning til fluency-nedbrud i afkodningen for den enkelte læser. Hvis senere undersøgelser kan skabe klarhed over i hvor høj grad fokusord for den enkelte elev modsvarer ord eleven, havde svært ved at læse, så følger der muligheder for at anvende fokusord som baggrund for undervisningens tilrettelæggelse.

I denne undersøgelse viser vi, at en enkelt kort læsning i en letlæst bog kan afspejle relevante forskelle i elevernes stave- og læsefærdigheder på fjerde klassetrin på to skoler i to forskellige kommuner. Resultaterne kan dog ikke frit generaliseres til andre elevgrupper. Fremtidige studier kan med en stikprøve, der er repræsentativ for danske skoleelever på flere forskellige alderstrin og læse- og staveniveauer, svare på, om resultaterne kan generaliseres til flere elevgrupper

Perspektiver for praksis

Denne undersøgelse adresserer to nødvendige skridt på vejen mod praktisk og metodisk modenhed for eye-tracking-teknologien som et redskab, der gennem optagelse af elevernes øjenbevægelser ved læsning af selvvalgte tekster, kan generere viden, som lærere og læsevejledere kan anvende i arbejdet med at tilpasse læseundervisningen til den enkelte elev.

Det første skridt er at afdække, om forskelle mellem eleverne i øjenbevægelser optaget under selvstændig hverdagslæsning faktisk afspejler forskelle i elevernes læseforståelses- og stavefærdighed. De moderate til stærke sammenhænge mellem klassiske eye-tracking-mål, fokusordsmålene og målene af henholdsvis læseforståelse og stavning underbygger dette. Resultatet bidrager dermed til dette første skridt på vejen mod teknologiens praktiske og metodiske modenhed. Samtidig er det også væsentligt at fremhæve at eleverne i denne undersøgelse alle læste samme tekster. Vi ved derfor endnu ikke om sammenhængen mellem læseforståelse, stavning og fokusord ville være mindre i andre situationer. Denne undersøgelse er endvidere en her-og-nu undersøgelse. Vi kan derfor ikke belyse, om optagelser af øjenbevægelser reflekterer elevernes læse- og staveudvikling over tid.

Et spørgsmål i relation til dette første skridt kan fx være, om elevernes læse- og staveudvikling kan følges, i en ramme der svarer til almindelig frilæsning, mellem tidspunkterne for de standardiserede test, og dermed hjælpe skolen med løbende at opdage og hjælpe elever, der har brug for særlig støtte. Resultaterne, præsenteret her, er baseret på, at eleverne læser samme tekst i et begrænset tidsrum. Med designet lægger undersøgelsen sig tæt op af den ramme som forudsættes under standardiserede tests, og giver dermed så præcist et sammenligningsgrundlag som muligt mellem øjenbevægelsesmåling og standardiserede mål for læsning og stavning, der er almindeligt tilgængelige i en skolekontekst. En oplagt fortsættelse er at undersøge, hvordan øjenbevægelsesmål kan bruges med elever, der læser forskellige tekster.

Det næste skridt mod teknologiens praktiske modenhed i en skolekontekst, er at udvikle mål, der kan tolkes og anvendes med den eksisterende specialviden, som læsevejlederen allerede besidder. Fokusordsmålet er et bud på sådan et mål. I denne undersøgelse finder vi, at fokusord er et robust mål. Det underbygges dels af de moderate til stærke korrelationer med læseforståelse og stavning, og dels af at fokusord forklarer variation i læse- og stavemål ud over den, som de klassiske mål kan forklare. Undersøgelsen belyser fokusords sammenhæng med læse- og stavemål på gruppeniveau. Videre undersøgelser, der sammenligner analyser af elevers højtlæsning og de udpegede fokusord på individniveau, kan besvare i hvor høj grad fokusord udpeger ord, som den enkelte elev har svært ved at læse. Denne videre analyse er nødvendig for at belyse, hvordan udpegningen af fokusord kan anvendes i lærerens og læsevejlederens arbejde med at tilpasse læseundervisningen til den enkelte elev.

På den baggrund tager denne undersøgelse et skridt på vejen mod et længere proces omkring at validere brugen af eye-tracking, som en måde at få relevant viden om elevernes skriftsproglige udvikling i et naturligt skole-setup, og som en mulighed for at få viden om den enkelte elevs læsninger, som baggrund for at tilrettelægge undervisningen.

Referencer

Alexander-Passe, N. (2008). The sources and manifestations of stress amongst school-aged dyslexics, compared with sibling controls. Dyslexia: An International Journal of Research and Practice, 14(4), 291-313. https://doi-org.ez-ucs.statsbiblioteket.dk/10.1002/dys.351

Andersen, D. (2005). 4 år efter grundskolen – 19-årige om valg og veje i ungdomsuddannelserne. København: AKF-forlaget

Benfatto, M. N., Seimyr, G. Ö., Ygge, J., Pansell, T., Rydberg, A. & Jacobson, C. (2016). Screening for dyslexia using eye tracking during reading. PLoS ONE, 11(12), 1-16. https://doi.org/10.1371/journal.pone.0165508

Bundsgaard, J. & Puck, M. R. (2016). Nationale Test – Danske Lærere Og Skolelederes Brug, Holdninger Og Viden. DPU, Aarhus Universitet.

Børne- og undervisningsministeriet (2019). *Dansk Fælles Mål.* Lokaliseret d. 12. januar 2023 på https://emu.dk/sites/default/files/2020-09/GSK_F%C3%A6llesM%C3%A5l_Dansk.pdf
Børne- og undervisningsministeriet (2022). *Folkeskolens Nationale Færdighedstest.* Lokaliseret d. 12. januar 2023 på https://www.uvm.dk/folkeskolen/test-evaluering-og-skoleudvikling/test--og-evalueringsredskaber/folkeskolens-nationale-faerdighedstest
Caravolas, M., Hulme, C. & Snowling, M. (2001). The Foundations of Spelling Ability: Evidence from a 3 Year Longitudinal Study. *Journal of Memory and Language - J MEM LANG*. 45, 751-774.
Carroll J. M. & Iles J. E. (2006). *An assessment of anxiety levels in dyslexic students in higher education.* British Journal of Educational Psychology, 76, 651–662.
Christensen, V. T. (2019). *PISA 2018 – Danske unge i en international sammenligning.* København: VIVE. Det Nationale Forsknings-og Analysecenter for Velfærd.
Elbro, C. (2021). *Læsevanskeligheder*. Hans Reitzels Forlag. København. ISBN: 9788741263359
Elbro, C. (2014). *Læsning og læseundervisning.* Hans Reitzels Forlag. København. ISBN: 9788762501799
Franzen, L., Stark, Z. & Johnson, A. P. (2021). Individuals with dyslexia use a different visual sampling strategy to read text. *Scientific reports*, *11*(1), 6449.
Gran Ekstrand, A. C., Nilsson Benfatto, M. & Öqvist Seimyr, G. (2021, March). Screening for Reading Difficulties: Comparing Eye Tracking Outcomes to Neuropsychological Assessments. In *Frontiers in Education* (Vol. 6, p. 643232). Frontiers Media SA.
Hinkle, D. E., Wiersma, W. & Jurs, S. G. (2003). *Applied Statistics for the Behavioral Sciences.* Boston, MA: Houghton Mifflin Company.
Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H. & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford.
Juul, H. (2019). *Vejledning til staveprøve 1-3*. Hogrefe Psykologisk Forlage
Kousholt, K. (2015). "Børns Gætterier Ved Nationale Test". *Cepra-striben, Nationale Tests, nr. 18:* 46-57.
Laerd Statistics (2018). Spearman's correlation using SPSS Statistics. *Statistical tutorials and software guides*. https://statistics.laerd.com/
Larsen, M., Jakobsen, V. & Rosdahl, A. (2022). Voksnes basale færdigheder. Udvikling i læse- og regnefærdigheder de seneste 10 år. VIVE. e-ISBN: 978-87-7582-071-9
Meister, C., Pimentel, T., Clark, T., Cotterell, R. & Levy, R. (2022). Analyzing Wrap-Up Effects through an Information-Theoretic Lens. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)* (pp. 20-28).
Møller, L. (2013). *Vejledning til tekstlæseprøve 1-8.* Hogrefe Psykologisk Forlag
Nelson, J. M., Lindstrom, W. & Foels, P. A. (2015). Test Anxiety Among College Students With Specific Reading Disability (Dyslexia): Nonverbal Ability and Working Memory as Predictors. *Journal of Learning Disabilities, 48(4),* 422-432. https://doi.org/10.1177/0022219413507604

Ravn, K. (2015). *"Duer ikke: En femtedel af opgaverne i de nationale test kasseret – Folkeskolen.dk"*. folkeskolen.dk. https://www.folkeskolen.dk/572751/duer-ikke-en-femtedel-af-opgaverne-i-de-nationale-testkasseret

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological bulletin*, *124*(3), 372.

Rosdahl, A., Fridberg, T., Jakobsen, V. & Jørgensen, M. (2013): *"Færdigheder i læsning, regning og problemløsning med IT"*. København: SFI – Det Nationale Forskningscenter for Velfærd

Rosdahl, A., Fridberg, T., Jakobsen, V. & Jørgensen, M. (2013): *"Færdigheder i læsning, regning og problemløsning med IT. Sammenfatning af resultater fra PIAAC"*. København: SFI – Det Nationale Forskningscenter for Velfærd.

Sperlich, A., Schad, D. J. & Laubrock, J. (2015). When preview information starts to matter: Development of the perceptual span in German beginning readers. *Journal of Cognitive Psychology*, 27(5), 511-530.

Swalander, L (2012). Selvbillede, motivation og dysleksi. *In Samuelsson (Ed.) Dysleksi og andre vanskeligheder med skriftsproget* (pp. 178-191). Dansk Psykologisk Forlag.

Von der Malsburg, T. & Vasishth, S. (2011). What is the scanpath signature of syntactic reanalysis?. *Journal of Memory and Language*, *65*(2), 109-127.

Wangebo, K. (2011), Running record en systematisk og detaljeret iagttagelse af læseadfærd. *Viden om Læsning, 9*, 23-30. Nationalt Videncenter for Læsning.

Bilag 1

Kriterier for øjendatakvalitet

I vores forsøg har vi tilstræbt at eleverne kunne gennemføre læsningerne med eye-tracking på egen hånd, og der kan derfor være opstået afvigelser under optagelsen som gør de registrerede øjendata uegnede til brug i denne undersøgelse. Det kan fx. ske hvis eleven er blevet forstyrret, har misforstået opgaven, hvis lyset har ændret sig markant eller trackeren er blevet flyttet undervejs. Vi har valgt kriterier for datakvalitet, der tager højde for en skolehverdag hvor en gennemlytning af hver optagelse ikke er realistisk. De følgende fire kriterier har afgjort hvilke optagelser der skulle inspiceres manuelt og evt. ekskluderes fra analysen:

— *Andel af fikserede ord er under 85%*
  Dette forekommer fx. når læseren bladrer videre uden at læse i bogen og hvor eye-trackingen er blevet upålidelig.
— *Læsehastighed er målt til at ligge uden for intervallet 20-180 ord/ minut*

Høj læsehastighed forekommer under almindelig stillelæsning mens lav hastighed måles hvis teksten er alt for svær, hvis der er foregået noget andet end læsning, eller hvor trackingen er blevet upålidelig.

— *Kalibreringspræcission er målt til lavere end 4 ud af 5 mulige* Dette forekommer når kalibreringen af eye-trackeren fejler. Det kan skyldes at instruktionen ikke er fulgt eller at det billede som eye-trackeren optager af læserens øjne, ikke viser pupillen tydeligt nok.

Kriterierne førte til at 9 (11%) af optagelserne blev inspiceret manuelt. Alle 9 blev udelukket fra analysen.

Optagelser med upålidelige øjendata kan stadig afspilles med lyd og manuelt evalueres af læreren, men det er ikke validt at foretage beregninger baseret på upålidelige optagelser af øjenbevægelser.

Detaljeret beskrivelse af læse- og stavetest i undersøgelsen.

Tekstlæseprøve 5 (Møller, 2013): Tekstens karakteristika er beskrevet med antal ord i teksten (946), ord pr side (59), sider (19), ord med mere end seks bogstaver (97), andelen af lange ord (10,2), sætninger (113), gennemsnitligt antal ord pr sætning (8,3) og lix (19). Der er ikke en tidsgrænse på testen, men eleverne opmuntres til, at løse så mange opgaver de kan på 20 minutter. Elever, der bruger mere tid, får lov at fortsætte. Den samlede besvarelsestid måles. Testen beregner 1) antal korrekt besvarede per minut, 2) hastighed = antal forkert + korrekt besvarede per minut, 3) præcision = korrekt besvarede / (korrekt + forkert besvaret), 4) passerede (korrekt + forkert + oversprungne) og 5) antal oversprungne.

Staveprøve 3 (Juul, 2019): Items belyser om staveren staver lydligt acceptabelt, kan bruge hyppige betingede stavemåder, grammatiske suffikser og ordspecifikke stavemåder, samt mestrer stavning af sammensatte ord. Testen er ikke tidsbestemt. Testen beregner både andel korrekt stavede og lydligt acceptabelt stavede ord. Den beregner også andel korrekt stavede for betingede stavemønstre (herunder sammensatte ord), grammatiske suffikser og ordspecifikke stavemåder og stavning af sammensatte ord.

Statistisk kodning af nul fokusord

I dette valideringsstudie har vi valgt at udelukke 8 læsninger med nul fokusord for at opnå et datasæt med mindst mulig støj til validering. Nul Fokusord er den mest støjfyldte score fordi den ville tælle som den bedste score, og samtidig have stor risiko for at fange eye-tracking med mange udfald, uopmærksom hurtiglæsning uden forståelse og stillelæsning. Uden manuel inspektion, må scoren derfor antages at være tvetydig.

Scoren 0 er også problematisk i rangkorrelationsanalysen da den vil skabe et loft i den ene ende af skalaen, med atypiske ties for de tidsbaserede varianter af fokusord. Ved at udelade den potentielt tvetydige 0-score og dermed regne læsninger med 1 fokusord, besøgt i 1 sekund som optimalt, kan vi pålideligt vurdere forskellen på at beregne fokusordsprocent af tiden og af antallet af ord, samt forskellen på fokusordenes placering i sætningen.

I praksis vil en lærer ved manuel gennemlytning og inspektion af optagelsens kvalitet kunne vurdere om resultatet nul fokusord er pålideligt.

Detaljer om optagelse med eye-tracker

Trackeren blev placeret med magnet under skærmen, og den blev kalibreret i ni punkter inden optagelserne.

Fikseringer er beregnet på gennemsnittet af højre og venstre øje med I-VT-algoritmen (Komogortsev et al., 2010), med en nedre grænse for fikseringstid på 50 ms og en saccade-hastighedsparameter på 0,5°/s.

Bilag 2

Detaljeret beskrivelse af antagelser bag de Hierarkiske regressionsanalyser

(Tabel 5 og 6)

Antagelserne for den multiple hierarkiske regressionsanalyse for forudsigelsen af stavning er i høj grad overholdt, i det der var en rimelig lineær sammenhæng mellem outcome-variablen og hver af de to prædiktorvariable for sig og sammen. Dette vurderede vi på baggrund af visuel inspektion af de delvise regressionsplot og studentized residualer-plottet mod de forudsagte værdier. Analysen havde en Durbin Watson-værdi på 1,85, hvilket ikke indikerede, at de enkelte observationer var forbundet til hinanden, hvorfor residualerne vurderes at være uafhængige. Vi vurderede på baggrund af visuel inspektion af studentized residualer-plottet mod de forudsagte værdier, at der var varianshomogenitet for stavning. Toleranceværdien på 0,46, indikerede ikke multikollinearitet for nogle af variablene, hvorfor den stærke sammenhæng mellem gennemsnitlig varighed af første fiksering pr. ord og mediale fokusord (procent af læsetid) vurderes til at kunne indgå i samme analyse uden at overtræde antagelsen om multikollinearitet. For stavning er der en enkelt ekstrem værdi med studentized slettet residual på -3,66. Analysen af særligt indflydelsesrige værdier viste, at ingen leverage-værdier var større end 0,27, hvilket er acceptabelt, men samtidig pegede på to indflydelsesrige datapunkter med værdier for Cooks afstand på 0,15 og 0,12. En gennemgang af data indikerer ikke at disse datapunkter er fejl og vi beholder dem i analysen. Sidst vurderede vi på baggrund af visuel inspektion af et PP-Plot, at antagelsen om normalfordelte residualer blev overholdt.

For forudsigelsen af læseforståelseseffektivitet er næsten alle antagelser overholdt. Det gælder dog ikke antagelsen om varianshomogenitet, og der er endvidere et indflydelsesrigt datapunkter med værdier for Cooks afstand på 0,36. Vi har fastholdt analysemetoden dette til trods, da vi vurderer det som væsentligt, at de to outcome-variable læsning og stavning belyses med samme analyse.

Bilag 3

Analyserne i tabel 7 og 8 er gennemført som en eksplorative analyser efter designet af undersøgelsen. Formålet med analyserne er at belyse bidraget fra eye-tracking-mål i sammenligning med simpel læsehastighed. På den måde vil vi imødegå, at vi med valideringen af det nye eye-tracking-mål op ad eksisterende eye-tracking-mål kommer til, at fremstille resultaterne uden at læseren har en fornemmelse af, om det nye eye-tracking-mål forklarer forskelle i læse- og stavefærdighed, som rækker ud over det velkendte bidrag fra simpel læsehastighed. Læsehastighed er i denne undersøgelse opgjort som antal læste ord per minut.

Tabel 7.
Multipel hierarkisk regression.

	Andel korrekt stavede ord			
	Model 1		Model 2	
Variabel	B	β	B	β
Konstant	4,443		38,977	
Læsehastighed (ord per minut)	,517**	,539	,275(*)	,286
Mediale fokusord (procent af læsetid)	-	-	-2,586*	-,396
R^2	,290		,383	
F	27,011**		20,208**	
ΔR^2	,290**		,093*	
ΔF	27,011**		9,801*	

n=67, (*)p=,027 *p<,01 **p<,001. Note. Niveauet for signifikans er bonferroni justeret til p=,025 for to outcomevariable i samme datasæt.
Forklaret variation i andel korrekt stavede ord fra læsehastighed og mediale fokus-ord.

Tabel 8.
Multipel hierarkisk regression.

	Læseforståelseseffektivitet			
	Model 1		Model 2	
Variabel	B	β	B	β
Konstant	-,187		,443	
Læsehastighed (ord per minut)	,032**	,684	,028**	,590
Mediale fokusord (procent af læsetid)	-	-	-,045	-
				,146
R^2	,468		,480	
F	58,831**		30,451**	
ΔR^2	,468**		,012	
ΔF	58,831**		1,570	

n=68, *p=,001, **p<,001. Note. Niveauet for signifikans er bonferroni justeret til p=,025 for to outcomevariable i samme datasæt.
Forklaret variation i læseforståelseseffektivitet fra læsehastighed og mediale fokusord.

Abstract

Formålet med artiklen er at undersøge danske gymnasieelevers anvendelse af digitale læremidler. Undersøgelsen tager udgangspunkt i observationer og interviews med klasser fra de fire gymnasiale retninger, HHX, STX, HF og HTX, med henblik på at udvikle designprincipper, der understøtter elevers multimodalitet i digitale læremidler. Danske gymnasieelever udviser, på baggrund af en rig individualiseret medieøkologi, en høj grad af multiplexing og singletasking i anvendelse af digitale læremidler, hvilket medfører, at Illeris didaktiske trekant er udfordret i sin traditionelle forstand. Ud af undersøgelsen træder fire digitale elevtyper frem, som adskiller sig fra traditionelle elevtyper ved bl.a. at anvende digitale læremidler til transformation af de læringsrum, de befinder sig i. Artiklen diskuterer på baggrund af undersøgelserne, hvordan elevadfærden kan styrke fællesskaber i transformerende læringsrum, hvordan elevers multimodale blik kan kultivere Expert-like Learner-kompetencer, og hvordan den nye digitale virkelighed understøtter de traditionelle læringsteoretikere, som den danske gymnasieskole bygger på.

The article investigates Danish high school students' use of digital teaching aids to develop design principles supporting multimodality. Danish high school students exhibit high multiplexing and single-tasking in digital teaching aid use, challenging traditional didactic features. The study identifies four digital student types reshaping learning spaces. It discusses how student behavior can enhance community-building, cultivate Expert-like Learner competencies through multimodal engagement, and aligns with traditional learning theories in Danish upper secondary education.

Danske gymnasieelevers multimodale anvendelse i fysiske og digitale læringsrum

Et observationsstudie af danske gymnasieelever, med henblik på undersøgelse af deres faglige læsning

Michael Juul Nielsen, HF & VUC Fyn, & Lasse Bo Jensen, Praxis

Indledning

Igennem det seneste årti har digitale læremidler, bl.a. under begrebet digitale bøger, gjort sit indtog i flere dele af det litterære område fra skønlitteratur til faglitteratur. Denne udvikling har ligeledes gjort sit indtog i den danske gymnasiale sektor, hvor de enten erstatter eller supplerer analoge fagbøger (Danmarks Evalueringsinstitut, 2017, s. 20). I den gymnasiale sektor er digital læsning ikke noget nyt, men særligt udviklingen og implementeringen af digitale læremidler har en række iboende teknologiske potentialer, der ikke understøttes i eksisterende fagmaterialer. Denne artikel vil som udgangspunkt pege hen mod en række læringsaktiviteter, der kan muliggøres ved brug af digitale læremidler, og er skrevet på baggrund af udarbejdelsen af masterprojektet "Læsning og læring i danske gymnasier – med fokus på understøttelse af forskellige læsestrategier og arbejdsformer gennem teknologier i digitale læremidler" (Jensen & Nielsen, 2023) ved Master i IKT og Læring.

Artiklen er imidlertid også del af en større diskussion, idet digitale og analoge teknologier i undervisningen livligt bliver diskuteret overfor hinanden, særligt i forhold til om den ene teknologi er bedre end den anden i læringsaktiviteter. Dette er understreget af undervisningsminister Mattias Tesfaye, der i januar 2023 italesatte Regeringens opgør med digitale teknologier i den danske folkeskole:

> ” Jeg vil godt opfordre til, at undervisningen som hovedregel er analog. Og digital, når det kan begrundes. Så lav det benspænd for jer selv, og overvej, om det kan begrundes. Hvis ikke det kan begrundes, så sluk skærmen, og se hinanden i ansigtet, lyder det.
> (Nørgaard, 2023)

Baggrund – digital læsning og elevers brug af læremidler

Digital læsning (læsedidaktik/læsestrategier, dannelse)
At læse er ikke bare at sætte sig med en tekst og derefter starte med det første bogstav og være færdig, når læseren når det sidste. Der er forskellige genrer i gymnasiet (Carlsen & Hansen, 2015) og forskellige medier, der hver kræver sin tilgang fra den læsende. Ud fra en affordanceanalyse, der dækker over, hvordan teknologier henholdsvis understøtter, men også kan afvige fra det læringsmål, som teknologien skal understøtte (Bower, 2008), opstår der ved digitale bøger en række affordances, som analoge bøger ikke har. Digitale bøger kan f.eks. transformere læsning til en social oplevelse med interaktioner, kollaborationer og diskussioner. De sociale oplevelser kan rammesættes i en situeret læsedidaktik (Carlsen & Hansen, 2015), hvor læringen i klasserummet transformeres til, at eleven skal tilegne sig strategier for at deltage i forskellige læringsrum (klasserummet, studierummet, træningsrummet, projektrummet og værkstedsrummet). Denne deltagelse kræver kompetencer, der dækker over elevernes evne til at deltage i forskellige former for samarbejdende fællesskaber i og udenfor undervisning samt i og med digitale teknologier. Samarbejdet er en hjørnesten i gymnasieloven, og deltagelseskompetencer forstærker og kultiverer en samarbejdskultur, der også har til sigte at pege videre ud i arbejdslivet (Caviglia & Dalsgaard, 2020). Disse indlejrede interaktioner giver den enkelte elev mulighed for at arbejde med teksten på deres egen unikke måde (Lim & Toh, 2020), dog kræver effektiv digital læsning, at den læsende har erhvervet lineære læsestrategier såsom hukommelsesstrategier, organiseringsstrategier, elaboreringsstrategier og overvågningsstrategier samt de metakognitive strategier planlægning, overvågning og evaluering (Carlsen & Hansen, 2015; Samar & Dehqan, 2013). Derudover peger Carlsen og Hansen (2015) på de strategier, der er unikke for digitale tekster: intertekstuelle læsestrategier,

der handler om, hvordan eleven knytter forbindelser mellem forskellige informationskilder i det digitale tekstunivers for at forstå en tekst, og læremiddelstrategier, som handler om elevens strategi i at bruge læremidler til at hjælpe med deres tekstforståelse. Eleven skal mestre disse strategier, da digital læsning kan vænne læseren til at skimme og browse (overfladelæsning), når der søges efter nøgleord, links eller anden særlig information, der skal bruges for at bevæge sig videre i teksten. Digital læsning er præget af, at læserens opmærksomhed ofte flyttes fra sted til sted, dette har betydning for den læsendes fordybelse (Balling, 2017). Den digitale transformation af læringsrummet kan ifølge Marte Blikstad-Balas betyde, at den øgede digitalisering i skolen gør skolens literacy mere individorienteret og subjektiv, og dermed mindre forankret i traditionelle institutionelle tekstpraksisser. Eleven har et større spillerum til selv at vælge og foreslå tekster, der kan arbejdes med. Fordi elever tidligere har været begrænset af indholdet på et skolebibliotek og de værktøjer, læreren har stillet til rådighed, så betyder internettet og de digitale tekster, at hver elev har fri adgang til faglige tekster, Google og Wikipedia, som de kan bruge i undervisningen (Blikstad-Balas, 2016; Blikstad-Balas & Hvistendahl, 2013). Her sættes fokus på elevens informationskompetence, der bl.a. taler ind i elevernes evne til kritisk at søge faglig viden på nettet, med anvendelse af det udvidede tekstbegreb i en virkelighed med mediekonvergens. Dette dækker bl.a. over kompetencerne til kritisk at foretage søgninger efter viden på f.eks. databaser og søgemaskiner, samt at orientere sig i fagligt materiale mhp. fagligt at bearbejde problemstillinger eller undersøgelser (Caviglia & Dalsgaard, 2020).

Brug af læremidler (kompetencer, medieøkologier, instinkt, erfaring)

Digital læsning betyder samtidig, at materialiteten omkransende læsning ændres, for mens det er let orientere sig i en analog bog om, hvor langt henne i bogen læseren er, kan det være mere vanskeligt i en digital bog, hvor læseren kan opleve haptisk dissonans, der dækker over, at læsning for mange elever er forbundet med en taktil følelse, der ikke kan reproduceres i en digital bog, hvilket vil påvirke læseoplevelsen negativt (Balling, 2017). Dog kan digitale teknologier bidrage i undervisningen, hvis de didaktiseres korrekt. Eleverne kan opnå en dybere læring ved målrettet at arbejde på samme opgave, men ved brug af flere medier samtidig, og dermed blive i stand til at udføre multiplexing, der defineres som "the situation in which we have our attention on the same intentional object using more than one medium" (Tække & Paulsen, 2019, s. 12), og ikke med flere opgaver og flere medier samtidigt, så det bliver klassisk multitasking, der netop vil opleves som forstyrrende for den enkelte elev op påvirke kvaliteten af deres opgaveløsning negativt (Tække & Paulsen, 2019). De opnår dermed at kunne bevæge sig simultant rundt i flere dialogiske rum i undervisningen, dette kunne være det formelle rum skabt af læreren eller et uformelt rum skabt af elever og typisk placeret på et eller flere sociale medier samt et vekslende dialogisk rum, hvor det digitale og analoge kombineres (Bülow, 2020). Samme betragtninger om deltagelse i forskellige rum i og udenfor undervisningen findes i Caviglia og Dalsgaard (2020), hvor der også identificeres tre rum, en gymnasieelev skal mestre at deltage i. (1) Åben delingskultur; som finder sted i klasselokalet, videndelingskompetencer, feedback igennem digitale teknologier, kommentering i f.eks. delte dokumenter og evnen til at indgå i dialog med andre om viden, produkter m.m. 2) Dialog med andethed; som kan understøttes og medieres igennem teknologier, hvormed elever kan indgå i dialog med noget ukendt eller fremmed i digitale teknologier. (3) Deltagelse i online rum; som dækker over elevernes indgåelse og deltagelse i forskellige onlinefora på f.eks. sociale medier eller nyhedsplatforme. Her kan eleverne deltage i f.eks. debatter som led i undervisningen, men så er det væsentligt at facilitere deltagelsen, så den tager udgangspunkt i dialog frem for selviscenesættelse.

Denne brede anvendelse af medier og deltagelsesformer i forbindelse med multiplexing-adfærd er ofte elevinitieret, sådan at de selv finder ud af, hvilke værktøjer de har brug for i forbindelse med deres uddannelse (Caviglia et al., 2018). De væsentligste værktøjer, eleverne finder, er: samarbejdsværktøjer, kommunikationsværktøjer, tekstbehandling, Learning Management Systems, læse- og noteværktøjer, søgeværktøjer og videotjenester (Caviglia et al., 2018).

Transformationen mod det digitale kræver som tidligere skrevet nye strategier fra elever (og lærere), men nye strategier kommer ikke af sig selv. De hænger sammen med, at eleven skal opøve en række kompetencer, hvilket belyses i DiDaK-projektet (Dalsgaard et al., 2020). Her betyder handlingsduelighed, erfaringer med tidligere enslydende arbejdsopgaver samt elevernes digitale instinkt, at de tilegner sig en adfærd, der såvel konstruktivt som til tider ukonstruktivt gør dem i stand til at anvende teknologier som kognitive informationspartnere eller ukritiske brugere at informationsteknologier (Dalsgaard et al., 2020).

Der mangler imidlertid viden om, hvordan læsning i digitale læremidler kan være med til at styrke et undersøgende fællesskab i gymnasieskolen. Vi finder, at der inden for dansk forskning er en forholdsmæssig begrænset belysning af danske gymnasieelevers anvendelse af digitale læremidler, sammenlignet med den forskning, der har fundet sted indenfor almen didaktik og pædagogik på gymnasieområdet.

Undersøgelsens teoretiske grundlag

Til at besvare artiklens undersøgelsesspørgsmål:

Hvordan læsning i digitale læremidler kan være med til at styrke et undersøgende fællesskab i gymnasieskolen.

Der er flere teoretiske positioner, der skal tydeliggøres.

Som teoretisk ramme er der valgt Randy Garrisons *Community of Inquiry Framework* (COI) (Garrison, 2016, s. 53 - 65), hvilken vi finder kan danne udgangspunkt for et virkelighedsrelevant design af digitale læremidler til gymnasieskolen.

COI tager sit udgangspunkt i et undersøgende fællesskab, hvor læring er en undersøgende proces, i tråd med Lipmans og Deweys teorier om læring gennem undersøgelser samt Vygotskys socialkonstruktivistiske teori. Mens Lipman opfattede målet som den kritiske refleksion, der sker i sociale fællesskaber, der er undersøgende sammen, så Dewey undersøgende læring som central for refleksiv tænkning, hvilket han betragtede som en uundværlig praksis i læringssituationer (s. 53). Vygotsky var ligesom Lipman og Dewey optaget af det undersøgende, og i tråd med Lipman var han optaget af, at individets læring gennem en undersøgende tilgang er interdependent med de sociale læringssammenhænge, som denne indgår i (s. 54 - 55). COI´s teoretiske fundament bygger dermed på læringsteoretiske traditioner, som Garrison anvender til at udvikle en model, der bygger på undersøgelsesfællesskaber, hvor teknologier medierer aktiviteterne i fællesskabet. Dette ud fra at "commuities of inquiry make use of technological affordances of a rapidly evolving digital society (...)" (s. 54). Garrisons model bygger dermed på en integration af undersøgelsesfællesskaber i digitale teknologier, så elever og undervisere kan udvikle kompetencer, der ligger udover den forståelse, som de førnævnte teoretikere af forskellige årsager ikke havde mulighed for at indtænke i deres teorier (s. 54-59).

COI-modellen er opbygget ud fra de tre hovedkategorier Social Presence, Teaching Presence og Cognitive Presence, som er dependente og samtidig særegne kategorier, der har forskelligt fokus alt efter aktiviteten, men samlet kan anvendes i forskellige undersøgelsesfællesskaber. Medieringen af aktiviteter ved hjælp af teknologi i fællesskabet eller i det individuelle rum er af Dalsgaard og Ryberg (Dalsgaard & Ryberg, 2022) defineret som digitale læringsrum. Her tages udgangspunkt i digitalisering som kognitiv læringspartner og medieøkologi som fundament for deres definition af det individuelle rum. Deres udgangspunkt er, at digitale teknologier kan styrke den enkelte elevs handlekraft, være kognitiv partner og understøtte den individualitet (s. 66), der også er en væsentlig del af læringsdesignet i den danske gymnasieskole. Ryberg og Dalsgaard forbinder det individuelle læringsrum med læringsaktiviteter indenfor kategorierne "undersøgelse, konstruktion og kommunikation i det individuelle rum" (s. 66).

Til at støtte COI-frameworket suppleres med teorien bag samarbejdsprincippet (collaboration principle), som skrevet hos Kirschner et al. (2014). Princippet anvender kognitive perspektiver på kollaborativ læring i et multimodalt miljø og kommer med tre underliggende principper, der giver en retning for, hvornår og under hvilke betingelser kollaborativ læring vil have en positiv effekt på læring (effektivt forstået som hvornår eleverne får det største læringsudbytte af at samarbejde).

Principper, der sikrer at kollaboration i multimodal-læring er effektivt:

1. Læringsopgaven kræver tilstrækkelig kognitiv belastning til at kræve kollaboration og dermed effektiv brug af en kollektiv arbejdshukommelse (s. 548)
2. De kognitive processer og informationer, der er nødvendige for læring, deles efficient og effektivt blandt gruppemedlemmerne (s. 553)
3. Det multimedia-miljø (medieøkologi), der er til rådighed, giver de nødvendige værktøjer til efficient og effektiv kommunikation om indholdet af opgaven samt koordination og regulering af de involverede processer for at minimere transaktionsaktiviteter (s. 561).

Målet med multimodalitet i kollaborativ læring er at optimere forholdet mellem de transaktionsomkostninger, der vil være, når flere skal koordinere en opgave, og de distributionsfordele, der samtidig kan opnås, ved at flere kan samarbejde. Der er ikke nogen garantier for, at samarbejde nødvendigvis giver et bedre udbytte end individuelt arbejde (Kirschner et al., 2014):

> ” However, research shows that simply placing learners in a group and assigning them a task does not guarantee that they will work together, engage in effective collaborative learning processes and/or show positive learning outcomes. At times, it has even proved to be detrimental to learning.
> (Kirschner et al., 2014, s. 548)

Foruden at styrke de kollaborative arbejdsgange, står det også centralt at frembringe et design, der reducerer uvedkommende støj. Til det arbejdes der med teksten "Principles for Reducing Extraneous Processing in Multimedia Learning: Coherence, Signaling, Redundancy, Spatial Contiguity, and Temporal Contiguity Principles" af Mayer og Fiorella (2014), der består af en gruppe principper, der alle har til formål at reducere uvedkommende støj (extraneous processing) i multimodal læring. Principperne har mest af alt karakter af gode råd, der dog alle er funderet i kognitive teorier og afprøvet empirisk. Teksten sætter fokus på, at undervisningsdesignere bør være opmærksomme på begrænsningerne i arbejdshukommelsen ved at være forsigtige med mængden og layoutet af information, der præsenteres for eleverne. Designere bør tilstræbe at minimere mængden af unødvendige detaljer i det grafiske og tekstlige materiale i multimodale beskeder. Uvedkommende materiale bør fjernes, når det er muligt, så kernen i det essentielle materiale er tydeligt for eleven. Alternativt bør der inkluderes ledetråde for at styre eleven kognitivt mod det essentielle materiale (s. 279). Dermed kan extraneous kognitiv belastning defineres som den belastning, elever oplever grundet indhold, der ikke direkte vedrører målet for læringen.

Undersøgelsens metodiske tilgang

Det empiriske materiale, der ligger til grund for denne artikel, består af et litterature review, observationer på tre gymnasier (STX, HTX og HHX), hvor en klasse hvert sted blev fulgt en skoledag, udarbejdelse af levede erfaringsbeskrivelser fra elever samt workshops med to elevgrupper á fire elever i forbindelse med validering af designprincipper og en workshop med fire elever til test af udviklede prototyper.

Dette gøres for at kunne svare på, hvordan danske gymnasieelever i 2023 anvender digitale læremidler som udgangspunkt for deres faglige læsning såvel som del af undervisningen og i forbindelse med lektielæsning og opgaveløsning udenfor for de fysiske rammer, som skolen danner. Metoden, vi anvendte, var *feltstudiet* (Sharp et al., 2019), idet vi med udgangspunkt i nedenstående observationsguide indarbejdede en række fokusområder, der havde til hensigt at fastholde vores dataindsamling ud fra en *passiv observatør*, idet vi kun observerede handlinger (Sharp et. al., 2019, s. 290 - 291). Dette kombinerede vi med den *etnografiske tilgang* (Sharp et. al., 2019, s. 291 - 294), sådan at vi igennem en indledende socialisering med eleverne og deres undervisere fik kendskab til elevgruppen, der blev anvendt til at kvalificere observationerne vedrørende adfærd med og anvendelse af digitale læremidler, der blev anvendt. Denne guide blev yderligere konkretiseret i observerbare fænomener, der kunne iagttages i klasserummet og kan ses i Bilag 1.

Figur 1.
Indledende observationsguide til brug for dataindsamling.

Guide til adfærd og anvendelse (s. 289)	Guide til etnografisk observation (s. 294)
Overordnet observationsspørgsmål: Hvordan anvender eleverne digitale læremidler i undervisningen? **Undersøgelsesspørgsmål:** Hvilke teksthændelser og tekstpraksisser kan observeres i undervisningen? Hvilke læsestrategier (kognitive, sociokulturelle) anvendes af eleverne? Hvordan fremstår elevernes brug af tekster, både i og udenfor skolens literacy? Hvilken genre læremiddel arbejdes der med i den observerede undervisning? Hvor meget tid fylder anvendelsen af digitale læremidler i den samlede observerede stilladserede undervisning?	Klarlægge overfor eleverne, hvorfor og hvordan vi er til stede Klarlægge, hvordan de adfærdsmæssigt skal forholde sig til os som observatører Dokumentere såvel formel som uformel samtale mellem eleverne i anvendelse af digitale læremidler Dokumentere og beskrive det fysiske læringsmiljø Beskrive teknologier, der anvendes i kombination med digitale læremidler

For at udvide vores dataindsamling til også at give os indblik i elevernes arbejde med digitale læremidler uden for skolens fysiske rammer i forbindelse med lektielæsning, bad vi ligeledes eleverne om at udarbejde *levede erfaringsbeskrivelser* som beskrevet hos Van Manen (1990). Her var der fokus på indsamling af skriftlige beskrivelser, mens der hos eleverne er fokus på refleksion over anvendelsen af digitale læremidler i forbindelse med lektielæsning, samt hvornår og i hvilket miljø de læser i digitale læremidler. Hertil blev der udarbejdet en specifik guide til eleverne, som de kunne støtte sig til, ud fra det forhold, at *levede erfaringsbeskrivelser* er en kompleks individuel skriveøvelse, som vi ville sikre os, at eleverne kunne gennemføre mhp. på såvel egen refleksion over egen praksis og kvalificerede beskrivelser til brug for vores dataindsamling.

Vores indsamlede kvalitative data blev analyseret ved brug af en tematisk analyse.

> ” Thematic analysis is a method for identifying, analysing and reporting patterns (themes) within data. It minimally organizes and describes your data set in (rich) detail. However, frequently it goes further than this, and interprets various aspects of the research topic.
> (Braun & Clarke, 2006, s. 79)

Denne havde til formål at identificere mønstre i vores observationer og de udarbejdede levede erfaringsbeskrivelser, der skulle danne grundlag for den videre analyse af elevernes anvendelse af digitale læremidler. Ydermere skulle analysen bruges i forbindelse med udvikling af *designprincipper* fra metodologien i Design Based Research (DBR) til anvendelse af udarbejdelse af prototyper, der kunne hjælpe med at styrke det undersøgende og producerende fællesskab hos danske gymnasieelever ved faglig læsning i digitale læremidler. Designprincipperne har gennemgået en transformation i løbet af projektet, hvor vi startede med de fire nedenfor efter gennemgangen af relevant tilgængelig litteratur.

— Læremidlet skal fokusere på problembaseret læring i en rig ressourceøkologi
— Læremidlet skal invitere til teksthændelser, hvor elever læser og løser opgaver sammen
— Der skal holdes et skarpt fokus på opgaverelateret kognitiv belastning, og urelateret støj skal undgås
— Der skal skabes rum for refleksion over anvendelse af ressourceøkologi.

Efterfølgende er de blevet beriget med data fra observationer fra undervisning, yderligere teoretisk forankring samt to workshops med to grupper elever, der blev afholdt med udgangspunkt i feedback-capture-grid-metoden (Dam & Siang, 2022), hvorefter vi afslutter med fem nye principper, der behandles senere i artiklen.

Ud fra den indsamlede data, vil vi præsentere to hovedpunkter, nemlig Digitale elever og Gymnasieskolen er ikke under forandring.

Digitale elever

I det efterfølgende analysearbejde med observationerne udkrystalliserede sig fire temaer, der gik igen på tværs af al empiri.

Tema 1: Digitale arbejdsformer
Observationerne på tværs af de fire gymnasieretninger viste, at eleverne i meget høj grad anvender og mestrer digitale kollaborative arbejdsformer, idet de vha. forskellige teknologier kan bearbejde en pragmatisk konstrueret læringsaktivitet, eksemplificeret vha. nedenstående observation:

> ” En gruppe arbejder direkte videre med deres præsentation i PowerPoint. En fra gruppen har præsentationen på sin computer, mens den anden har i-bogen samt andet google-søgt materiale åben på sin computer, som de snakker ud fra.
> (Observation HHX)

Her udviser eleverne en høj grad af kollaborativ *handlingsduelighed* som beskrevet af (Dalsgaard et. al., 2020, s. 11 - 13), i og med at eleverne er i stand til at handle på baggrund af teknologier i deres læringsaktivitet. Igennem vores observationer indså vi imidlertid, at kooperation i højere grad var foretrukket frem for kollaboration, når det drejede sig om elever, der i fællesskab skulle udarbejde læringsaktiviteter. Det mest tydelige eksempel på dette var en aflyst undervisning på HTX, hvor eleverne fik en instruks fra underviseren om, hvordan de skulle arbejde videre med en opgave. Her inddelte eleverne opgaven i mindre bidder, som de bearbejdede hver for sig og delte med hinanden. Her opstod der ikke fælles videnskabelse, og en enkelt gruppe resignerede hurtigt over for opgaven, som vist af nedenstående dialog mellem to elever:

> ” En gruppe på to elever, der har delt spørgsmålene op mellem sig, begynder at udvise frustration over opgaven og begynder følgende dialog: Elev 1: ”Skal vi ikke bare aflevere nu?” Elev 2: ”Jo, det kan vi godt, der er bare 12 trætte spørgsmål tilbage”. Elev 2: ”Det er mega træls, for det er bare sådan nogle ja/nej spørgsmål”. Aflevering efter 56 ud af 90 minutters modul. Gruppen sad resten af tiden på deres egne devices.
> (observation HTX)

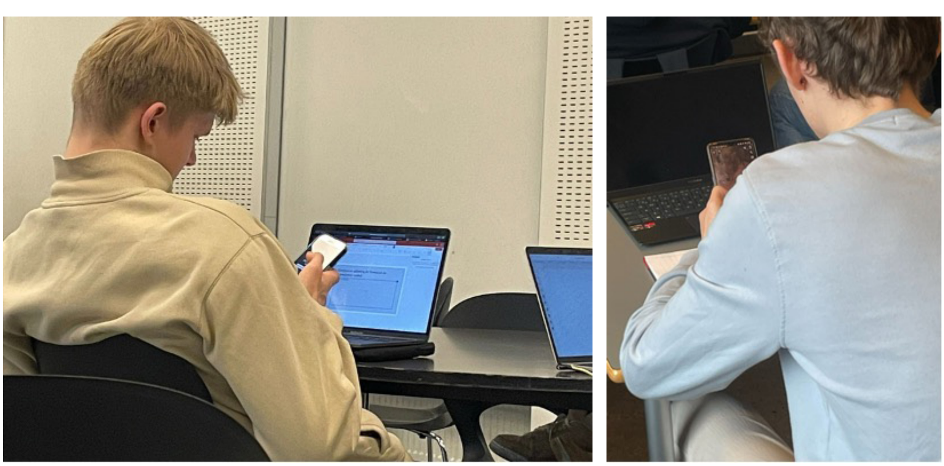
Et andet eksempel var det faciliterede og stilladserede gruppearbejde, som undervisere som led af deres undervisning havde designet deres undervisning ud fra. Særligt gruppearbejdet som læringsaktivitet var meget udpræget i den observerede undervisning.

Tema 2: Mangfoldighed

Mangfoldighed anvendes til at beskrive elevernes brede anvendelse af læremidler i undervisningen på gymnasiet, men også i forbindelse med lektielæsning hjemme. Adgangen til udvalget af læremidler skyldes primært integreringen af internettet i undervisningen, hvilket leder tankerne hen på de tre digitale bølger af Tække & Paulsen (2017), fordi vi med vores observationer kunne se udfoldelsen af bølge 1 og 2, mens den tredje bølge var fraværende. Foruden de tre digitale bølger står begreberne multitasking og multiplexing centralt i vores analyse. Multitasking defineres som “a situation in which we try to follow other intentional objects than the one we are trying to learn about, a situation which does not have a constructive influence on our understanding and learning” (Tække & Paulsen, 2019, s. 12).

Nedenfor blev der observeret en elev, der til sin opgavebesvarelse undersøgte, hvordan Normals hjemmeside så ud på en telefon. Formålet var at finde forskelle mellem visningen på en stor skærm på computeren og den noget mindre telefonskærm – altså et eksempel på multiplexing-adfærd. Ved siden af til højre kunne en anden elev observeres bruge sin telefon som spejl for at få sat sit hår tilfredsstillende, hvilket ikke kan siges at have et fornuftigt læringsudbytte og må karakteriseres som multitasking-adfærd.

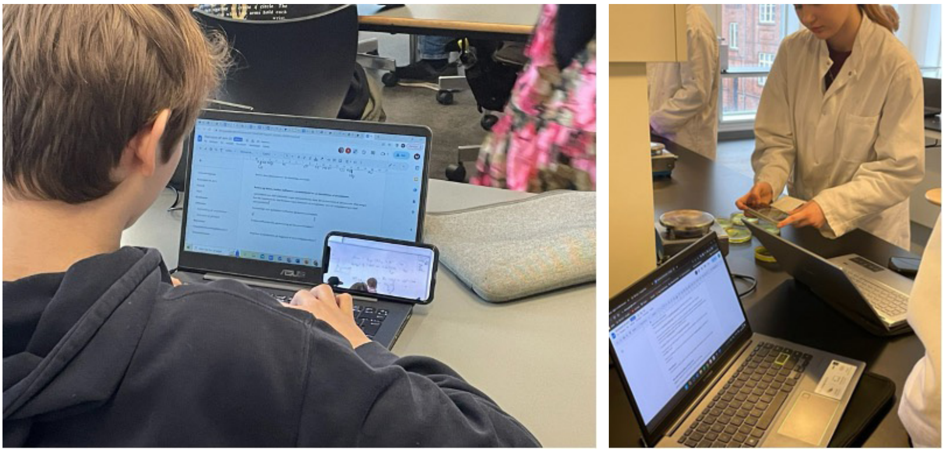
Figur 2.
Elevers mangfoldige anvendelse af digitale enheder.



Generelt bruger eleverne telefoner i deres læring, men formålet varierer, som Tække & Paulsen (2019) bemærkede. Telefonen kan være et nyttigt værktøj i undervisningen, men når den bruges til ikke-relaterede formål, forbedrer den ikke kvaliteten af læringen. Men anvendt didaktiseret af læreren i undervisningen eller med egne stærke strategier for læring kan brugen af telefonen styrke elevernes læring. Nedenfor ses udførelsen af et konkret forsøg i laboratoriet, hvorefter billedet sendes til et andet gruppemedlem, der efterfølgende sætter billedet ind i deres rapport, som i øvrigt udarbejdes i Google Docs, hvor alle arbejder sammen.

> ” Under skrivning af rapport spørger en elev, om den anden ikke vil sætte billederne ind i rapporten (de sidder begge i det delte Google Docs-dokument). Eleverne, der indsætter billederne i rapporten, placerer dem i dokumentet ud fra anvisninger fra den anden elev i gruppen.
> (observation HTX)

Figur 3.
Brug af mobilenhed til at hjælpe med løsningen af opgaver.



På billedet til højre ses, hvordan en elev anvender sin telefon i forbindelse med hukommelsesstrategien (Carlsen & Hansen, 2015, s. 78), hvor den har fungeret som et værktøj til at tage noter i undervisningen. “En elev har placeret sin telefon med et billede af en tavle med noter på sin computer – han anvender flere medier til produktion af sin poster. I samme gruppe bliver der søgt efter billeder på Google til opgaven” (observation HTX).

Vi observerede flere gange, at eleverne i individuel samtale med lærer eller i plenumdiskussioner anvendte læremiddelstrategien, eksemplificeret ved multiplexing.

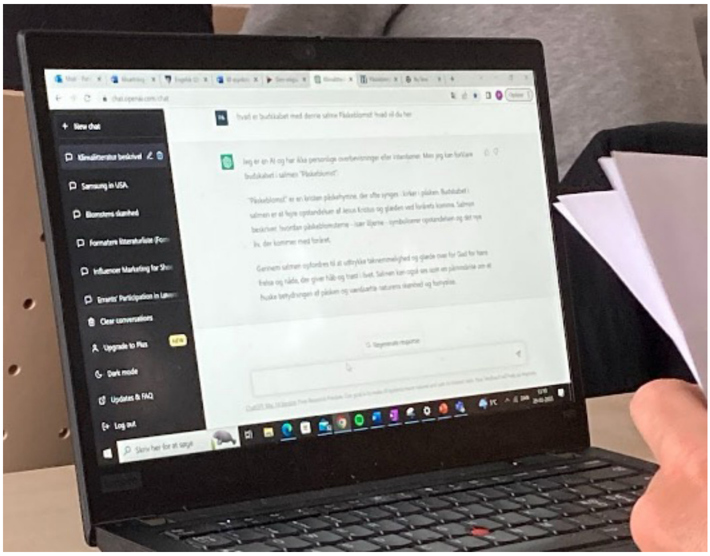
”En gruppe på tre elever snakker med læreren om finanspolitik i forhold til den artikel, de arbejder med. Undervejs i snakken Googler en elev ”ekspansiv finanspolitik” og læser det øverste resultat” (observation HTX). Men hvorfor ikke bare spørge læreren? Vi tolker adfærden sådan, at eleven ikke ønsker at udstille sin manglende viden i det forum, som eleven var en del af, og vælger derfor at finde den viden et “sikkert sted”, hvor det ikke bliver tydeligt for læreren eller de andre elever i gruppen.

Elever kan også vælge andre læremidler grundet glemsomhed eller overbevisning om, at deres egen udviklede strategi er den mest hensigtsmæssige, hvilket Dalsgaard et al (2020, s. 21) kalder *long-time-practitioner* adfærd, der hos eleverne blandt andet kom til udtryk ved følgende observation:

> ” En pige bruger Google til at slå nogle ord op, som hun ikke forstår. Det samme mønster ses hos andre. Det virker til, at de har glemt beskeden fra læreren om, at der er hjælpeord under salmen. To drenge bruger tiden på facebook i stedet for teksten.
> (observation HHX)

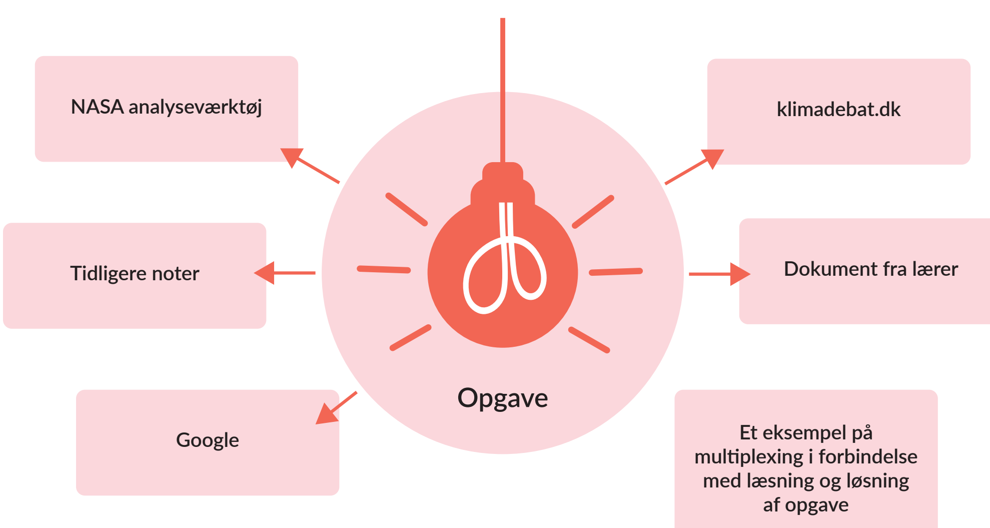
Et værktøj, der blev observeret anvendt bredt i klasserummene, var ChatGPT, ikke kun i forbindelse med lærerinitieret brug, men også for at tilegne sig en læringsressource og deltage i debatten, og for at få løst en opgave uden at skulle bruge egne kognitive ressourcer.

Figur 4.
ChatGPT's tilstedeværelse i klasserummet.



Eleverne havde også beskrevet deres lektielæsning hjemme, og her ser vi den mangfoldighed af ressourcer, en elev anvender til at løse en hjemmeopgave.

Figur 5.
Eksempel på en elevs multiplexing ved lektielæsning



Der er én opgave i spil, og ud fra denne opgave opsøger eleven fem forskellige ressourcer, og det må antages, at eleven hver gang skal bruge lidt tid på at omstille sig til det nye format. Det kan umiddelbart virke uoverskueligt, men eleven angiver ikke nogen form for frustration i sin erfaringsbeskrivelse, hvilket tolkes derhen, at der ikke er oplevet extranenous kognitiv belastning, og at eleven har kunnet bygge ny viden ovenpå sin allerede eksisterende viden om emnet. Vi kan se, at eleven genbesøger sine noter fra tidligere og dermed påbegynder en elaboreringsstrategi og i den grad også anvender en overvågningsstrategi i sin søgen efter andre kilder end dem, læreren har stillet til rådighed, og reflekterer over, om den nye viden er forstået, f.eks.: " Så sammenlignede jeg med en graf for Grønland og fandt forskelle der" (erfaringsbeskrivelse).
Der er observeret en rig og dyb anvendelse af læremidler i undervisning og lektielæsning, og i langt de fleste tilfælde har eleverne anvendt dem meningsfuldt. Hvis vi holder teorien bag kognitiv læring fra Mayer (2014, s. 52) for øje, så vil der pågå et større indre arbejde med at opfatte ord og lyde (sensory memory) for derefter at udvælge nogle af dem til arbejdshukommelsen, hvor de skal organiseres i modellen for afslutningsvist at blive integreret med elevens eksisterende viden i langtidshukommelsen.

Tema 3: Individualitet
Foruden de nævnte arbejdsformer under temaet "Digitale arbejdsformer" kunne der observeres individuelt arbejde, hvilket trådte tydeligst frem i de levede erfaringsbeskrivelser. Lektielæsningen foregår for hovedpartens vedkommende hjemme på værelset og primært i sengen, enkelte angiver også, at de sidder ved et skrivebord (4 ud af 27). Det kan samtidig ses, at langt den meste læsning foregår i forbindelse med løsningen af en konkret opgave, og ikke nødvendigvis læsning for læsningens skyld, hvilket må antages at have betydning for valg af læsestrategi. Det er samtidig et billede på, hvor sammenfiltret (Orlikowski, 2010) deres liv som gymnasieelev er med de digitale enheder. Denne sammenfiltring eksisterer på et rent materielt plan, i en praksis-sammenfiltring og ikke mindst på det kulturelle plan.

I elevernes individuelle lektielæsning, der oftest sker i forbindelse med en konkret opgave (ikke læsning), der skal løses, er det de didaktiserede læremidler (Hansen, 2010, s. 93), der indtager den førende position hos eleverne, men som er under pres fra andre sider, da eleverne i deres læsning samtidig anvender læremiddelstrategien (Carlsen & Hansen, 2015, s. 83) som en digital læsestrategi, der integrerer andre læremidler end de af læreren stillet til rådighed. Her er nogle læremidler mere hensigtsmæssige end andre, f.eks. "TikTok viste mig ChatGPT og lærte mig at bruge den". (erfaringsbeskrivelse)

Flere elever beskriver deres læsesituationer således, at der træder tydelige tegn fra den kognitive læsedidaktik (hukommelsesstrategien) frem (Carlsen & Hansen, 2015, s. 78), nemlig det at gentage læsningen, indtil de oplever at have forstået stoffet, eller at tage grundige noter, oftest i OneNote, hvoraf det kunne ses, at mange af disse noter havde karakter af punktopstillinger. I undervisningen sås det, at eleverne typisk anvender deres skærm todelt, så de kan have teksten i den ene side og OneNote til at tage noter løbende på den anden side af skærmen. Der observeres også scrollen frem og tilbage i teksten, for at eleverne kan genlæse, hvilket også kunne kobles sammen med de kognitive hukommelsesstrategier. Samtidig får lærebogen til det enkelte fag mere karakter af leksikon end en bog, der skal læses fra start til slut, hvilket ses af følgende udsagn: "Jeg kunne bruge afsætningsbogen til at slå ting op" og "Jeg brugte vores digitale afsætningsbog til at finde fagbegreber". (erfaringsbeskrivelser)

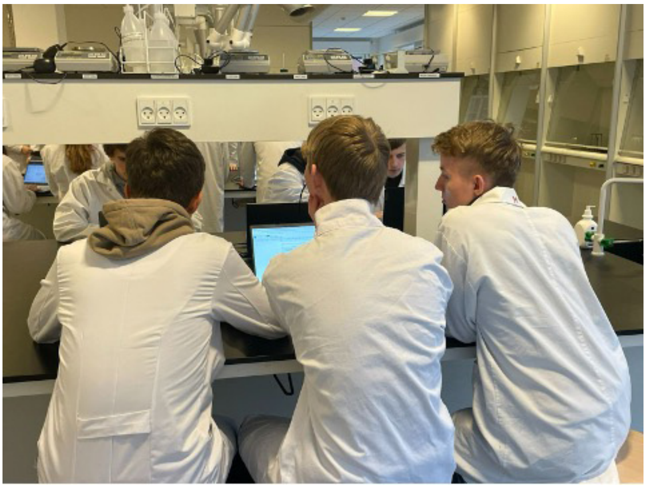
I lektielæsningen (opgaver) bar opgaverne ofte præg af, at eleverne arbejdede med selvvalgt stof, hvilket kom frem i flere udsagn: "Udover det skulle man selv finde information og kilder om emnet. Umiddelbart var den måde at arbejde med en opgave på spændende nok." (erfaringsbeskrivelse). Dette taler ind i pointen fra Blikstad-Balas (2016) om den øgede individualisering og dermed af-institutionalisering af læring, men det kan samtidig give eleverne digitale informationskompetencer som omtalt hos Caviglia et al. (2018) og udgør en væsentlig del af elevers læsning.

I det rammesatte individuelle arbejde på gymnasiet kunne læremiddelstrategien og multiplexing ses anvendt til løsning af opgaver. "En dreng finder modellen (omverdensmodellen) via Google og efterfølgende lærebogen og læser om den for at kunne svare på spørgsmålene i boost." (observation HHX). I observationen her arbejder eleven i læremidlet Boost, der er bygget op, så læremidlet viser den korrekte information, hvis du svarer forkert, men alligevel vælger denne og flere elever at finde svaret via andre læremidler (Google og lærebog), inden de svarer.

Som afslutning på dette tema kommer en betragtning om brug af software. På alle gymnasier kunne det observeres, at “Når man kigger på hvordan de læser teksten, så åbner de teksten via deres browser (safari) og ikke pdf-visningssoftware” (observation STX). Her kunne eleverne have valgt at indlæse deres pdf-dokumenter direkte i f.eks. OneNote og dermed kunne skrive noter direkte i teksten – blot digitalt.

Tema 4: Distraktioner
På alle gymnasier og i hovedparten af erfaringsbeskrivelserne var det meget tydeligt, at eleverne blev distraheret i deres læsning. I én læringssituation kunne vi ikke observere distraktioner hos eleverne, hvilket var første halvdel af et modul på HTX i faget kemi, hvor eleverne var i laboratoriet og udførte et forsøg.

Figur 6.
En gruppe elever samlet om computeren i laboratoriet på HTX



Den manglende distraktion kan ikke tilskrives, at eleverne ikke havde adgang til de teknologier der sædvanligvis distraherede dem – det havde de.

Vi vurderer, at forklaringen i stedet skal findes i disse forhold.

— Fastholdelse i få teknologier, idet læreren havde didaktiseret teknologibrugen i opgaven.
— Praksisrettet undervisning/læremidler, hvor eleven tager rollen på

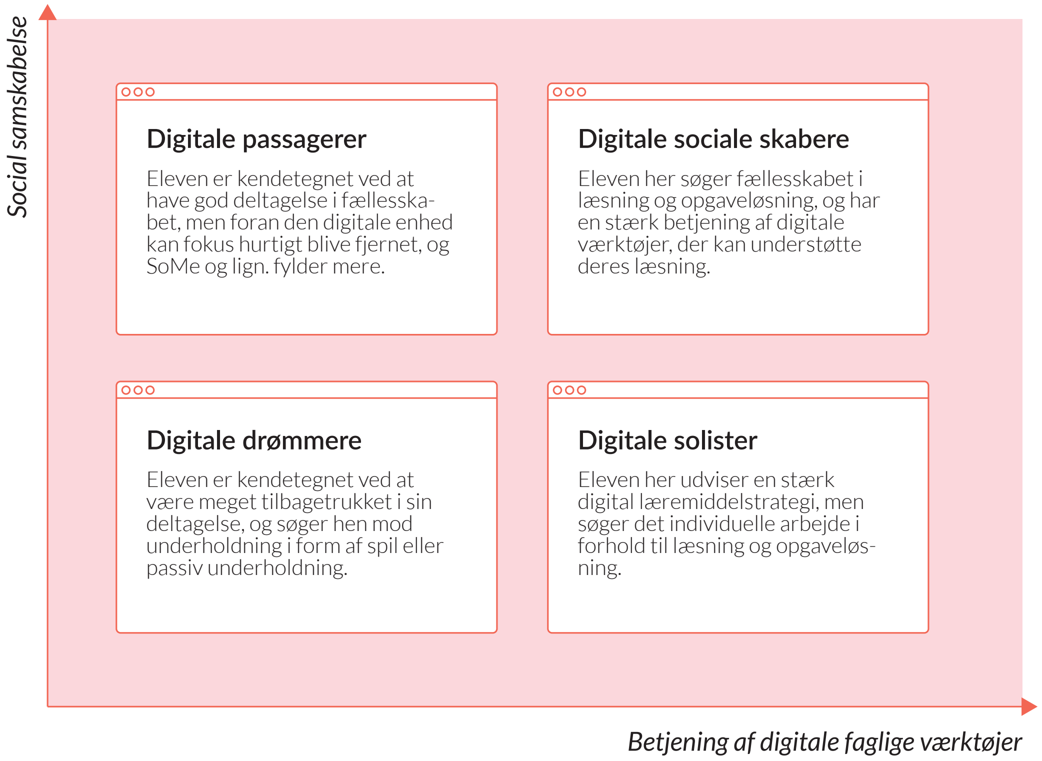
sig som forsker og dermed aflægger sin klassiske elevrolle.
— Projekt- og/eller værkstedsrum, der i sig selv kan tilbyde en anden form for aktivering og fordybelse end det almindelige klasserum.

Derudover oplever flere elever at blive distraheret/frustreret, når de er i tvivl om, hvad det er, læreren gerne vil have de laver, f.eks. "Jeg skulle lave hjemmeopgave i engelsk. Jeg sad derfor før træning i min seng og skrev. Det var frustrerende, fordi jeg synes opgavebeskrivelsen var minimalt forklaret, og opgaven i sig selv var svær." (erfaringsbeskrivelse). Her kan vi læse, at eleven oplever både at være intrinsic og extranenous kognitivt belastet, hvilket ifølge Mayer (2014) vil resultere i en dårligere opgavebesvarelse, end hvis læringsvejen havde været tydelig for eleven.

Hvad fortæller det os om gymnasieelevers læsning?
Vi kunne se, at gymnasieelever læser meget og mange forskellige genrer – men ikke nødvendigvis, hvad Carlsen & Hansen (2015, s. 78) benævner studielæsning, men derimod nærmere læsning for at løse en konkret opgave.

I analysen af empirien af elevernes læsning trådte der en række sammenfald i elevernes adfærd frem, så det var muligt at identificere fire elevtyper (Figur 7).

Figur 7.
De fire digitale elevtyper



Social samskabelse og Betjening af digitale faglige værktøjer er valgt som akser til kategoriseringen af eleverne. Den sociale dimension handler om den enkelte elevs evner til at indgå i læsesituationer, hvor kollaboration indgår kontra en mere individcentreret tilgang. Betjenings-dimensionen skal vise, hvorvidt eleverne formår at anvende de forskellige læremidler med udgangspunkt i multiplexing og ikke multitasking.

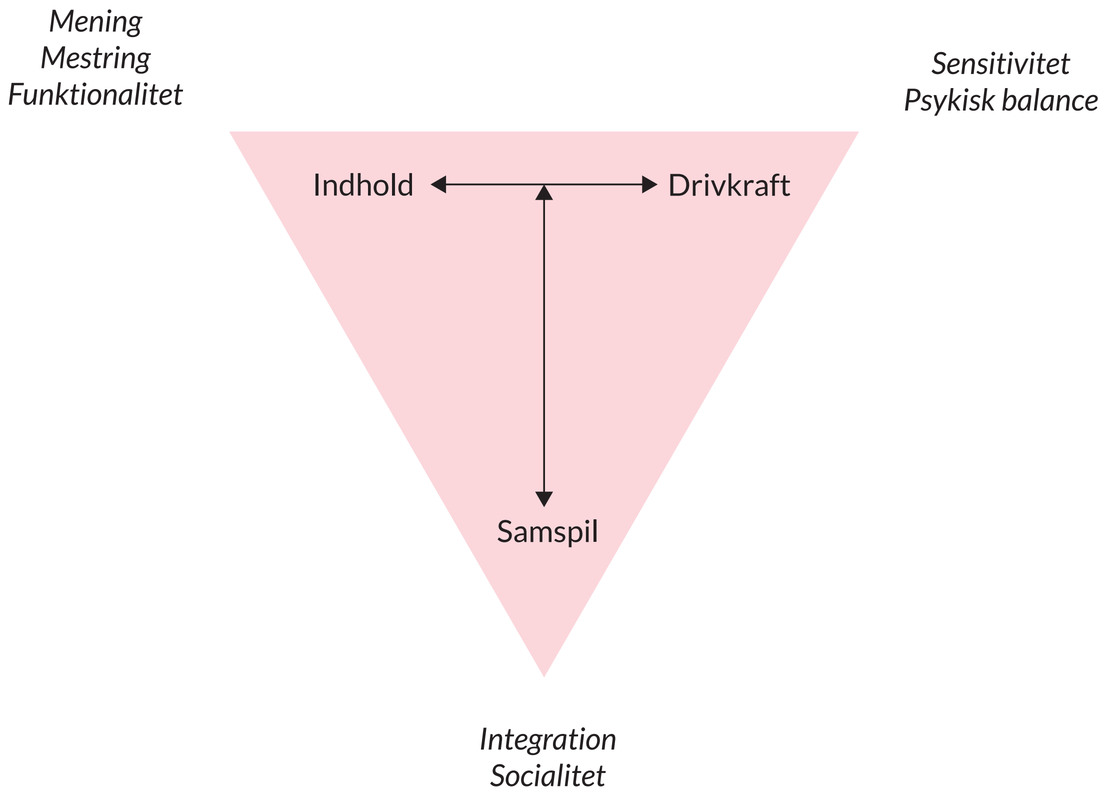
Alle fire typer var til stede i alle observationer, hvor vi i klasserummet identificerede flest digitale passagerer, der er kendetegnet ved, at de har stærke evner til at arbejde i et fællesskab, men mangler evner til at drive læsningen/læringen fremad ved hjælp af digitale værktøjer.

De kan f.eks. sagtens i fællesskab søge hen mod alternativ underholdning. De digitale sociale skabere indgår i arbejdsfællesskaber med de digitale passagerer og kan være den kraft, der kan drive opgaven frem mod løsningen, da de er kendetegnet ved at have en bred og dyb medieøkologi. De digitale solister er stærke til at anvende faglige værktøjer, men søger ikke fællesskabet i løsningen af opgaven, de arbejder hellere selv, men deler gerne deres viden, når der f.eks. skal samles i slutningen af et modul. Vores sidste kategori er de digitale drømmere, der ikke deltager i fællesskabet og ofte lader sig passivt underholde af forskellige videotjenester eller andre sociale medier. De ses også spille, men da vil det ofte være spil uden social interaktion.

Et sidste forhold, der blev identificeret i observationerne, var elevernes hurtighed til selv at opsøge viden, undervisningsrelevant eller ej, og det gælder alle typer af læsere. Tages denne betragtning seriøst, hvad vi mener den skal, så vil det have en enorm indvirkning på deres læring. Knud Illeris har udviklet en model for læring, der bygger på en omfattende undersøgelse af læringsteorier. Modellen indeholder de væsentligste dimensioner af læring:

den indholdsmæssige, drivkraftsmæssige og samspilsmæssige dimension (Illeris, 2013).

Figur 8.
Læringstrekanten med de tre dimensioner. (Illeris, 2013, s. 42)



I denne artikel vil vi kun folde dimensionen *Indhold* ud, da det er den dimension hos Illeris, som vi mener bliver mest påvirket af den adfærd, vi har observeret i forbindelse med elevernes læsning. Den indholdsmæssige dimension defineres sådan:

> ” Indholdet er det, der læres. Man kan ikke meningsfuldt tale om læring, uden at der er et læringsindhold, et eller andet, der skal læres. Det kan fx have karakter af viden, kundskaber, færdigheder, forståelse, indsigt, mening, holdninger eller kvalifikationer, og der kan også bruges andre betegnelser.
> (Illeris, 2013, s. 37)

Elevernes brug af egne selvvalgte digitale medier i deres læsning betyder, at lærerne og det enkelte læremiddel kan udøve mindre indflydelse på denne dimension, og at eleverne skal oplæres til at kunne foretage en kvalificeret vurdering af den enkelte kilde, de anvender, som ligger udover, hvad læreren har bygget sin undervisning eller faglige læsning op omkring. Som beskrevet kan forholdet have positive egenskaber i form af kvalificeret multiplexing, men også mindre positive effekter i form af distraktioner fra det intenderede formål med læsningen.

Gymnasieskolen ikke under forandring

Vores undersøgelser, designprincipper og prototypers iteration viser en dansk gymnasieskole, der på den ene side ikke har ændret sig i sit læringsteoretiske ståsted og på den anden side udfordres af elevers digitale vaner, læsestrategier og teknologier, der finder vej ind i læringsrummet.

Gymnasieskole som uforandret læringsarena
Med de observationer, vi foretog, levede erfaringsbeskrivelser, der blev indhentet, og interviews afholdt som led i prototypernes iterative designudvikling, var det tydeligt, at særligt to læringsteorier er dominerende for den undervisning, vi har observeret som del af projektet. Disse er den socialkonstruktivistiske læringsteori og den pragmatiske læringsteori, der kom til udtryk i følgende aktiviteter:

Figur 9.
Observerede aktivitetstyper på de tre gymnasieretninger

	STX	HHX	HTX
Undervisningsdesign	Individuel læsning Gruppearbejde	Individuel øvelse Gruppearbejde	Laboratorieforsøg Gruppearbejde
Læringsaktivitet	Opgaveløsning	Opgaveløsning	Opgaveløsning Fællesskrivning

Det var med vores observationer tydeligt, at undervisningsdesign baseret på socialkonstruktivismen var dominerende. Her var det læringsteoretiske socialkonstruktivistiske udgangspunkt Lev Vygotskys teori om individets kognitive lærings interdependens med den sociale og sproglige læring, der opnås igennem dialogen med andre og andet ud fra den lærendes zone for nærmeste udvikling (Harasim, 2017, s. 68 - 70). Læringsaktiviteten, der blev knyttet til det socialkonstruktivistiske udgangspunkt, var i høj grad opgaveløsning af en pragmatisk karakter. Her var det primært det pragmatisk læringsteoretiske udgangspunkt, som vi finder hos John Dewey, der som de centrale præmisser i sin teori har opdelingen i *instrumenter* og *eksperimentel.* Han mener, at individet med udgangspunkt i dennes indlærte *instrumenter*, der består af teorier og begreber, og med en *eksperimentel* tilgang, der består af handling og refleksion over egen væren i verden, opnår kompetencer til at skabe mening samt forstå tvivl og usikkerhed iboende i individet. Den empiri, man som individ anvender til at skabe mening i sin egen verden, er de erfaringer, som de enkelte hver især besidder (Elkjær, 2019, s. 69-79).

At de observerede undervisningsdesigns bygger på henholdsvis socialkonstruktivistisk og pragmatisk læringsteori, er ikke overraskende, da Bekendtgørelse af lov om gymnasiale uddannelser (2022) bl.a. skriver:

> ” Paragraf 1, Stk. 2. Eleverne skal gennem uddannelsens faglige og pædagogiske progression udvikle faglig indsigt og studiekompetence. De skal opnå fortrolighed med at anvende forskellige arbejdsformer og opnå evne til at fungere i et studiemiljø, hvor kravene til selvstændighed, samarbejde og sans for at opsøge viden er centrale.

Vi finder særligt ordene ”studiemiljø” og ”samarbejde” centrale i denne sammenhæng, hvor begge ord indikerer en social dimension som central for uddannelsens aktiviteter. Her er det også interessant, at selvstændighed er en vigtig kompetence, men vi forstår ikke nødvendigvis selvstændighed som en isoleret arbejdsform, men derimod sådan at selvstændighed også kan dyrkes som en del af f.eks. kollaborative aktiviteter, hvor selvstændige bidrag i en gruppe også kan kvalificere såvel den enkelte som gruppens læringsudbytte.

Ydermere er det heller ikke overraskende, at pragmatisme dominerer de læringsaktiviteter, vi har set, da der er fire retninger indenfor den danske gymnasieskole, der på forskellig vis kan redegøres for som følgende (Dolin et al., 2017, s. 21-28):

— *Almen retning.* Indenfor det almene gymnasium ligger såvel STX og HF som uddannelsesretning. Begge har almen dannelse, men med udgangspunkt i en forskelligartet faglig og didaktisk tilgang til det almendannende felt. Studentereksamen (STX) har sit faglige fokus på det humanistiske, samfundsvidenskabelige og naturvidenskabelige dannelsesperspektiv, sådan at fagenes anvendelse kommer i spil ud fra et teoretisk udgangspunkt i undervisningen. Her ligger den faglige fordybelse på anvendelse af videnskaberne. Højere Forberedelseseksamen (HF) har det samme dannelsesperspektiv som STX, men for HF er fokus i undervisningen lagt på sammenhæng mellem fagene i et professionsorienteret perspektiv, hvor fagenes anvendelse sættes i relation til videreuddannelse og arbejdsmarkedet. HF er desuden den eneste af de fire, der ikke giver direkte adgang til universitetsstudium, men i sit uddannelsesdesign i stedet har professions- og akademiuddannelser som sit primære sigte.

— *Merkantil retning.* Højere handelseksamen (HHX) har sit faglige fokus på det merkantile og internationale dannelsesperspektiv med anvendelsen af innovation og problemløsning på virkelighedsnære problemstillinger i undervisningen. Dette ud fra en faglig fordybelse, der anvender dannelsesperspektivet til analyse af virkelighedsnære cases gennem en kombination af teoretisk viden og analytisk beredskab.

— *Teknologisk retning.* Højere Teknisk Eksamen (HTX) har sit faglige fokus på det teknologiske og naturvidenskabelige dannelsesperspektiv, hvor innovation, produktudvikling og problemløsning anvendes i kombination mellem teori og praksis i undervisningen. Dette udført i værksteder og laboratorier ud fra en faglig fordybelse, idet man anvender teknologisk og naturvidenskabelig teoretisk viden til analyse af virkelige cases.

De fire retninger er designet med henblik på at inddrage et pragmatisk læringsteoretisk perspektiv på dannelsesperspektiver og virkelighedsnære problemstillinger, som de hver især har som DNA i sin uddannelse, og alle med undtagelse af STX har fokus på enten professionsrettede eller virkelighedsnære undersøgelser, der fordrer en mere pragmatisk orienteret læringsteoretisk tilgang.

Af vores observationer fremtræder det tydeligt, at kernen i gymnasieskolen er intakt, hvad angår til undervisningsdesigns fundering i den traditionelle læringsteori, med udgangspunkt i den særlige DNA, der hører til de enkelte uddannelsesretninger indenfor gymnasieskolen. Udfordringerne til denne virkelighed er til gengæld til stede i form af bl.a. de digitale læremidler, som eleverne anvender, hvilket påvirker det klassiske ekkokammer fra både et lærer-, elev- og bekendtgørelsesperspektiv.

Gymnasieskolens ekkokammer under udfordring

Som nævnt tidligere udviser eleverne en høj grad af singletasking og multiplexing i de læringsaktiviteter, de arbejder med. Observationerne underbygger ligeledes, at eleverne i høj grad anvender en bred vifte af teknologier, når de arbejder som led i deres undervisning. Samlet falder dette ind under begrebet *medieøkologi*, der dækker over de digitale teknologier, som eleverne aktiverer som del af egen studielæsning og opgavebesvarelse (Caviglia et al., 2018). Elevernes medieøkologi er meget forskelligartet, alt efter hvor teknisk interesserede og kapable de er, men fælles for vores observationer er, at eleverne har en tendens til f.eks. at foretage Google-søgninger frem for at læse i det fagkuraterede udleverede materiale fra underviseren.

Selvom det på den ene side er en god kompetence at kunne udøve for eleven, er det problematisk i forhold til gymnasieskolens struktur, der bygger på underviserens tilrettelæggelse af undervisning med det fagmateriale, som denne finder relevant, hvilket eleverne prøves i ved en afsluttende eksamen eller årsprøve. At eleverne derfor, som led i deres læringsaktivitet, anvender teknologierne i deres egne medieøkologier, medfører, at undervisningsdesignet og læringsaktiviteterne ikke altid korrelerer, hvilket medfører, at underviserens overblik over sine elevers faglighed er udfordret. Dernæst har særligt VUC´erne indført uddannelsesdesigns, der bl.a. bygger på blended learning, hybrid learning og fjernundervisning, hvor eleverne fysisk ikke er i samme rum som underviseren, hvilket kan medføre et ringere fagligt overblik for undervisere, der ikke besidder de nødvendig IT-didaktiske kompetencer.

For at imødegå dette indarbejdede vi i vores endelige designprincipper elevernes medieøkologi, idet vi ud fra analyse af undersøgelser og diskussion udvidede vores første udgave af designprincipper, som vi med fordel vurderer er relevante i udarbejdelsen af digitale læremidler. Projektets samlede afsluttende designprincipper så ud som følgende:

- Læremidlet skal invitere til teksthændelser, hvor elever i et projekt- eller værkstedsrum læser og løser opgaver sammen i en rig medieøkologi.
- Læremidlet skal holde et skarpt fokus på indholdsfaglig kognitiv belastning, og urelateret støj skal undgås.
- Der skal skabes rum for refleksion over anvendelse af egen medieøkologi hos eleverne.
- Læremidlet skal i sit design udvikle og understøtte elevernes dybde- og normallæsningskompetence.
- Læremidlet skal indlejret kunne understøtte studietekniske kompetencer.

Ved at indarbejde den læringsvirkelighed, vi har undersøgt i læremidlet, fastholdes elevernes faglige fokus i det materiale, som er udvalgt af og anvendes af såvel undervisere som de institutioner, de er tilknyttet. Læremidlet er, jf. vores designprincipper, et lukket miljø, men imødekommer den adfærd og de kompetencer indenfor særligt multiplexing og medieøkologi, som ligeledes er en væsentlig kompetence indenfor læring i det 21. århundrede, men det anderledes i vores designprincipper er, at udgangspunktet er det fagkuraterede, didaktiserede materiale, som danner grundlag for det fælles faglige udgangspunkt i fagene, hvor underviseren har den faglige kontrol.

Undervisernes kompetencer

For at være fastansat gymnasielærer kræves et gennemført pædagogikum, som er et uddannelsesforløb, der forløber over et skoleår, og som består af en teoretisk og en praktisk del. Her skal underviserne igennem uddannelsen introduceres til såvel didaktik og pædagogisk teori indenfor de gymnasiale uddannelser, de er ansat på, samt fagspecifik læringsteori og praksis indenfor de fag, de har undervisningskompetence i (STUK, 2018). Pædagogikum for gymnasielærere danner dermed det læringsteoretiske, pædagogiske og praktiske grundlag for deres virke som undervisere, med udgangspunkt i den akademiske fagekspertise, de har erhvervet sig på de universiteter, hvor de er uddannet. En af de udfordringer, vi imidlertid ser med den øgede teknologiske påvirkning af læringsrummet, er, af:

— Pædagogikum for gymnasielærere skal skabe reflekterende praktikere som gymnasielærere, med den traditionelle læringsteori og didaktik som omdrejningspunktet. Det forhold, at digitale læremidler for alvor er brudt igennem indenfor det seneste årti, betyder, at didaktisering af teknologier i forhold til analoge læremidler, er en lille del af pædagogikums samlede formål, da det er en forholdsvis ny og uberørt virkelighed i gymnasieskolen.
— Underviserne har dermed ikke fået det nødvendige kompetenceløft til at kunne transformere teknologierne til udvikling og forlængelse af faglighed, hvilket ligeledes kommer til udtryk i DiDaK-projektet (Dalsgaard et. al., 2020, s. 58 - 67).
— Kompetenceudvikling af undervisere er dermed et ansvar for de enkelte uddannelsesinstitutioner, hvor udviklingen de seneste år har været hæmmet af besparelser på de gymnasiale uddannelser, hvorfor fokus ikke alle steder har været kompetenceløft af undervisere indenfor bl.a. digitale læremidler.

Derfor vurderer vi, at designprincipperne, vi foreslår for de digitale læremidler, suppleres med kompetenceløft af undervisere, mhp. at anvende teknologier som udgangspunkt for deres praksis frem for et element, der supplerer praksis.

Her er det væsentligt, at kompetenceløftet finder sted med udgangspunkt i den didaktiske og pædagogiske virkelighed og tradition, der eksisterer i forvejen i gymnasieskolen, hvilket som nævnt er solidt forankret i den klassiske læringsteori og pædagogik. Derfor er det væsentligt at forholde sig til teknologier som læringsteknologier, der i højere grad moderniserer og aktualiserer en læringsvirkelighed, frem for at transformere til noget andet, der i værste fald kan underminere gymnasiebekendtgørelsen.

Konklusion

Vores empiri har anskueliggjort, at elevernes digitale faglige læsning er præget af dels forskellighed, dels mangel på studielæsning. Vores observationer i undervisningslokaler viser, at eleverne i højere grad scanner end dybdelæser i undervisningssituationer, og de levede erfaringsbeskrivelser understøtter denne tendens ved elevers læsning hjemme. Ydermere har vi erfaret, at elevers anvendelse af digitale læremidler i høj grad præges af en long-time-practitioner-adfærd, styret af vaner, og at læsning i høj grad er opgavecentreret og ikke fordybelsesorienteret. Særligt vores observationsstudier afdækkede anvendelse af digitale læremidler som et konstruktivistisk værktøj, der oftest var socialkonstruktivistisk faciliteret ud fra en pragmatisk funderet lærerstillet opgaveløsning. Derfor er brugen af digitale læremidler i udpræget grad forbundet med en lærerstillet opgave, og her agerer elever meget forskelligt, hvilket bunder i en differentieret anvendelse af elevernes personlige medieøkologier. Vi har på baggrund af vores empiri identificeret og redegjort for, hvad vi har karakteriseret som fire digitale elevtyper, digitale solister, digitale drømmere, digitale passagerer, digitale sociale skabere, der hver især anvender digitale læremidler forskelligt i læringssituationer. Interessant for de digitale elevtyper er, at de ikke er statiske, men at samme elev kan skifte digital elevtype efter fag og fysisk lokation. Eksempelvis viste det sig, at elever, der var digitale drømmere i stillesiddende undervisning, agerede digitale passagerer i værksteder, f.eks. et kemilaboratorie på HTX. Projektet har igennem dets mange delelementer og DBR som metodologi for vores udvikling fra domænekendskab, teoretisk forankring, indsamling af empiri til brug for designprincipper og slutteligt udvikling af prototyper medført, at vores indledende forståelse af gymnasieelevers faglige digitale læsning i digitale læremidler er en større kompleksitet, end vi indledningsvist havde troet. Særligt elevernes forskelligartede anvendelse af digitale teknologier viser os, at de i høj grad supplerer de digitale læremidler, der stilles til rådighed fra gymnasierne, hvor bl.a. Google søgninger m.m. er en indgroet del af deres udarbejdelse af faglige aktiviteter, hvilket er belyst igennem projektets enkelte faser. Dette rykker imidlertid ved, hvilke indholdsdimensioner der er henholdsvis udvælgelsen og anvendelsen af digitale teknologier, hvilket betyder, at eleverne fagligt kan udvikle sig på måder, som de analoge læremidler ikke understøtter, men også at undervisernes påvirkning af det faglige indhold i den didaktiske trekant bliver udfordret, såfremt digitale læremidler eller undervisningsaktiviteter ikke forankres i kuraterede teknologier. Underviserne har ikke været inddraget i projektet, men såfremt projektet skulle udvides til yderligere faser, vil en inddragelse af underviserne være en naturlighed, idet der skulle være særskilt fokus på undervisernes kompetencer, også i forhold til forståelse af egne kompetencer overfor den teknologiske virkelighed de befinder sig i mhp. kompetenceudvikling.

Referencer

Balling, G. (2017). Unges læsning i en digital tidsalder. LEARNING TECH – TIDSSKRIFT FOR LÆREMIDLER, DIDAKTIK OG TEKNOLOGI, (3), 56-84.

Blikstad-Balas, M. (2016). Skolens nye literacy – Hvordan endres skolens tekstpraksiser når digital teknologi er tilgjengelig i klasserommet. Learning Tech, 1(1), 77-98. https://doi.org/10.7146/lt.v1i1.107724

Blikstad-Balas, M. & Hvistendahl, R. (2013). Students' Digital Strategies and Shortcuts. NORDIC JOURNAL OF DIGITAL LITERACY, 8(01-02), 32-48.

Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research In Psychology, 3(2), 77-101. 10.1191/1478088706qp063oa

Bekendtgørelse af lov om de gymnasiale uddannelser. (2022). Bekendtgørelse af lov om de gymnasiale uddannelser § 1, stk. 2 https://www.retsinformation.dk/eli/lta/2022/957

Bower, M. (2008). Affordance analysis – matching learning tasks with learning technologies. Educational Media International, 45(3). 10.1080/09523980701847115

Carlsen, D. & Hansen, J. J. (red.). (2015). Digital læsedidaktik (1. udg.). Akademisk Forlag.

Caviglia, F. & Dalsgaard, C. (2020). Introduktion til digitale kompetenceområder. Center for Undervisningsudvikling og Digitale Medier, Aarhus Universitet.

Caviglia, F., Dalsgaard, C., Davidsen, J. & Ryberg, T. (2018). Studerendes digitale læringsmiljøer: læringsplatform eller medieøkologi? Tidsskriftet Læring og Medier (LOM), 11(18).

Dalsgaard, C., Caviglia, F., Boie, M. A. K., Færgemann, H. M. & Thomsen, M. B. (2020). Digitale kompetencer i gymnasiet. Center for Undervisningsudvikling og Digitale Medier, Aarhus Universitet.

Dalsgaard, C. & Ryberg, T. (2022). *Digitale læringsrum*. Samfundslitteratur.

Dam, R. F. & Siang, T. Y. (2022, 2. maj). *Test Your Prototypes: How to Gather Feedback and Maximize Learning*. Interaction Design Foundation. Hentet 19. april, 2023, fra https://www.interaction-design.org/literature/article/test-your-prototypes-how-to-gather-feedback-and-maximise-learning

Dolin, J., Ingerslev, G. H. & Jørgensen, H. S. (red.). (2017). Gymnasiepædagogik: en grundbog (3. udg.). Hans Reitzel.

Elkjær, B. (2019). Pragmatisme – Læring som kreativ fantasi. I K. Illeris (red.), 15 aktuelle læringsteorier. Samfundslitteratur.

Garrison, D. R. (2016). *Thinking Collaboratively: Learning in a Community of Inquiry*. Taylor & Francis Group, an Informa business.

Hansen, J. J. (2010). Læremiddellandskabet: fra læremiddel til undervisning (1. udgave). Akademisk.

Harasim, L. (2017). Learning Theory and Online Technologies. Routledge.

Illeris, K. (2013). Læring (2. udgave). Roskilde Universitetsforlag.

Jensen, Lasse Bo & Nielsen, Michael Juul: *Læsning og læring i danske gymnasier - med fokus på understøttelse af forskellige læsestrategier og arbejdsformer gennem teknologier i digitale læremidler.* https://projekter.aau.dk/projekter/files/535196383/Masterprojekt___Lasning_og_laring_i_danske_gymnasier.pdf

Kirschner, P. A., Kirschner, F. & Janssen, J. (2014). The Collaboration Principle in Multimedia Learning. I R. E. Mayer (red.), *The Cambridge Handbook of Multimedia Learning* (pp. 547-575). Cambridge University Press.

Lim, F. V. & Toh, W. (2020). How to teach digital reading? Journal of Information Literacy,14(2). http://dx.doi.org/10.11645/14.2.2701

Mayer, R. E. (2014). Cognitive Theory of Multimedia Learning. I R.E. Mayer (ed.), The Cambridge Handbook of Multimedia Learning (pp. 43-71). Cambridge University Press. doi:10.1017/CBO9781139547369.005

Mayer, R. E. & Fiorella, L. (2014). Principles for Reducing Extraneous Processing in Multimedia Learning: Coherence, Signaling, Redundancy, Spatial Contiguity, and Temporal Contiguity Principles. I R. E. Mayer (ed.), *The Cambridge Handbook of Multimedia Learning* (pp. 279-315). Cambridge University Press.

Nørgaard, T. B. K. (2023, 22. januar). Tesfaye: »Hvis jeg var skoleleder, havde jeg kørt en meget restriktiv skærmpolitik«. Politiken. Hentet 28. januar, 2023, fra https://politiken.dk/indland/art9170439/%C2%BBHvis-jeg-var-skoleleder-havde-jeg-k%C3%B8rt-en-meget-restriktiv-sk%C3%A6rmpolitik%C2%AB

Orlikowski, W. J. (2010). The sociomateriality of organisational life: considering technology in management research. Cambridge Journal of Economics, (34), 125–141. 10.1093/cje/bep058

Samar, R. G. & Dehqan, M. (2013). Sociocultural theory and reading comprehension: The scaffolding of readers in an EFL context. International Journal of Research Studies in Language Learning, 2(3), 67-80. 10.5861/ijrsll.2012.183

Sharp, H., Rogers, Y. & Preece, J. (2019). Interaction Design: Beyond Human-Computer Interaction (5. udgave). Wiley.

STUK (Styrelsen for Undervisning og Kvalitet) (2018): Bekendtgørelse om pædagogikum i de gymnasiale uddannelser. https://www.retsinformation.dk/eli/lta/2018/1169

Tække, J. & Paulsen, M. (2017). Undervisning og digitale medier – Distraktion, koncentration og engagement. Tidsskriftet Læring Og Medier (LOM), 11(18), 1-22. https://doi.org/10.7146/lom.v10i18.26242

Tække, J. & Paulsen, M. (2019). Distraction and digital media – multiplexing, not multitasking in the classroom. Tidsskriftet Læring og Medier, (21).

Van Manen, M. (1990). Researching Lived Experience: Human Science for an Action Sensitive Pedagogy. State University of New York Press.