UiT The Arctic University of Norway

Faculty of Humanities, Social Sciences and Education

GenAI in the English subject – from the pupils' perspective

A mixed method study on pupils' actual and wanted use of generative artificial intelligence in the English subject

Sondre Ingerøyen

Master's thesis in English didactics, LER-3902, May 2024

Acknowledgements

Deciding to write a Master's thesis about a subject such as GenAI, which had not been done before, put me firmly on a journey of discovery for the past months, on which I often felt alone. Luckily for me, I have been surrounded by fellow students, family and friends who have consistently shown great interest in my thesis, encouraging me with comments on how current and interesting they find the subject matter. In a way, this has become my motivation for this journey, knowing that part of the reason why it is difficult is also what makes it interesting. Thank you all.

I would like to thank the teachers at the school where the research was conducted, who were enormously accommodating and helpful throughout the data collection process, and the pupils who participated, without which there would be no research to write about.

I would also like to thank my supervisor, Hilde Brox, for her continuous enthusiasm for the project and for her always solution-oriented advice. Not all supervisors send emails with links to news segments at nine o'clock in the evening on a Thursday, thinking it could be relevant to the project, but you do, and it made me feel that you wanted me to succeed in my work. Thank you.

Sondre Ingerøyen

Tromsø, May 15th 2024

Abstract

This Master's thesis aims to explore how pupils use generative artificial intelligence (GenAI) in the English subject and how pupils would like to use the technology in the future. This research objective was based on the lack of research into both the pupils' perspectives on using GenAI, as well as subject-specific research on the use of the technology. The thesis is based on the findings of a mixed-method research project, where a survey was conducted to explore pupils' use of GenAI in the English subject, and group interviews were conducted to explore their views on the future use of GenAI. The research project was conducted at an age-mixed school in northern Norway in January and February 2024, with 39 participants, of which 18 took part in group interviews.

The findings presented in this thesis show that a majority of the pupils taking part in the survey have used GenAI for tasks in the English subject. Different types of use are also shown, with a majority of the pupils saying they have used GenAI for explanation of difficult material, as a "discussion partner" for inspiration, translation, and to proofread text. The findings also show a difference in the use of GenAI between pupils with different levels of perceived competence in the English subject. A thematic analysis of the interview data was conducted, and the emerging themes of how pupils want to use GenAI in the future are presented. These themes of wanted use mostly mirror the use cases from the survey, namely: inspiration, explanation, help, information, and improving language.

In the discussion, the findings are seen in the context of national and international guidance documents on the use of GenAI in education, as well as previous research on the subject. The pupils' statements in the group interviews show a high level of insight into the possible dangers of becoming over-reliant on GenAI, and their reflections on the subject are reminiscent of several points highlighted in the different guidance documents. The thesis concludes with a call for further research into the field of GenAI so that educated decisions and policies can be made.

Keywords: *Education, schools, GenAI, AI, ChatGPT, English language learning, mixed method research, pupils' perspective, chatbots, achievement levels, wanted use.*

Sammendrag

Denne masteroppgaven tar sikte på å utforske hvordan elever bruker generativ kunstig intelligens (GenAI) i engelskfaget og hvordan elevene ønsker å bruke teknologien i fremtiden. Dette målet bygger på mangelen på forskning både når det gjelder elevenes perspektiver på bruk av GenAI, samt fagspesifikk forskning om bruken av teknologien. Oppgaven er basert på funn fra et forskningsprosjekt med både kvantitative og kvalitative metoder. En spørreundersøkelse ble gjennomført for å utforske elevenes bruk av GenAI i engelskfaget, og gruppeintervju ble gjennomført for å utforske deres syn på fremtidig bruk av GenAI. Forskningsprosjektet ble gjennomført ved en aldersblandet skole i Nord-Norge i januar og februar 2024, med 39 deltakere, hvorav 18 deltok i gruppeintervjuer.

Funnene som presenteres i denne oppgaven viser at flertallet av elevene som deltok i spørreundersøkelsen har brukt GenAI til oppgaver i engelskfaget. Forskjellige typer bruk vises også, der flertallet av elevene sier de har brukt GenAI til forklaring av vanskelig materiale, som en "diskusjonspartner" for inspirasjon, oversettelse og korrekturlesing av tekst. Funnene viser også en forskjell i bruken av GenAI mellom elever med ulik opplevd kompetansenivå i engelskfaget. En tematisk analyse av intervjudataene ble gjennomført, og de fremtredende temaene for hvordan elevene ønsker å bruke GenAI i fremtiden er presentert. Disse ønskede bruksområdene ligner i stor grad på bruksområdene fra spørreundersøkelsen, nemlig: inspirasjon, forklaring, hjelp, informasjon og forbedring av språk.

I diskusjonen sees funnene i sammenheng med nasjonale og internasjonale veiledningsdokumenter om bruk av GenAI i utdanningen, samt tidligere forskning på feltet. Elevenes uttalelser i gruppeintervjuene viser en høy grad av innsikt i de mulige farene ved å bli for avhengig av GenAI, og deres refleksjoner om emnet minner om flere punkter som er fremhevet i de ulike veiledningsdokumentene. Oppgaven konkluderer med en oppfordring til videre forskning på feltet GenAI, slik at informerte beslutninger og retningslinjer kan utvikles.

Nøkkelord: *Utdanning, skole, GenAI, AI, ChatGPT, engelsk språklæring, mixed methods, elevperspektiv, chatbots, prestasjonsnivåer, ønsket bruk.*

Table of Contents

List of Figures

Glossary

Large Language Model (LLM): A type of AI that can process and produce natural language text. It learns from a massive amount of text data such as books, articles, and web pages to discover patterns and rules of language.

Generative Pre-Trained Transformer (GPT): A type of LLM that is pre-trained on even larger amounts of data, which allows the model to capture the nuances of language and generate coherent context-aware text.

Chatbot: In this thesis, the term chatbot refers to applications that utilise LLMs to simulate human conversation by producing human-like text in response to user prompts. Examples include ChatGPT, Copilot and Gemini.

Artificial Intelligence (AI): A term for technology that enables computers and machines to simulate human intelligence and problem-solving capabilities.

Generative AI (GenAI): Generative AI includes all artificial intelligence that can generate text, images, videos, or other content based on the data they were trained on. Examples of GenAI include all the chatbots mentioned above, but also text-to-image models such as DALL-E, Midjourney and Stable Diffusion, and text-to-video models such as Sora.

ChatGPT: ChatGPT is a chatbot, a task-specific model which uses GPT-3.5 as its foundational Large Language Model and is fine-tuned for conversation.

1 Introduction

As I write this introduction, it has been 18 months since the popularisation of generative artificial intelligence, or GenAI. This happened in the wake of the public release of OpenAI's ChatGPT in November of 2022, and there are still no clear guidelines for how these tools should be approached in schools. The lack of guidelines leaves it up to teachers to decide to what degree one should educate pupils in the use of such tools and to what degree one should allow and encourage the use of such tools for school purposes. The speed of innovation is still staggering in this field, and brings further difficulties to the creation of rules and regulations regarding GenAI. For example, in contrast to the GPT-3 powered ChatGPT that was trained on language scavenged from the internet up to 2021, the GPT-4 based GenAI-tools of today can search for information on the internet and use this information in real-time. This enables them to find and cite recent sources accurately and suggest further reading on the subject in question. In the final days of writing this thesis, GPT-4o was released, introducing features such as communicating through audio and video with the model (OpenAI, 2024).

Knowledge on the subject of AI is lacking, both among pupils, teachers, and society as a whole. A process has been started to regulate the use of GenAI, but this is a slow process, as it is a revolutionary development with many facets and, therefore, many different views to consider. Politicians ask for research to guide the decision-making process, but research takes time, and there are no simple answers as to how we should approach the innovations happening in the GenAI field. The Norwegian government have expressed the need for greater competence in the field and more research so that the education sector can be prepared for the extensive changes we stand before (Ministry of Education and Research, 2023). Earlier research on technological aids in schools exists, and can be used to guide research into GenAI. Still, GenAI is so unlike any previous technological advancement that generalisation from such research will be of limited value compared to targeted research into this groundbreaking technology.

1.1 Background and motivation

Several factors contributed to my motivation towards choosing GenAI as the theme for this research project; my own interest in the fields of technology and Computer Science being a prominent one. This interest fuelled my motivation to follow the development of different

GenAI's through the spring of 2023, and to experiment with several of the popular GenAI's and the services they had to offer. The insight I gained during this period made me realize that GenAI would change the way I would have to think about the English subject as a whole, as research, essays and other student work was forever fundamentally changed. I was, however, in the position to lay the foundations for that work as a part of my own master thesis project. Motivated by this prospect, I sought out to find a way in which I could make a meaningful contribution to the field, however small, which might in turn inform later research to some extent.

After deciding on GenAI in the English subject as the theme for my research, I did some preliminary research into GenAI in schools to help inform my decision on narrowing down my overarching theme into something more specific. The research available at this time (August 2023) was very limited, and what limited research existed was largely done by private firms, who sold access to the data at a premium. One example of this is KANTAR, which, according to their website, charges 30.000 NOK for access to their report on media habits for Norwegian 15-24-year-olds, which includes information regarding the use of ChatGPT (KANTAR, 2023). Some of the results from this research were, however, published by news outlets, e.g. the Norwegian Broadcasting Organization - NRK (Zulic, 2023), which gave me a small insight into the data. It showed some statistics on pupils' use of GenAI, but gave little information on the nature of the research, and the data shown was exclusively of a quantitative nature. None of the published results from this report, or from any other research that I could find, focused on subject-specific use of GenAI, i.e. GenAI within the English subject. This helped inform my decision in regard to what research was already being done, and as a result, steered me in the opposite direction- toward a mixed-method design with a focus on qualitative data, with specific regards to the English subject.

1.2 Thesis question and research questions

My overarching goal for this research project is to gain a better understanding of pupils' use of, and wanted use of, GenAI in the English subject. As such, I formulated the following thesis question:

In what ways do pupils use GenAI in the English subject, and how do pupils believe it should be used?

To answer this main thesis question, I chose to further operationalise it into the following research questions:

1. **To which tasks and to what degree are pupils using GenAI in the English subject?**
2. **In what ways, if any, does the use of GenAI differ between pupils of different levels of achievement in the subject?**
3. **In what manner do pupils believe they should be able to use GenAI in the English subject?**

Research questions one and two are connected, as one would need data on both the use of GenAI and the pupils' level of achievement in the English subject to draw any conclusions on the subject. These questions were mainly chosen as I believe they offered an interesting angle to answering the part of the thesis question regarding the ways in which pupils use GenAI in the English subject. The other reason behind this choice is that I could not find any previous research that had explored pupils' use of GenAI in light of their levels of achievement in the English subject, although experts have alluded to GenAI as something that could "increase segregation in society" (Zulic, 2023).

The third research question focuses on the latter part of the thesis question, which is a dimension to the questions on the use of GenAI in the English subject and in education as a whole, which, as far as I could assess at the time, was unexplored. As the pupils currently in school are directly affected by the rise of GenAI as a tool, it seemed natural to me that their views on the subject warranted some consideration. This is a research question that cannot be answered by the same data as the first two, as data concerning pupils' current use of GenAI in the English subject would not capture potential use under different circumstances, e.g. if rules and regulations locally or nationally changed in regard to using GenAI.

1.3 Introduction to GenAI in education

To be able to discuss the findings of this research project, it is important to have an understanding of the status of the field regarding GenAI in education. As a result of the novelty of the technology, the choice to not include dedicated chapters on theory and previous research was made, opting instead to call it an introduction to the theme of GenAI in education.

The remainder of this chapter presents a summary of the subject of GenAI in Norwegian schools. It is divided into three parts: the first speaks to the uniqueness of GenAI compared to earlier technological implementations in schools, the second presents information on what constitutes an ideal implementation of the technology, and the third presents information on what we know about pupils' use of the technology.

1.3.1 GenAI is like no previous technology

Bigum (2012) summarises the ways in which educational institutions respond to technological innovations as an ongoing loop, wherein a new technology comes to market; arguments are made concerning the improvements the technology will make to existing teaching/learning practices; the acquisition of the technology is legitimised; the technology is implemented, and finally the technology is integrated into existing practices, or is rejected. Bigum also points out that while schools often approach new technology believing that it can improve something about schooling "(…) schools often use those technologies in old and familiar ways: integrating them into existing routines, deploying them to meet existing goals and, generally, failing to engage with technologies in ways consistent with the world beyond the classroom" (2012). While this cycle may have been historically accurate regarding schools' response to innovations as computers, digital learning resources and iPads, the popularisation of GenAI and its adopted use have not followed such a cycle. Being released for free to the general public, GenAI services like ChatGPT did not go through either legitimization or acquisition by the schools, as it simply became available to anyone with an internet connection. This means that, based on this loop, educational institutions simply had to jump to the final step; integrating the technology into existing practices, or reject it.

Simultaneously, as schools and teachers have had to take a stance on whether to integrate GenAI into their teaching practice or reject it, the debate on the subject has been widespread,

with teachers, researchers, politicians and students all voicing their opinions. A characteristic of this debate is that knowledge and research around GenAI in education is sparse, and such knowledge is substituted in the debate by anecdotal evidence. This is discussed by Krumsvik who, in his editorial, discusses the blending of everyday and research discourse on what he calls Large Generative Language Models like ChatGPT and GPT-4, emphasizing the need for systematic testing and research-based discussions. “The everyday discourse, where anecdotes dominated without being rooted in a knowledge base, and where no distinction was made between the capabilities of ChatGPT and GPT-4, shone through strikingly often.” (Krumsvik, 2024). Incidentally, the same editorial furthers its own point, by itself being inaccurate with terminology on the subject: The author refers to ChatGPT and GPT-4 as Large Generative Language Models, which in and of itself is not a term. ChatGPT is a chatbot which uses GPT-3.5 as its foundational Large Language Model and is fine-tuned for conversation, not simply a LLM. Similarly, comparisons are made between ChatGPT and BioGPT without mentioning that BioGPT is a model that is designed to be downloaded and run locally using specialised software (Luo et al., 2022), in contrast to a chatbot like ChatGPT which can be accessed via an internet browser or an app (OpenAI, 2022).

While some scholars call for more knowledge on the subject of GenAI in schools, experts in the field of artificial intelligence are also prominent in the debate. In a news story from the Norwegian Broadcasting Organization in August 2023, Professor Morten Goodwin, described as one of the most recognized figures in Norway on the subject of artificial intelligence, voiced several positive views on AI in a school context (Johansen, 2023). He points out that for pupils, artificial intelligence is much more than cheating: it is technology that can help the pupils improve. Goodwin exemplifies this by saying that pupils can improve their writing, possibly getting motivation to write longer texts that earlier have been too difficult, and burnt-out pupils might re-ignite their spark.

1.3.2 Ideal use

While experts in the field of AI voice their opinions on the promises of the technology, there is still limited information for schools to base their decisions on how to best implement the

technology in their practice, and what constitutes an ideal use of GenAI in a school setting. This part of the chapter presents both international and Norwegian guidelines available on the subject of use of GenAI in education, and their views on what constitutes best practice regarding such use.

International guidelines on GenAI in education

While the release of ChatGPT in November 2022 certainly took educational institutions, and the world, by surprise, high level discussions on artificial intelligence and education had already taken place years before. In May 2019, the International Conference on Artificial Intelligence and Education took place in Beijing, People's Republic of China. The participants included "50 government ministers and vice ministers, as well as around 500 international representatives from more than 100 Member States, United Nations agencies, academic institutions, civil society and the private sector" (UNESCO, 2019). The outcome document of this conference is titled "Beijing Consensus on Artificial Intelligence and Education", which consists of a list of 44 numbered consensuses. This list provides recommendations for actions that governments can consider implementing "in response to the education-related opportunities and challenges presented by AI", as well as recommendations for "international organisations and partners active in the field" and actions the Director-General of UNESCO is invited to seek to implement (UNESCO, 2019).

The recommendations for governments included planning and development of "system-wide strategies for AI in education that are aligned and integrated with education policies, within a lifelong learning perspective" (UNESCO, 2019) as well as an entire point on integrating AI into curricula:

> *14. Be cognizant of trends regarding the potential of AI to support learning and learning assessments, and review and adjust curricula to promote the in-depth integration of AI and transformation of learning methodologies. Consider applying available AI tools or developing innovative AI solutions, where the benefits of AI use clearly outweigh the risks, to facilitate well-defined learning tasks in different subject areas and supporting the development of AI tools for interdisciplinary skills and competencies (UNESCO, 2019).*

This point further Bigum's (2012) view discussed earlier that educational institutions often integrates new technology into existing routines and deploy them to meet existing goals, but is accompanied by another point in the consensus that brings attention to the lack of research in the field of AI in education, and recommends the following:

> *31. Be mindful of the lack of systematic studies on the impacts of AI applications in education. Support research, innovation and analysis on the effects of AI on learning practices and learning outcomes, and on the emergence and validation of new forms of learning. Take an interdisciplinary approach to research on AI in education. Encourage cross-national comparative research and collaboration (UNESCO, 2019).*

This is a stark contrast to the cycle outlined by Bigum (2012), where new technology is integrated into existing routines, in that already in 2019, the Beijing Consensus postulated that research in the field of AI in education could lead to the emergence of new forms of learning.

The consensus also includes several points on equitable and inclusive use of AI in education, one of which brings up the concept of a digital divide:

> *23. Ensure that AI promotes high-quality education and learning opportunities for all, irrespective of gender, disability, social or economic status, ethnic or cultural background, or geographic location. The development and use of AI in education should not deepen the digital divide and must not display bias against any minority or vulnerable groups (UNESCO, 2019).*

The Beijing Consensus was created before the mass availability of GenAI, and after the technology became available was used as background for the creation of UNESCO's Guidance on GenAI in Education and Research. In September of 2023, UNESCO published the "first-ever global Guidance on Generative AI in Education and Research, designed to address the disruptions caused by Generative AI technologies" (UNESCO, 2023). This book "aims to support countries to implement immediate actions, plan long-term policies and develop human capacity to ensure a human-centred vision of these new technologies" (Holmes & Miao, 2023).

One section of the book is titled “Towards a policy framework for the use of generative AI in education and research”, in which specific measures are proposed for the planning of policies on GenAI in education and research. A category of such measures are labelled “Protect human agency”, and is introduced by highlighting the danger of dependency on GenAI:

> *As more individual users use GenAI to support their writing or other creative activities, they might unintentionally come to rely upon it. This can compromise the development of intellectual skills. While GenAI may be used to challenge and extend human thinking, it should not be allowed to usurp human thinking (Holmes & Miao, 2023).*

One of the measures proposed in this section speaks to the need for need for research into the pupils’ perspective on the use of GenAI:

> *Consult researchers, teachers and learners about their views on GenAI and use the feedback to decide whether and how specific GenAI tools should be deployed at an institutional scale. Encourage learners, teachers and researchers to critique and question the methodologies behind the AI systems, the accuracy of the output content, and the norms or pedagogies that they may impose. (Holmes & Miao, 2023)*

Another measure proposed here says to “Use GenAI tools to minimize the pressure of homework and exams, rather than to exacerbate it.” (Holmes & Miao, 2023)

In the book, UNESCO also raises several uncharted ethical issues that need to be examined, among which is the issue of access and equity. One of the policies that are recommended concerning this, calls for action to be taken to “promote universal connectivity and digital competencies in order to reduce the barriers to equitable and inclusive access to AI applications”.

Norwegian guidelines for using GenAI in education

In April 2023, the Norwegian government released their strategy for digital competence and infrastructure in kindergartens and schools (Ministry of Education and Research, 2023). In the

foreword to the strategy, ChatGPT is mentioned as an example of how new frameworks and technology can challenge central parts of how education is being done in Norway. The purpose of the strategy is described as charting a new course, enabling (school) owners and employees to take control of the development, and to support them in the decisions that need to be made (Ministry of Education and Research, 2023).

In the chapter of the strategy concerning digital praxis in elementary and secondary schools, the strategy points to research overall showing that the use of digital devices and educational materials has a moderate positive impact on students' learning, and that this effect increases with age (Munthe et al., 2022). The strategy goes on to point out that we still lack knowledge about the effect on specific grade levels, individual subjects, and early education. The importance of clear goals and leadership in schools is also highlighted, as well as ensuring that teachers have up to date competencies.

Artificial Intelligence is addressed under a dedicated sub-heading in the chapter mentioned above, where several aspects of the implications the technology has on education are brought up. After listing several examples of uses of AI in schools, it is noted that each use case presents pedagogical possibilities, but also dangers to ensuring the pupils' learning if the teaching is not adequately planned. The strategy goes on to say that the knowledge base regarding the pedagogical opportunities and challenges of AI in education is limited, and that competence development and research in this area must be strengthened so that the education sector is prepared for the extensive changes we are facing (Ministry of Education and Research, 2023). It is also noted that the framework conditions and requirements necessary for AI-supported education and assessment to yield pedagogical and didactic benefits must be further investigated.

In February of 2024, the Norwegian Directorate for Education and Training (UDIR) published a guidance for schools and school-owners to support them in their work with AI. The guidance states that as society changes as a result of AI, schools must change also. It concedes that as the technology develops, our praxis, knowledge and laws concerning AI will lag behind, and that some challenges must be taken in stride while others must be worked on over time.

The Norwegian Directorate for Education and Trainings guidance provides nine points of advice on AI in schools, with in-depth information on each point, stating that each point have different levels of relevance for school-owners, leaders, teachers and parents (Norwegian Directorate for Education and Training, 2024c). A theme that is recurring in the list is teacher autonomy, as it is stated that the questions concerning if, when and how AI tools are used in education are pedagogical questions of which the teachers must take ownership. This entails deciding what use of AI is appropriate based on aims and values in the teaching, relevance to the subject, and the group of pupils in question.

One of the explicit pieces of advice the guidance gives, is to clarify the framework for the use of any AI tools prior to a learning session or period, including activities that will be assessed. It also says to consider whether it may be relevant to limit students' access to AI based on the planned learning in order to properly gauge the student's competence.

In addition to the guidance, The Norwegian Directorate for Education and Training has also created a digital competence package to support teachers, leaders and school owners in their work with AI (Norwegian Directorate for Education and Training, 2024a). They describe the package as an aid to lay the foundation for reflections to be made in professional learning communities concerning challenges and opportunities this technology will introduce to learning work. The competence package includes information about what AI is, how AI-based services work, as well as several challenges and opportunities the technology presents concerning teaching, assessment, source criticism, privacy and information security. The competence package is continuously updated and was relaunched in January 2024 with the inclusion of an module addressing challenges and opportunities regarding GenAI and assessment.

The competence package is recommended to be used in collaborative work in professional learning communities (Norwegian Directorate for Education and Training, 2024a), and is published on the Canvas platform and is accessible through login via FEIDE or by providing an email address to create a user (Norwegian Directorate for Education and Training, 2024b). At the time of writing in May 2024, the competence package is only available in Norwegian, and changing the preferred language on the platform does not translate the content of the package.

1.3.3 Data on pupils' use of GenAI

As mentioned in the Background chapter of this paper, I was mostly unable to find previous research on the subject of pupils' use GenAI in August 2023 when I started the work of this research project, and that the little I found was inaccessible behind a paywall. I relied on news organizations' coverage of the release of reports to get insight to their contents, as in the example mentioned in Chapter 1.1 regarding Kantar's media trend report. The Norwegian Broadcasting Organization (NRK) reported Kantars findings, stating that 18% of a representative sample of youths aged 15-24 use ChatGPT for schoolwork weekly (Zulic, 2023). During the past months, however, more reports have been published in both Norway and from all over the world concerning pupils' and students' use of GenAI.

Several surveys of Norwegian students have included questions regarding use of GenAI, and the largest one to do so as of writing was published in February 2024 by The Norwegian Agency for Quality Assurance in Education, called "The Student Survey", in which 29,000 students from universities and university colleges from all over Norway responded to a questionnaire (Hauge & Øygarden, 2024). Their findings regarding GenAI show that "60% of students report that they have used AI in their studies, while 40% have not" (Hauge & Øygarden, 2024). The survey also found that the most common use case of AI among students was to "explain subjects, syllabi, concepts, terminology etc." which was selected by 67% of students who had used AI, with around 40% selecting the options; as a "discussion partner", inspiration for tasks, brainstorming etc.; quality assurance/editing self-written text; shorten texts (articles, syllabus, own texts, etc.). 23% answer that they use AI to generate new text.

The finding in the Student Survey, that 60% of students have used AI, is almost identical to another survey conducted on behalf of the Norwegian Broadcasting Organization. This survey, conducted on a sample of 1000 people showed that 59% of 18-29 year olds had used ChatGPT for school, studies or work in the past year (Vik, 2024). The same survey show that one in eight of 18-29 year olds have used AI for something they are unsure whether their place of study or work would allow. Four percent of the same age group answer that they have used AI for something they know is prohibited.

While most reports that I have found concerns students in higher education, one report on the use of generative AI in British secondary schools was found. In January 2024, the British Department of Education published a report on the use of generative AI in education, based on insights from educators, experts, quantitative data sources and grey literature (The Open Innovation Team & British Department for Education, 2024). The report presents statistics gathered from the British Department for Education's Parent, Pupil and Learner Panel survey, where 14% of pupils years 7-13 (i.e. aged 11-17) are showed using AI tools for schoolwork. Among those who answered that they had used AI tools for schoolwork 61% reported using these tools for homework, 40% to learning at home, and 29% in lessons at school.

The report also points to Ofcom (OFCOM, 2023), who reports that 79% of online teenagers aged 13-17 and 40% of online 7–12-year-olds reporting that they had used ChatGPT, Snapchat My AI, Midjourney or DALL-E. The most commonly used GenAI tool was in this instance found to be Snapchat My AI, which was reported having been used by 72% of children aged 13-17, and 30% of children aged 7-12. The data is from the CHILDWISE Summer Omnibus, which is "based on fieldwork conducted from June to July 2023, from 1,504 online interviews with boys and girls aged 7-17, plus 19 tweens and teens who identify outside the gender binary. The final data was weighted to restore complete balance by age and gender." (OFCOM, 2023).

The report also points out that other surveys conducted in Britain have figures higher than what is found in the Department for Educations survey, e.g. RM Technology who found in a survey of 1,000 secondary school students that 67% admitted to have used chatbots such as ChatGPT to write essays or do work for them (RM Technology 2023). In their own methodology, the British Department for Education states that as this field is rapidly evolving, the figures and findings of this report may become out of date quickly.

The data shown here on pupils' use of GenAI have several limitations, which highlight the lack of data in this area in the larger field of GenAI. One such limitation is that none of this data says something specifically about pupils' use of GenAI within the English subject. Another limitation is that little research have been done with children and GenAI, and especially within a school setting; while surveys of university students are readily available. This is, however,

the nature of exploratory research, and while not directly comparable, the available data serves as reference to the use of GenAI among other age groups, and other nationalities.

2 Research method

The following chapter presents the research method used in this research project. The chapter is divided into three parts according to data collection, data analysis, and research reliability and validity. Each part refers to the methodological considerations that were taken in the planning and execution of this research project, aiming to bring transparency to the study as a whole.

2.1 Data collection

2.1.1 Sampling

The process of choosing participants for the study, called sampling, can mainly be split into two different types: probability sampling and non-probability sampling (Cohen et al., 2018, p. 214). For this study, a non-probability sample was used, meaning that I as a researcher have chosen to exclude parts of the wider population, not giving them the same probability to participate as the ones included in the sample. This was done with the intention of keeping the group of participants at a reasonable size, as the wider population in the context of the research question would include all pupils with English as a subject. As Cohen (2018, p. 217) states, non-probability samples are far less complicated to set up, and can prove perfectly adequate where researches do not seek to generalize their findings beyond the sample in question. There are several ways of conducting non-probability sampling, and for this project, a hybrid of convenience sampling, purposive sampling and sequential mixed methods sampling was chosen.

Convenience sampling involves choosing the respondents based on availability and accessibility at the time until the required sample size has been obtained(Cohen et al., 2018, p. 218). As the participating school was chosen in part because a fellow student worked at the school, and subsequently put me in contact with the English department at the school, I would characterize the initial sampling as convenience sampling. This entails that this sample does not represent any group apart from itself, and any generalizability to a wider population in this type of sampling is negligible (Cohen et al., 2018, p. 218). As group interviews were to be held, this dictated that smaller samples would have to be formed within the larger sample. For this, purposive sampling was used.

Purposive sampling involves handpicking cases to be included in the sample on the basis of the researchers judgement of their typicality or possession of the particular characteristic(s) being sought, and is assembled to meet the researchers specific needs (Cohen et al., 2018, p. 218). As I did not have any knowledge of the individual pupils participating in the study, I received assistance from the English teacher for each class in facilitating that the groups interviewed were representative in regards to their achievement level in English.

In sequential mixed methods sampling, it is typical that the results from the first strand inform and influence the methodology (e.g., sample, instrumentation) used in the second strand (Tashakkori et al., 2020, p. 189). In the instance of this research project, this was done by conducting group interviews with parts of the larger group of pupils participating in the study, as the interviews were thought to expose issues that could be further explored in the questionnaire.

2.1.2 Mixed method

As the research question for this study is twofold, asking both how pupils are using GenAI in the English subject, as well as how they believe it should be used, a research method that could provide insight to both facets was necessary. Whereas a pure quantitative method design would have been suitable to answer how pupils are using GenAI, the same method would not be suitable to provide insight on the pupils' thoughts on how GenAI should be used in the English subject; for this, a qualitative aspect was needed in the method design. A mixed method approach was therefore decided upon, wherein group interviews represented the qualitative aspect and a questionnaire represented the quantitative aspect. As all the research was to be conducted by a single person over time, not simultaneously, an exploratory sequential design that took advantage of this linear approach was decided upon.

An exploratory sequential design in its simplest form would seek to collect qualitative data first (usually with a small sample), with quantitative data from a larger sample used to generalize the findings (Cohen et al., 2018, p. 39). For this study, a version of this is done, but in addition to generalising the findings from the qualitative, the quantitative seeks to answer parts of the research question on its own. There are also other facets of the study that take advantage of the sequential nature of this design. Since the group interviews were conducted before the

questionnaire, the discussions in the interviews might have illuminated conditions related to the pupils' use of GenAI that were previously unknown to the researcher, which would later be incorporated into the questionnaire.

Figure 1 - Exploratory design - as shown in Cohen et al., 2018

2.1.3 Group interview

"Hochschild (2009) notes that the interview can do what surveys cannot, which is to explore issues in depth(…) They can be used to cast further explanatory insight into survey data, or indeed to set up a survey."(Cohen et al., 2018, p. 508). Alas, the interviews in this research project were used to inform the survey questions and explain the survey data, in addition to exploring an issue in depth. The issue that was meant to be explored in the interviews was the third operationalization of the research question, namely "In what manner do pupils believe they should be able to use GenAI in the English subject?"

The participants for this research project were 39 pupils from years 7-10 at an age-mixed school in Tromsø. From this pool of participants, three smaller groups of six were selected to form the groups that were to be interviewed, all while considering their level of competence in English to ensure a mix of different levels in each group. This was done with the intention of making each group as heterogeneous as possible, so as to give a greater chance for opposing views to emerge and the discussion that follows. As far as it was possible, an equal mix of age and gender was also sought after for the same reasons, but as some of the possible participants were absent of the day of the interview, this was not possible for all three interviews.

There are several practical reasons for choosing group interviews over individual interviews, some of which were considered in the planning of the research project, while others became apparent during the course of the study. A practical reason for choosing the group interviews is time; as a single researcher, I reasoned in the planning stage that I could interview a much larger number of pupils by conducting group interviews than individual ones. A practical reason that only became apparent during the study was the flexibility group interviews gives in a school setting where pupils might be absent on the day of the interview for any number of reasons. By simply informing all participants in the study of the time at which the group interviews were planned to be conducted, and that they might be chosen to participate, I as a researcher could simply substitute any absent pupil with another, ensuring that the interviews were conducted in a timely fashion. This proved to be essential, as a winter storm prevented several of the pupils to attend school on the day the interviews were to be conducted.

The decision of conducting group interviews instead of individual interviews was not only made because of its practicality, as it was also influenced by the advantages the group interview as a method proposed. Vaughn et al (1996) lists several advantages that the focus group interview offers researchers over the individual interview:

> *1.synergism (when a wider bank of data emerges through the group interaction),*
>
> *2.snowballing (when the statements of one respondent initiate a chain reaction of additional comments),*
>
> *3.stimulation (when the group discussion generates excitement about a topic),*
>
> *4.security (when the group provides a comfort and encourages candid responses), and*
>
> *5.spontaneity (because participants are not required to answer every question, their responses are more spontaneous and genuine). (Vaughn et al., 1996)*

On the matter of group interviews Gleiss and Sæther states that conversing on a subject in such a setting creates new thoughts and the participants recollect previous experiences and get different associations (2021, p. 81).

In addition to the benefits that group interviews have over individual interviews in general, group interviews have several strengths that are particularly beneficial when interviewing children. Group interviews encourages interaction between the group, as opposed to direct questioning from an adult, and might be less intimidating for them than individual interviews (Greig and Taylor, 1999, cited in Cohen, 2018 p.529). Similarly, Houssart and Evens notes that "… interviewing a group of children together can equalize more the power differentials between interviewer and children" (2011, cited in Cohen p.529).

2.1.3.1 Interview guide

I decided to conduct semi-structured interviews, which Brinkmann and Kvale say "attempts to understand themes of the lived daily world from the subjects' own perspectives" (2018, p. 14). As the nature of the research question that was to be explored in the interviews was open, meaning that I as a researcher could not anticipate every theme that could surface in the interviews, some flexibility in the interview approach was needed. To that end, an interview guide was constructed. Conducting semi-structured interviews with an interview guide is a very flexible form of interviewing and consists of a more or less detailed overview of topics and questions with a loose structure (Bjørndal, 2017, p. 109, my translation). For me, this structure involved rigid main questions with suggestions for follow-up questions, which could be deviated from or added to during each interview.

The interview guide was constructed following the procedures outlined by Arksey and Knight (1999, p. 97) which entailed devising a wide range of questions, eliminating any that were unlikely to contribute towards answering the research question, piloting to see how well the remaining questions worked, and to revise accordingly. A draft consisting of seven questions was used in a pilot group interview with a group of fellow English teacher students. The choice of using fellow students was made as I did not have access to any pupils who could participate in a pilot interview at this time in the research project. The pilot interview helped to inform further revision of the language used in the interview guide, both for clarity, but also to ensure

that the questions elicited responses relevant to the research question. At a later stage in the revision process, the English teacher at the participating school was invited to give feedback on the wording of the questions, ensuring that it was appropriate for the ages of the participating pupils.

In devising the questions for the interview guide, I consulted the outlines given by Arksey and Knight (1999, pp. 93-95), which suggest that attention has to be given to several aspects of the wording and phrasing of the questions. Some of these aspects are; keeping the vocabulary simple, precise and unambiguous; avoiding making assumptions, as well as double-barreled questions; whether to use hypotheticals or speculative questions; and whether to ask or avoid personal questions. The aspects of hypotheticals and personal questions were especially relevant for this interview, and special care was taken to phrase the questions in a manner that did not require the interviewees to discuss their private use of GenAI. This was deemed necessary, as admitting to using GenAI might have been considered cheating by the other pupils in the group. In practice, this meant that the phrasing of the questions referred to pupils in general, as in "Do you think that using GenAI affects the learning of pupils who use it?", or were hypothetical as in "What do you think can be possible pros or cons of using GenAI in the English subject". A fictional case was also used as a base for some of the questions, giving the pupils in the group interview a fictional pupils' use of GenAI to discuss instead of their own. In addition to making sure the questions were worded appropriately to avoid such answers, an introduction was added to the interview guide, ensuring that I explicitly instructed the group to avoid naming names, discussing their possible use of GenAI, as well as outing others.

The nature of the questions in the interview guide was predominantly open, and the ones that might have elicited yes/no responses were followed up with further open probes to better assess what the interviewee based their answer upon. Cohen et al. (2018, p. 513) list several advantages to open-ended questions; e.g. flexibility, the ability to probe for more information or to clear up misunderstandings, encourage cooperation, and allowing the interviewer to make a truer assessment of what the respondent truly believes. The advantage most relevant to me is, however, that open-ended situations can produce unexpected or unanticipated answers which may suggest hitherto unthought-of relationships or hypotheses (Cohen et al., 2018, p. 513). As the nature of my research involves exploring a relatively unexplored subject, this possibility of

the respondents suggesting unthought-of relationships and hypotheses in the interviews was seen as a positive addition to the research, as it could inform the construction of the following questionnaire, as previously mentioned in Chapter 2.1.2. In addition to the aforementioned advantages of open-ended questions, Cohen et.al (2018, p. 530) list several reasons why open-ended questions are preferable when interviewing children, mainly that they avoid the single-answer type of responses, the answers are usually more accurate than answers to closed questions as they are respondent-driven responses, and they can better accommodate responses from children with limited linguistic and cognitive abilities.

2.1.3.2 Conducting the interviews

As a part of the planning process leading up to the group interviews, I visited the school to familiarize myself with the space, and to find a suitable spot to conduct the interviews. This process was based on Gleiss and Sæther's point that " When conducting an interview, the location may affect the interview outcome; a quiet, practical location for the informant, where he/she can relax, is preferential" (Gleiss & Sæther, 2021) The management at the school suggested one of their meeting rooms, which ultimately became the location for the interviews. The room was close to the pupil's classrooms, but was not usually used by pupils, so it was deemed unlikely that other pupils would disturb the interviews. The meeting table in the room was large enough for all participants to be seated comfortably. To ensure that the audio recordings of the interviews would be intelligible, I did a test recording of myself using the same equipment as in the interviews, switching between every seat around the table while speaking. Listening to the test recording, I concluded that the audio quality was good enough to clearly make out speech, and to differentiate between voices.

The group interviews were recorded using two separate methods; a smart phone recording via the app Nettskjema-diktafon, and a dedicated digital audio recorder. The app Nettskjema-diktafon is developed by the University of Oslo to enable users to record audio on smart phones, where the recordings are encrypted and sent directly to their online form service Nettskjema when the phone is online (University of Oslo, 2023b). The aforementioned test interview was sent to a test form that was created in Nettskjema, to familiarize myself with the entire process

of data collection before the actual group interviews. The dedicated audio recorder was used on the recommendation of the University of Oslo, who says that backup solutions, like dedicated audio recorders, are recommended when conducting interviews, as there are instances where the app might fail to deliver the recording to the online form service (University of Oslo, 2023b).

During the interviews, the interview guide was mostly followed in the order planned order of the questions. When the pupils in the interviews would want to answer an earlier question, they were encouraged to do so. While some of the follow up questions were planned beforehand, others were improvised based on the pupils responses. This was often done to assure that an answer was understood correctly, but also to make connections to what had been said earlier in the interview. This was done based on a statement from Arksey and Knight, which says that "The importance of mutual understanding of meaning cannot be over-emphasized. If meaning is not established, then the subsequent analysis stands to be flawed by unperceived misunderstandings or misinterpretations between the interviewer and the interviewee" (1999, p. 100). They suggest that one way the interviewer can improve the accuracy of their understanding of the interviewees' answers is to occasionally paraphrase or summarize what they understand from an interviewee's answer, and invite feedback.

The interviews were meant to be brief, so as to not induce fatigue in the pupils that participated. To ensure this, a timer was set on the computer where I also had the interview guide, and all interviews were aimed to last around 30 minutes. The length of the interviews ranged from 28 to 38 minutes.

2.1.4 Questionnaire

When planning the questionnaire, a version of the operationalizing process in Figure 2 was implemented. Where the five first operations were addressed in the creation of the research questions, the remaining operations were addressed specifically in the design of the questionnaire. When deciding on what kind of questions were required, I looked to the research questions, and created questions based on the themes of how often pupils use GenAI, what they use it for, and their level of achievement. At this point, I also included questions on attitudes toward GenAI, with the purpose of creating context to their frequency and nature of using

GenAI in the English subject. Earlier questionnaires on student's use of GenAI that I myself had responded to online was looked to for inspiration, both in terms of wording of the question, but also the scales used.

As mentioned in Chapter 2.1.2, the interviews were conducted before the questionnaire to give the interview data a chance to inform the creation of the questionnaire. This was only done to some degree, as a full analysis of the interview data was not complete before the questionnaire was conducted. As such, the preliminary findings were used to inform the part of the questionnaire regarding attitudes towards GenAI. The questionnaire was then piloted by a small group of fellow English students to ensure that the questions were understandable.

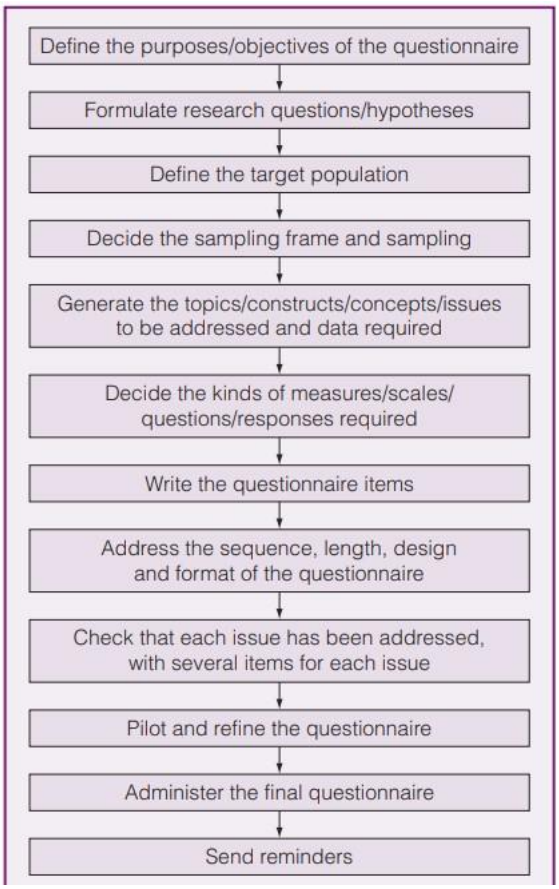

Figure 2- Stages in questionnaire design, as presented by Cohen et al. (2018)

Before conducting the questionnaire, I went through the same process as with the interviews to find a suitable location. An art classroom was decided upon, as it would be spacious enough for all 39 pupils to respond privately with their laptops at the same time. On the day of the questionnaire, the teachers at the school assisted me in sending all pupils that had given me their consent forms to the classroom, so that I could administer the questionnaire in a timely fashion. When all pupils were seated, attendance were taken based on the names on the consent forms to ensure that every pupil had given their consent. The pupils were subsequently given information on the questionnaires privacy, and were invited to ask for help at any time while filling out the questionnaire. The pupils were given an ample timeframe of 45 minutes, which proved to be excessive as the pupil who used the most time completed the survey in under 10 minutes. They were asked to show the completion screen of the questionnaire when they had finished, and were sent back to their classes on completion.

2.2 Data analysis

2.2.1 Quantitative analysis

The quantitative data collected in the survey part of this research project are used to answer research questions one and two, namely: How much, and to what purpose are pupils using GenAI in the English subject, and does the use differ between pupils with different levels of achievement in the subject. As the answers to these questions aim to describe the pupils' use of GenAI in the English subject, a univariate analysis was conducted. "Univariate analysis examines differences among cases within one variable" (Cohen et al., 2018, p. 730), which in my case meant looking at the number of respondents for each possible answer to the relevant questions, and report the numbers. Descriptive statistics are employed to report the collected data. Descriptive statistics make no inferences or predictions about population parameters, and simply report what has been found in a variety of ways (Cohen et al., 2018, p. 727). Regarding presentation of the data, Cohen et al. gives some guides on usage that became the basis of my decision of using bar graphs, namely that bar graphs are useful for presenting discrete data, and that they can be more accessible and comprehensible for some readers (Cohen et al., 2018, p. 754)

The quantitative analysis of the interview data was conducted by downloading the interview data from Nettskjema as a .xlsx file, where all further data handling was done in MS Excel. Not all questions in the survey proved necessary to answer the relevant research questions, and so the data from the part of the survey regarding attitudes towards GenAI was not utilized further in the quantitative analysis.

In the process of sorting the data and making graphs, the response time for each respondent was noted, and an outlier was found in the dataset. The outlier was a respondent who had used 18 seconds to complete the survey, compared to an average response time of 167 seconds. I chose to follow the advice of Maniaci and Rogge, who recommends screening the dataset for inattentive respondents (2014). They give an example of how this can be done in relation to response time, namely excluding participants who completed the survey in less than half the average time. Half the average time of all respondents here would be 84 seconds. They also point out that this is something that needs to be adjusted or optimized for any given study, which is what I did next. A benchmark was set by me, as to how fast I could possibly complete the survey, while also reading every question before I answered. I timed myself to 25 seconds and decided to remove any respondents with a shorter response time, which only included the outlier of 18 seconds. When inspecting the answers from this respondent, they had answered "No" or "disagree" to every question, or rather the leftmost alternative to every question.

2.2.2 Qualitative analysis

When the group interviews were completed, and the recordings were successfully uploaded to Nettskjema, the recordings were automatically transcribed by an integrated solution in Nettskjema called Autotekst. Autotekst is a tool developed by the University of Oslo for transcribing speech to text using OpenAIs automatic speech recognition system called Whisper (University of Oslo, 2023a). The output texts from Autotekst were downloaded and reviewed to ensure that they were as verbatim as possible. This was done by listening to the audio recordings while reading through the transcriptions, fixing any mistakes that the automatic transcription tool had made. The review process was conducted shortly after the interviews, ensuring that my recollection of events was as accurate as possible. A part of this process

entailed segmenting the texts by speaker, where each participating pupil was assigned a colour, and all text from each pupil was given these colours. This was done by discerning each voice in the audio recording, assigning colours, and repeatedly listening to the different "colours" to ensure that they only included statements from one pupil. The process of identifying each speaker was done following guidelines stated by Guest et al. who says "When analysing focus group transcripts, we want to know whether a few voices dominate the discussion or whether a given perspective is widely shared" (2012, p. 64). The colour codes made it possible to see later in the coding process which pupil said what, ensuring that if one pupil was quoted repeatedly, it was not mistaken as being several pupils' opinions.

After the interview transcriptions were reviewed, the process of coding and analysing the material ensued. Applied thematic analysis was chosen as the main approach to this work, which is defined in the following manner:

> *Applied thematic analysis as we define it comprises a bit of everything—grounded theory, positivism, interpretivism, and phenomenology—synthesized into one methodological framework. The approach borrows what we feel are the more useful techniques from each theoretical and methodological camp and adapts them to an applied research context. In such a context, we assume that ensuring the credibility of findings to an external audience is paramount, and, based on our experience, achieving this goal is facilitated by systematicity and visibility of methods and procedures. (Guest et al., 2012, p. 15)*

This approach is further described in the following section on analysing the interview data.

While the preliminary review of the interview data was done by downloading the text files from Nettskjema and using MS Word to edit them, a dedicated qualitative data analysing software (QDAS) was chosen for the subsequent coding process. As stated by Guest et al., "QDAS facilitates this process, enhancing the efficiency, consistency, and comprehensiveness of thematic analysis. These benefits apply equally to qualitative data management, coding, exploration, reduction, and reporting(…)" (2012, p. 220). I chose to use NVivo 14, as I could acquire the license for free through my university.

The first step in the coding process was to create structural codes based on each interview question. Guest et al. describes structural coding as a way "used to identify the structure imposed on a qualitative data set by the research questions and design" (2012, p. 55). By doing this, the interview texts were segmented based on the questions asked, with their accompanying responses. Although each interview followed the same interview guide, the participants could sometimes return to a topic that was covered earlier, or spontaneously answer a question that would come later in the interview. Guest et al. points out that "(…) one generally hopes for these kinds of moments in a qualitative interview or focus group as it is an indication that one is uncovering rich data (or thick descriptions)" (2012, p. 61). In such instances, the text was coded into a structural code based on its emergent content, without consideration for when in the interview the instance occurred. This was done because the structural codes were only meant to be an aid in the following process of creating thematic codes.

To find thematic codes in an applied thematic analysis strategy, Guest et al. states that "The first step (…) is to refresh your understanding of the analytic objectives. They will always frame how the text is viewed and ultimately determine which themes are worth the effort of tagging, defining, and coding"(2012, p. 65). The analytic objective here was to answer the research question "In what manner do pupils believe they should be able to use GenAI in the English subject?", and so the data was reread to find themes regarding different manners of use of GenAI. This work followed an iterative approach outlined in Guest et al. (2012, p. 70), were the texts were read, themes were proposed, and the themes were refined into codes. As a part of this process samples of the text was coded twice with a week between coding, which then were compared to adjust the codes if necessary. The purpose of this process is "To avoid conflation of what people say with our interpretation of what they said" (Guest et al., 2012, p. 70). To further check for clarity in the creation of the thematic codes, a group of fellow students were invited to read a sample of the text and propose codes during a seminar. Their suggestions were similar to the codes I had already assigned the same segments.

The process of developing and refining a codebook necessitated that much of the text had to be reread and recoded several times as new codes and code definitions emerged. To aid with this, a coding log was created, ensuring that when creating a new code or definition, or when throwing out codes that did not work, all relevant text were coded with the new conditions in

mind. When the codebook was considered final, all interviews were reviewed while using the "code panel" function in NVivo to ensure that all relevant parts of the interview were assigned a code.

2.3 Research quality

2.3.1 Reliability and Validity

As stated by Gleiss and Sæther, the researcher is responsible for assessing and reflecting on the quality of their own research, which is usually done by evaluating the work according to the two terms reliability and validity (2021, p. 201). The following chapter will present considerations related to this research project's reliability and validity in the context of both the qualitative and quantitative research that has been conducted.

Reliability concerns the quality of a research process and whether the research is to be trusted (Gleiss & Sæther, 2021, p. 202). As the sole researcher conducting this research project, the project in its entirety is subject to my own subjectivity, meaning that my own views and possible biases will have affected the decision-making process which led to the questions asked in the interviews and questionnaire. One factor that helped remedy this in the development of the questionnaire was that there were previously conducted questionnaires that explored pupils use of GenAI that I could compare my own to. The effect my person had on the interviews is also a question of reliability, as the manner in which I presented the questions could have an effect on the pupils' answers. One example of this is my use of active listening to encourage participants to contribute to the conversation, which the pupils might interpret as agreeing with their views.

Validity is defined by Gleiss and Sæther as "the quality of the data and the researcher's interpretations and conclusions. Validity therefore concerns itself on how well the different parts of the research design are connected to another" (2021, p. 204, my translation). As a part of this research project, I have used different methods of data collection based on the different scopes outlined by each research question, trying to apply the strengths of each method. This included conducting a questionnaire to collect data on the pupils' previous use of GenAI, as this is information they could be hesitant to give in front of their peers. The qualitative data from the group interviews was also used to inform the questionnaire.

Regarding the presentation of the findings from the thematic analysis of the interview data, Guest et al. state that "[u]sing verbatim quotes increases the validity of findings by directly connecting the researcher's interpretations with what the participants actually said" (Guest et al., 2012, p. 85). Several measures were taken during the analysis process to ensure that this would be possible, including that the automatically generated transcripts from the interviews were reviewed and corrected to more accurately depict what was said in the interviews. This being said, the selection of quotes to be used in the presentation of the findings could be subject to critique, as well as the order in which they are presented.

2.3.2 Ethical considerations

When conducting research, there are always several ethical aspects that must be taken into account throughout the process. The National Committee for Research Ethics in the Social Sciences and the Humanities (2022) gives several guidelines concerning ethical obligations in research, many of which were considered applicable to my research project and were taken into consideration when planning and conducting my research project. Several of the guidelines relate to consent, stating that ethical consent to participate should be voluntary, informed, unambiguous and preferably documentable, and that to protect children in research, consent must be obtained from both the parents and the children themselves. Others relate to privacy and data protection, as well as avoiding causing harm and unreasonable disadvantage for the participants as a consequence of their participation.

To conform with the ethical guidelines and to ensure that the research project met the legal data protection requirements, a Notification Form was submitted to Sikt – the Norwegian Agency for Shared Services in Education and Research. Based on their templates, I created a consent form to be filled out by the participants and their parents and submitted it for assessment, along with copies of the interview guide and questionnaire. The Notification Form described how I intended to receive consent, how the data would be collected, how the data would be secured, and how anonymity would be guaranteed. The submission was approved by Sikt, with the advice that I reminded all interviewees not to mention teachers, parents or other pupils with regard to names or personal information. As a result, this was included as part of the interview guide. The interview guide, Notification Form and consent form can be found as appendixes to this thesis.

3 Findings

3.1 How are pupils using GenAI?

The findings presented in this chapter are the result of a univariate analysis of data gathered in the questionnaire, as described in Chapter 2.2.1. The pupils were asked two direct questions: “How often do you use GenAI in the English subject” and “What have you used GenAI for in the English subject?”. The latter question was only directed to the respondents who submitted in the former question that they had used GenAI to some degree in the English subject, i.e. the ones that answered “I have never used GenAI” to the former question did not get to answer the latter.

3.1.1 Frequency

To understand how pupils are using GenAI in the English subject, we first examine how often they have used it. The chart shown in Figure 3 shows the results from the question “How often do you use GenAI for tasks in the English subject?”. From reading this chart, we can see that the respondents who answered that they have used GenAI in the English subject are the majority, with 21 of the 38 respondents answering that they have used GenAI with some frequency in the subject, or 55%, as opposed to the 17 respondents who say that have never used GenAI in the subject, or 45%.

It is notable that within the group of 21 respondents that have used GenAI, 14 respondents have only used it a few times, while 7 respondents say they use it regularly; 3 of which at least monthly, and 4 of which at least weekly. This means that combined, the regular users, i.e. at least monthly, make up 19% of the total number of informants.

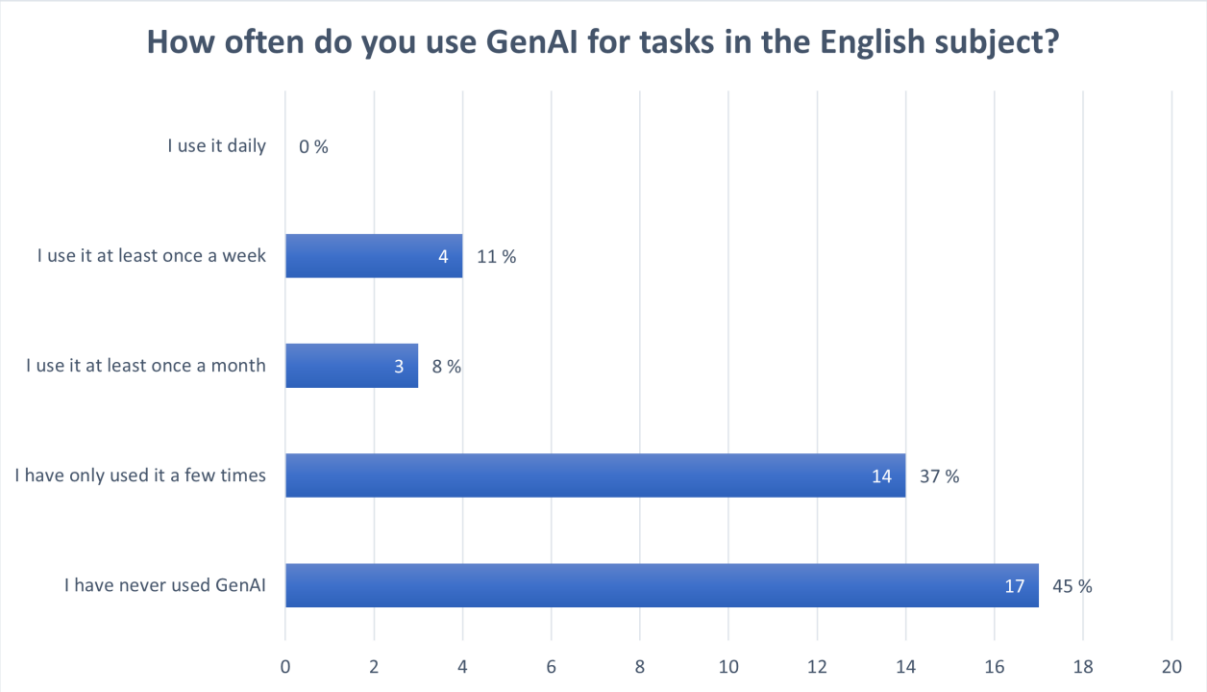

Figure 3- Frequency of using GenAI in the English subject

3.1.2 Types of use

The graph in Figure 4 shows the answer to the question "What have you used GenAI for in the English subject". As mentioned in the previous chapter, this question is only answered by the 21 respondents who previously stated that they have used GenAI in the English subject. The respondents had the option of selecting several alternatives for this question, which is why the number of responses exceeds the number of respondents. This is also the reason why the percentages shown beside each bar in the graph do not add up to one hundred per cent.

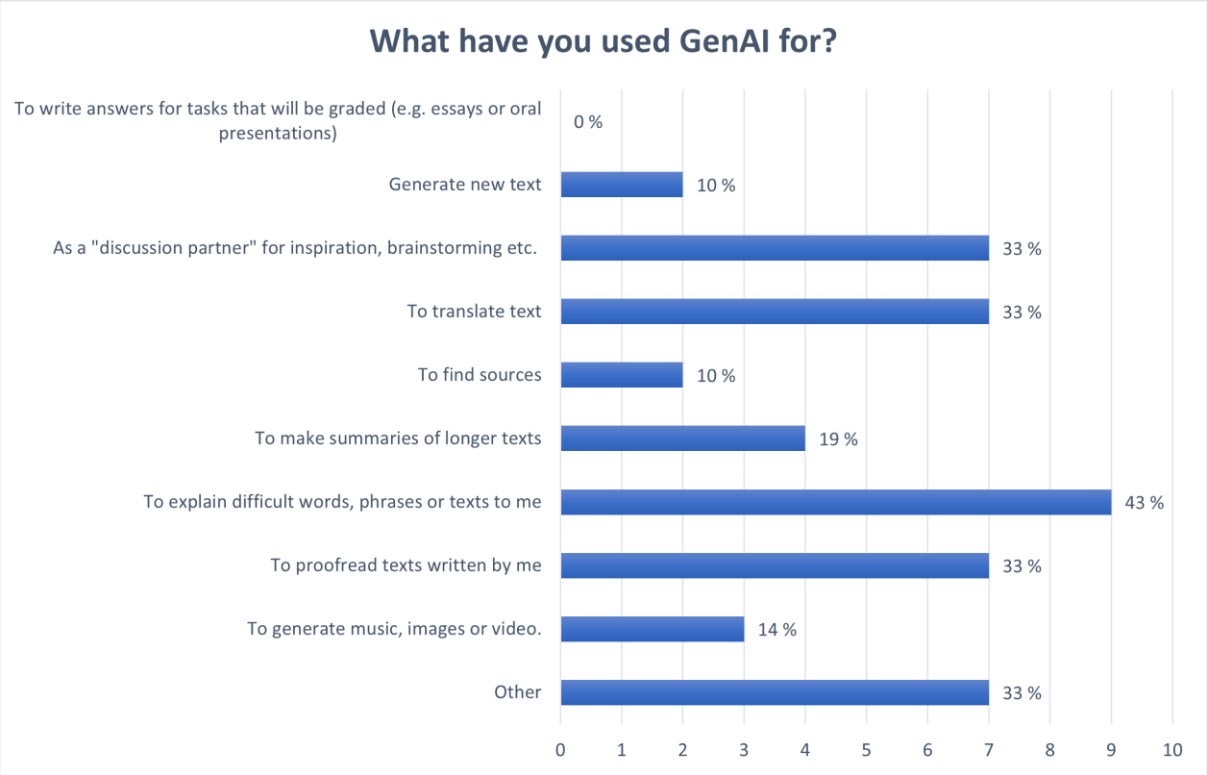

Figure 4 - Types of GenAI use

The data shown in Figure 4 shows that the most common use case for GenAI in the English subject among the respondents is “To explain difficult words, phrases or text to me”, which is answered by 9 out of 21 of the respondents to this question, or 43%. This is followed by three options with 7 counted responses each, or 33%, namely “ As a “discussion partner” for inspiration, brainstorming etc.; “To translate text”; and “To proofread texts written by me”. Four respondents answered that they have used GenAI to make summaries of longer texts, while three respondents answered that they have used it to generate music, images or video. Two respondents answered that they have used GenAI to generate new text, and two respondents answered that they have used it to find sources.

Figure 4 also shows a finding related to what the pupils have not used GenAI for; the only alternative that received a count of zero was “To write answers for tasks that will be graded (e.g. essays or oral presentations)”. This finding, that not one of the respondents answers that they have used GenAI for graded work, is given context in the findings from the interviews in

Chapter 3.4.2. Figure 4 also shows that seven respondents chose the alternative “other”, but they were given no option to elaborate on this.

3.2 Use of generative AI across different levels of achievement

The findings in this chapter are meant to answer research question 2, namely “In what ways, if any, does the use of GenAI differ between pupils of different levels of achievement in the subject?”. The participants had the option to choose more than one option for the question “What grade do you expect to get this semester in English?”. In the instances where more than one grade was chosen, an average was made for that individual. If the average was between two grades, the lower grade was used to represent the participant. The results were consolidated into two groups, namely “Grade 3 and 4” and “Grade 5 and 6”, with seven and fourteen participants in each group respectively.

3.2.1 Frequency of use

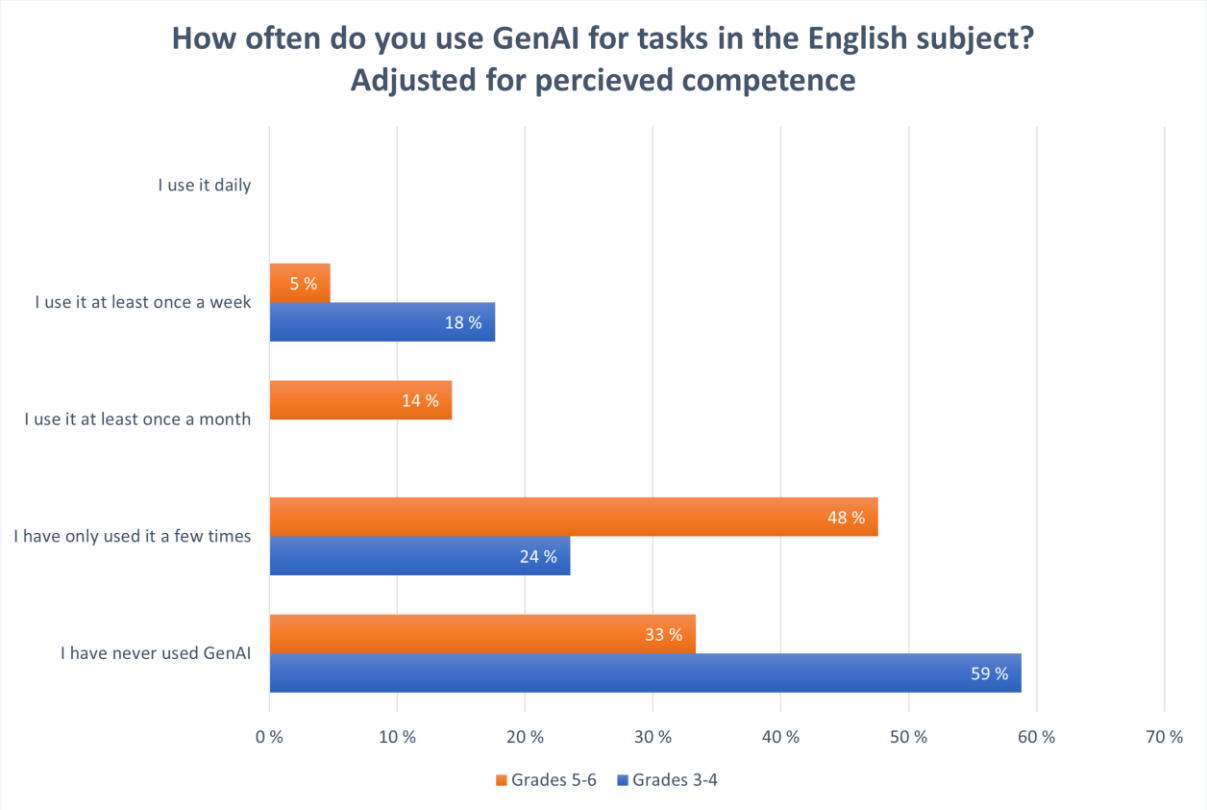

Figure 5 - Frequency of use, adjusted for grade

Figure 5 shows answers to the same question as Figure 3, but with the responses shown for the two aforementioned groups based on their expected grade in the English subject. This figure shows that within the group that answered that they expect to receive grades 3 and 4, 59% answered that they have never used GenAI in the English subject. This is nearly twice the percentage of the group that answered that they expect to receive grades 5 and 6, where 33% answered that they have never used GenAI in the English subject. Where 24% of respondents in the group expecting grades 3 and 4 say they have only used GenAI a few times, 48% of respondents in the group expecting grades 5 and 6 say the same. Grades 3 and 4 show a higher percentage in the "I use it at least once a week" column, with 18% compared to 5% for grades 5 and 6, but 14% of respondents expecting grades 5 and 6 have chosen the alternative "I use it at least once a month" whereas 0% of respondents expecting grades 3 and 4 did the same.

3.2.2 Type of use

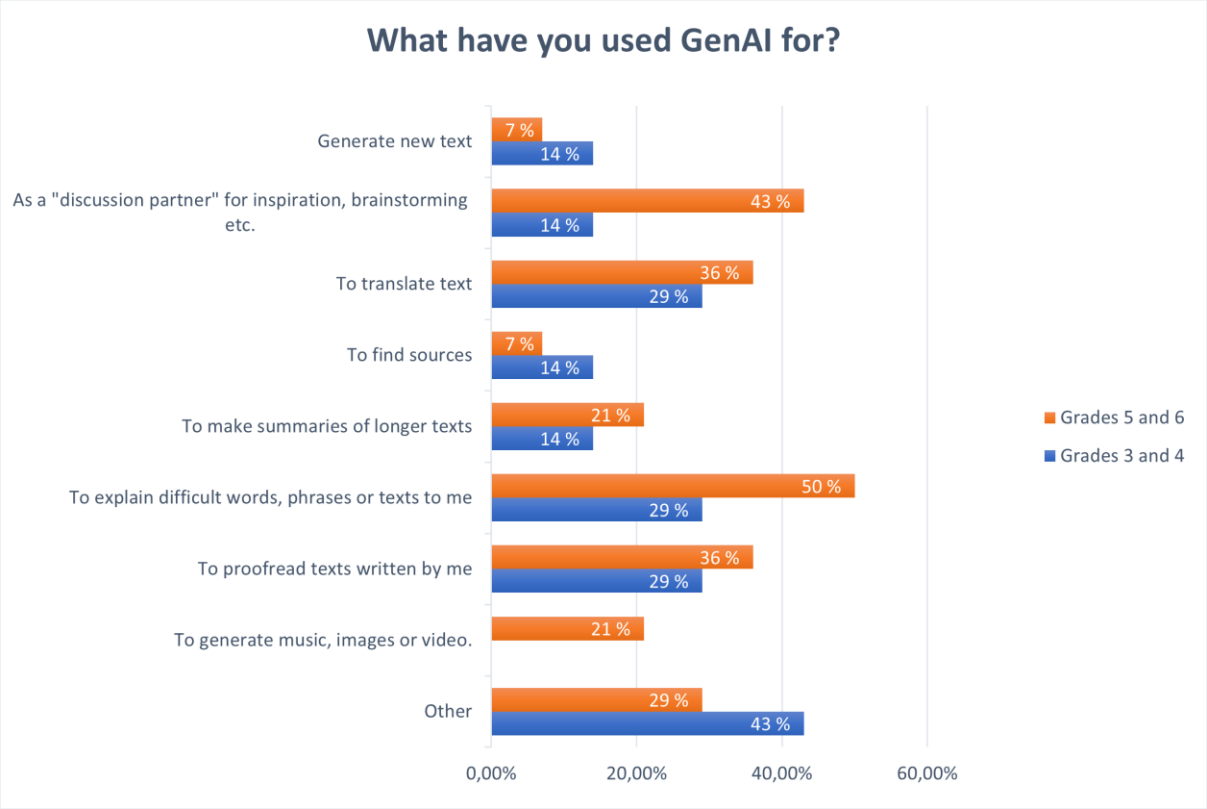

Figure 6 - Type of use, adjusted for grades

Figure 6 shows answers to the same question as Figure 3, but divided into the two grade groups. The respondents could as mentioned earlier choose several answers to this question, so the

numbers shown in the bars indicate the percentage from each group that included each alternative among their answers.

The largest disparity between the two groups is in their use of GenAI "As a "discussion partner" for inspiration, brainstorming etc." which was checked off by 43% of grades 5 and 6, and by 14% of grades 3 and 4. The second-largest disparity is in their use of GenAI "To explain difficult words, phrases or text to me", where 50% of pupils expecting grades 5 and 6 chose this option, compared to 29% of pupils expecting grades 3 and 4 did the same. While the disparities in other use cases are of lesser consequence, it is notable that pupils expecting grades 5 and 6 are represented by a higher percentage in six of the nine use cases that are presented.

3.3 Wanted use

The findings presented in this chapter are the result of a thematic analysis of the interview data, as described in Chapter 2.2.2. Based on this analysis, this chapter outlines seven reoccurring themes regarding how pupils want to be able to use GenAI in the English subject. Each theme will be presented in descending order, based on the number of occurrences of their respective codes in the interview data. After presenting each theme, examples from the three interviews will be used to show each group's thoughts on the respective theme.

3.3.1 Inspiration

Wanting to use GenAI for inspiration was a common theme in the conducted interviews. This was often proposed by the participants as a limit for what the pupils themselves should be able to use GenAI for, both in their homework and their essays. When questioned about possible rules for using GenAI in homework, one pupil in the interviews suggested the following:

> *I think maybe you can say that you have used AI so the teachers know (…) you have been writing it, but you got some ideas from ChatGPT.(…) I don't think you can copy-paste it directly. I would say if you get the idea from it, then it's fine.*

This sentiment of using GenAI to get ideas was echoed in another interview when the participants were asked the same question about possible rules for GenAI in their homework.

I think you could maybe just use ChatGPT for homework if you're stuck on ideas, (or) if you're wondering what kind of paragraphs to write. You just sort of like, oh, what would you say is a good paragraph starter for blah, blah, blah. Just like, helping you start going and not to use it to actually do all your homework.

Adding to this idea of using GenAI as inspiration in the form of ideas or paragraph starters, another pupil responded to the previous quote. They brought up the effect they thought GenAI could have on learning if they could use GenAI for these purposes. They said the following:

I think it's a great tool to help you to enhance your own learning from your homework instead of making the learning for you. So like creating guidelines or what she said.

During the interviews, the topic of using GenAI as inspiration for essays was also discussed in several instances, being specifically separated from topics such as homework assignments or tests. The recurring response when discussing essays in the interviews were that using GenAI for inspiration in the writing process of longer essays should be allowed, but not for larger parts of the essay, or aspects of the essay that they are graded on. When the participants in interview 2 were asked for an example of how the rules for using GenAI for homework and tests should be different, the following dialogue took place between the interviewer and two pupils, noted as I, P1 and P2 respectively.

I:You already said it, you shouldn't be able to have it write…

P1:No, because that's effectively plagiarism and cheating. But having it give you ideas and stuff like that, that's like, because when you're writing an essay, it's very difficult for some people to come up with ideas. So having that, using it as a tool to give you ideas, I don't think that's, I think that's okay.

P2:I think for written tests, it shouldn't be in use like whatsoever. But if it was for multiple page long essays, then you can use it for inspiration because written tests are like, I mean like written tests by hand. Yeah. For those, those are more intended to like strike your own creativity. But essays in particular, they're, you're meant to

like cite your sources and stuff. And if you were to get inspiration from generational AI, then I don't see a problem with it.

Wanting to be able to use GenAI for inspiration in tasks that are not necessarily homework or was also stated in another interview, where another participant mentions using GenAI for inspiration in doing creative tasks.

Yeah, I agree. I think we should be able to use it for generating ideas, especially in creative tasks when you don't know where to even start. It's really good for making design specifications or for making sentence starters and those kind of tasks. It can be really helpful to just get you started.

3.3.2 Explanation

The pupils who participated in the group interviews expressed several ways in which they believed GenAI could be used to explain both difficult task instructions and subject matter. On the matter of difficult task instructions, one pupil said the following:

And I think it could also, if you don't understand the task itself, if you get Gen AI to like explain it to you, like lay it out for you simply, it would benefit you.

When asked about the pros and cons of using GenAI, one pupil brought up GenAI's ability to explain in several different ways.

It's also good for explaining. Like, if it explains, you don't understand. You can ask it to explain it in another way.

And it's like, it has a lot of like, trained on a lot of data. So, probably, it can take very many different aspects. In which it can describe what you're trying to understand. If you're trying to understand like, what structure you need for your writing.

Another pupil had a similar thought when asked whether they believed there should be different rules for using GenAI for homework and tests. This pupil, however, explicitly speaks to using GenAI to explain difficult tasks in a homework situation:

I agree. For homework or tests, you might not understand the task. And no, for homework, you may not understand the task so you can get the AI to help you.

Another pupil made the same distinction between using GenAI to explain homework tasks as opposed to tests, saying this in agreement with earlier sentiments about using GenAI for help, but not copying from it:

I agree with you. You can use it if you don't understand the task. But when you're going to write a text or test, you should not use it.

Several of the pupils in the group interviews also highlighted GenAI's ability to explain difficult subject matter in a more accessible manner. The following quote is from an answer to the question "Do you think that using GenAI affects the learning of those who use it?":

Yes. The thing is it can be both positive and negative. If you just use it to replace your own answers, of course you don't get anything off that.But if you use it as it was meant to help you, like for example, taking a text that's written in a really difficult manner and recycling it through an AI to make it easier to understand, that can help you grasp some more difficult subjects through something else.

Another pupil answered the following to the same question:

P: Well, in the end it will affect you in the way that you decide to use it. If you decide to use it in a way that you're just using it to get something quick or do something like... Then you're not really learning, are you? You just read through the thing, replace some words and you got a text that is somewhat alright, but you don't really know anything about the subject, because you didn't write it.

I: Yeah.

I: But if you decide to take a text from Wikipedia, where sometimes they're quite difficult to read, you put it in there and you tell it to make it easier for you to understand, and it goes in there, changes some words to some easier words, shortens the text, keeps it relevant, and then you learn something.

3.3.3 Help

A category of use that featured heavily in the group interview responses was help, where several respondents mentioned that they want to be able to ask GenAI for help in various situations. The pupils generally made a distinction between asking GenAI for help and copying answers directly from a GenAI tool, believing they should be allowed to use GenAI for help and guidance in their work. Two pupils said the following regarding asking GenAI for help with their homework:

> *P1: I think it shouldn't be used to get you the answer, but help you, give you a way to get the answer. If you don't understand how to do something, it will show you how to do it.*
>
> *P2: Yeah, I agree with that.*

Another pupil used an example from mathematics to explain how GenAI can be used to help pupils in a way that search engines cannot, and ended their train of thought like this:

> *But on other tasks, for instance in math, maybe not in getting the answer, but it can be really helpful to show you a problem and then step by step how to solve it. Because in Google, you can't really give Google a question and tell me how to solve it. But for algebra, for instance, if you put it in ChatGPT, it can show you step by step how can you solve this question instead of just solving it. Instead of trying to scour the internet for a way to solve a question.*

The quote above serves to exemplify a view among several of the pupils interviewed, namely that using GenAI to get the steps needed to solve a problem, or examples, should be allowed.

3.3.4 Information

Some of the pupils participating in the group interviews proposed that they should be able to use GenAI to get information but with several suggested limitations. One such limitation is plagiarism, which was mentioned in all three interviews as an example of unacceptable practice. The following excerpt is in response to whether GenAI should be allowed to do homework:

P1: Yeah, I think you should be able to get information from ChatGPT, but you should not be allowed to copy-paste it into your document or something.

I: You agree?

P2: Yeah.

This view is further shown in the following excerpt, where a pupil resonates about a case they were given, in which a fictitious pupil, Ola, is given the task of writing an essay about the characters of a book. In this case, Ola copies text from ChatGPT which he then edits a bit. This case was discussed in light of several possible situations; if the teacher did not mention the use of GenAI, if the teacher prohibited the use of GenAI, and if Ola copied directly from ChatGPT without editing.

P: Yeah, you should write it yourself, but you should be able to use ChatGPT to ask it questions about a fantasy character. Or for more information, or words you can use.

I: But you should start writing yourself?

P: Yeah, you should write it yourself. And then you can ask.

I: Or Ola should do this.

P: Yeah, then Ola can ask for questions about the character, for fun facts.

Here, the pupil expresses a view that indicates that GenAI should be used to supplement the information in a writing process but not be the base of the writing, which the pupils should do themselves. Another pupil had an opposite view on the case, in that the information could be used as a base for writing, but also brings up some important caveats:

I think if he used it, like, to write paragraphs and then edit it. Or, he should at least go check if the information is right. ***Yeah, and check it.*** *Yeah, checking and... And, yeah, ask the teacher before you do it to get the information from ChatGPT.* ***Yeah****.*

Before you're just doing it. ***I see****. And in the last task or question, I think he shouldn't use it. If the teacher told him to not, then you shouldn't use it at all.*

The pupil in the excerpt above points out the importance of verifying the information one can get from GenAI, not trusting it blindly. This sense of academic integrity is further shown in that they believe one should ask the teacher before using GenAI to find information, even though the teacher had not mentioned the use of GenAI in this task.

Using GenAI to find information in the form of researching a topic was also mentioned in the interviews, where one pupil said the following when asked to reiterate what they want to be able to use GenAI for:

I: So you already said that you want to be able to*…*

P1: …use it, but for finding out, doing research on stuff, getting information, not just copy and pasting.

Research did not come up explicitly as a wanted use case in other places in the interviews, which can have its natural explanation in the following excerpt, in which another pupil talks about the change they have experienced in the nature of their writing tasks in the English subject after GenAI became available.

We still have assignments that are digital, but the rules are a lot more stricter, and our teachers are focusing more on tasks that require knowledge that you've acquired in class or from a textbook. The amount of essays we've written from research have gone down drastically. We don't do much research anymore.

3.3.5 Improving language

A suggested use case discussed in the group interviews was the use of GenAI to improve the language of texts written by the pupils. This discussion had several facets, one of which concerned how GenAI could help pupils better convey meaning in pupil-written texts. One of the pupils in the group interviews said the following when asked about possible pros and cons of using GenAI in the English subject:

P: Pros, it can definitely affect your language. Or pros and cons. Especially if you're trying to search for a very specific word and you know that it exists, but you don't know the specific word. It's really like just putting a little description of it into AI, it's really easy to get it back. And like, oh, that's like it can help you tell better what you're trying to say. But, yes.

I: ***So that's a pro.***

P: Kind of pro or con if it's like if you go too overboard and have no clue what you're saying, but you, you know, use too fancy words for yourself.

Here, the pupil describes a way of using GenAI to find more suitable words for a given situation. Another pupil in the interview elaborated on the same idea, but concerning rewriting entire sentences:

It can also like help to make like if you have like a sentence, I think it's just too like not, it's not good, like, worded like really badly, then you could make it inserted in like GenAI and it can make it more concise and more understandable for the readers, which is pretty helpful in a lot of circumstances.

In this instance, the context of the discussion is important to take into account, as the pupils are discussing *possible* pros and cons of using GenAI in the English subject. The pupil in the excerpt above does specify that using GenAI to rewrite sentences "is helpful in a lot of circumstances", which is to say, not in every circumstance. The same pupils showed awareness of possible negative aspects of relying too much on GenAI in their writing, which can be seen in the following excerpt, where another pupil in the group said the following when asked about possible cons of doing this:

Yeah, because then you don't know like how to make it more concise yourself. If you don't, like, if every sentence you write, you use chat GPT to make it as perfect and readable as possible, then you'll never know how to do that yourself.

The loss of the learning experience from rewriting and making texts more concise without GenAI are aspects of this type of use that were further discussed in the interviews. One pupil said the following about their experience of writing texts after GenAI became available:

> *[In essays] (…)the only thing it helps with, unless you use it to like write it for you, is to slightly speed up the process of polishing the text, I'd say. So just like tiny tweaks and make… you can cycle through and be like "hey, is there any grammatical mistakes?". That's something easy you can do instead of reading through it for the 50th time. We're asking somebody else to do that. But of course, it's still good to do it yourself before putting it through something.*

Here, the pupil describes using GenAI at a later stage in the writing process of an essay, where its purpose is to "polish" the text after the pupil has read through the text several times. The pupil points out that "it's still good to do it yourself before putting it through something", where "it" refers to polishing the text. What the pupil means by "good" is unclear, but from the context of the previous question, one can infer that this means that it ensures learning in the writing process.

The challenge of retaining learning in writing processes is also mentioned by another pupil who, when asked to sum up their thoughts on the subject, brings in aspects of monitoring and instruction to the discussion concerning the use of GenAI in the English subject:

> *I also feel like as long as it's monitored and at some point we're taught how to use it in a way that it's most efficient and most sustainable for us to have it by our teachers, so that those who haven't been taught how to use it or don't grasp the concept, we can say that, hey, you can use it to shorten your text or to just tighten up some grammatical mistakes instead of the most obvious "put idea in, get text out".*

The pupil in the excerpt above brings up several facets of the discussion on GenAI in the English subject, showing that they can be interconnected. While it may be unclear from the transcription, I understand the statement to say that without instruction in the use of GenAI, the

pupils who do not reflect on its impact on their learning, might not consider using GenAI in a manner that retains language learning, opting instead to use GenAI written text unscrupulously in their work. This point highlights that if instruction from the teachers in the use of GenAI is absent, some groups might be more affected than others, both in the long and short run, as they resort to using GenAI in an unsustainable fashion.

The final aspect discussed in the group interviews concerning the use of GenAI to improve language and grammar is the subject of shortening entire texts. This aspect is already mentioned in earlier findings regarding shortening difficult Wikipedia texts and in the previous excerpt regarding shortening pupil texts. One pupil, however, had explicit views on whether shortening one's texts with the use of GenAI should be considered cheating, if the text was written by the pupil. The following excerpt is from when a pupil in the group interview were asked to sum up their thoughts on how they want GenAI to be used in the English subject:

> *Honestly, I think it's just be used to maybe ask questions if you're wondering something or if you want to shorten a text. You should be able to use it to shorten it. Obviously, if it's already a text that you've written and you use it to shorten it, it's not really cheating or anything because it's your own text, but you're still getting help. It's basically like asking one of your peers to help you shorten a text. You just tell it to keep up with your wording and your spelling and just like, have it take the text and create a shorter version of it.*

3.4 Wanted restrictions

To better understand how pupils believe GenAI *should* be used in the English subject, this chapter describes the findings concerning the inverse, namely how pupils believe GenAI *should not* be used in the subject. As in the previous chapter, these findings are based on a thematic analysis of the interview data, described in Chapter 2.2.2. The themes will be presented in descending order based on the number of occurrences of their respective codes in the interview data, with examples from the interviews. These themes function as opposites of the themes

presented in Chapter 3.3, and many of the quotes used in Chapter 3.3 are also relevant to this chapter. These will not be presented again but rather referred back to if considered especially relevant. The themes that were found concerning pupils' wanted restrictions on GenAI related to copying, homework and tests.

3.4.1 Copying

The pupils that participated in the group interviews were joined in their beliefs that pupils should not be able to copy text directly from GenAI for tasks in the English subject. Their reasons behind this were many, and some have inadvertently been shown in other findings chapters in this thesis, as it was commonly used by the pupils in the interviews as contrary to how they believed they should be able to use GenAI in the English subject. One such reason was simply that copying text directly from GenAI is cheating, as exemplified by the pupils in the first quote used in chapter 3.3.4, who agree that they should be able to use GenAI to find information, but not to copy directly from.

The notion that copying text directly from GenAI is cheating came through in all three group interviews when they were given the fictitious case of Ola to discuss. Ola had used GenAI to complete an assignment, and the pupils were to discuss what they thought of Ola's use of GenAI. One of the questions the pupils were asked was what they thought of Ola's use of GenAI if he had copied directly from ChatGPT. One group answered the following:

> *I: How about if Ola did not edit anything from what he copied? If he just copied directly?*
>
> *P1: That's not allowed.*
>
> *P2: Because then it's cheating, then he didn't do the task.*
>
> *P1: Yeah, that's not allowed at all.*
>
> *(...)*

P3: I don't think it's okay, because he's taking another's work and pretending like it's his own. And that's... I don't think that's okay.

I:Yeah, I see you shaking your heads around the table. So I'm guessing...

P4:Yeah, I agree.

I: Yeah? You should not be allowed to just copy things.

P3: Mhhm.

I: Alright.

When the same scenario was presented to the second group, they said the following:

I: How about if Ola did not edit anything from what he copied from ChatGPT? Like he just copied it directly.

P1: Oh, that's not good. That's cheating.

I: That's cheating?

P2: If you get caught… Now it's cheating.

P3: Now it's cheating on both ways.

(…)

P1:I wasn't saying that the original one wasn't cheating, where he just like changed a bit. That's still…

P4: cheating. Alright. Alright.

I: But maybe just less cheating?

P4: Less cheating.

P1: Still cheating.

The pupils referred to as P1 and P3 in the excerpt above refer to the earlier question wherein Ola copies from ChatGPT, but edits the text a bit. When the third group was given the same scenario, they arrived at a similar conclusion, in that it does not matter if Ola changes the text just a little bit, or if he copies directly; it is cheating in both cases:

> *P1: I would say it does the exact same.*
>
> *P2: It's the exact same, I mean, it's the same thing.*
>
> *P1: It's the exact same, just like with the other version, it's just like some words are different. It's still the same exact consequences that it does.*
>
> *P3: I'd still call it cheating, but if you add, like, oh yeah, this person made it.*
>
> ***I: Yeah, like if you cited it, like put a reference?***
>
> *P2: I think even citing it, but it would still consider plagiarism if that's the only thing.*
>
> ***I: Yeah, if you cite your entire text.***
>
> *P1: It's just copied from ChatGPT.*
>
> *P4: It's just like copying a Wikipedia page, you know.*
>
> *P1: Also, ChatGPT isn't actually, like, you can't really cite ChatGPT.*

Another aspect of copying discussed in the interviews was the effect it would have on the learning outcome of the pupils. One pupil said the following about Ola's use of GenAI, which was echoed by several others in the same interview:

> *Well, I mean, he's not really doing the task, so if he uses ChatGPT to mostly do it, he might as well just not do the task. It's not useful to do the task. You're not learning anything. You're not getting anything out of it.*

In another interview, one pupil said the following about Ola's learning outcome:

Because he's not going to learn from that, you know, like the point was to analyze characters, he's not going to know how to analyze, now the teachers are going to have really high expectations for Ola, but, you know, because he did the whole thing with ChatGPT, he's never going to know how to do it if he doesn't have the opportunity to use it.

Here, the pupil in the interview brings up the idea that copying might not only affect Ola's learning outcomes but might also affect the teachers' perceptions of Ola's competence, which successively might be detrimental to Ola's learning of the skill he used ChatGPT to do for him. This notion of copying being harmful for the pupil in the long run was also mentioned by another pupil, who said the following:

I think that it can also affect like in the long run, because if you have tasks, if you get used to using it like as cheating, and you get tasks where you can't use it, then you haven't really learned anything previously, because you just cheated yourself out of it.

In addition to being harmful to the pupil, a wider societal aspect was brought up by another pupil in the same interview, who had the following example of how copying from ChatGPT might end up harming third parties as well as the pupil in question:

P: Imagine if like you have your doctor, for example, and you use ChatGPT to get yourself like through school.

I: Right. Yeah. So you wouldn't want your doctor to have been using ChatGPT.
P: Yeah. An inexperienced surgeon, for example. Yeah. Imagine if he used ChatGPT to just get through everything. That would have sucked.

Several of the pupils in the group interviews also considered the possibility of GenAI conveying wrongful- or misinformation as a reason as to why pupils should not be allowed to copy from GenAI in the English subject. One pupil said the following on the subject:

> *I believe that maybe if you use ChatGPT, you shouldn't use it to, I guess, use the answers that they give you. Because usually ChatGPT can come up with just like false answers sometimes because ChatGPT doesn't really know any of it. It browses through the internet and just, you know, finds it.*

Here, the pupil speaks to their understanding of how ChatGPT works; that it can give false answers, as it does not *know* anything. Another pupil also spoke about the unreliability of the answers of ChatGPT, and said the following:

> *But if you use it to a major thing, like for example writing your paragraphs or copy and pasting whatever ChatGPT said, it's not extremely reliable.*
>
> *Because some ChatGPT sources might be wrong or outdated.*
>
> *And so it's just best to use it as a tool and not like a full on machine doing all the work for you.*

This pupil highlights the possibility of ChatGPT's sources being outdated or wrong, echoing the sentiment from the previous pupil excerpt of not using the answers ChatGPT gives you as they may be false.

3.4.2 Homework and tests

In all three interviews, when asked, the pupils generally agreed on the fact that there should be rules concerning the use of GenAI in the English subject for both homework and tests. Their views on how they believe they should be able to use GenAI for these purposes have been described throughout chapter 4.3, so they will not be reiterated here when presenting the findings concerning restrictions for these use cases.

One restriction that was suggested was that GenAI should not be allowed to be used for aspects of the task that the pupil would be assessed on. One pupil said the following regarding using GenAI to write an essay:

> *Yeah. I think you should be able to use it to, like, if you're stuck, you should be able to use it to spike your creativity, but you shouldn't be able to use it to help you*

improve on things that you're being assessed on. So like if you're being assessed on the structure of your essay and you get Gen AI to structure it for you, you are, it's just not fair for everyone else or people that don't use Gen AI. And it's not fair for you either.

This stance on the use of GenAI for aspects of assignments that the pupil is assessed on was echoed by another pupil, who brought up the example of grammar:

And I definitely don't think you should be able to use it for grammar in like language assignments, but maybe in other subjects just to improve that for the teachers. But not in subjects where you're assessed on your grammar.

One view on restrictions revolves around restricting the use of GenAI in homework in the English subject because its use will mislead the teacher in assessing the pupils' progress. A pupil said the following on the subject:

I think that you should be restricted because homework is for the teachers, not for you. It's so that the teachers can see how well you're doing in certain things. And yeah, your classwork also contributes to that. But if you were to use AI on homework or any work in general, you're going to be showing that you know something that you don't really because you're using artificial intelligence to write or generate for you.

The view in the excerpt above that homework is for the teachers is unique to this pupil, but the view shown in the second part of the excerpt, where the pupil mentions using GenAI for "any work" will display knowledge the student does not possess, was more commonly found in other excerpts on the subject of tests; i.e. like the following excerpt:

But for tests or assignments that are graded, it's to see how much you understand, how well you understand the topic. And if you were to get AI to help you, it's kind of cheating because it's something you don't understand and you're showing that you do understand the task.

Another pupil who does not believe pupils should be allowed to use GenAI for tests gives insight into the testing praxis of the school while sharing their thought on the subject:

> *P: We don't usually do tests on the computer. If we're doing hand-ins, I don't think we should really be allowed to use ChatGPT on tests in general. Especially on sit-down tests, how would we even be able to get a ChatGPT to find the answer to our questions?*
>
> *I: Because in this example, sit-down tests is written by hand?*
>
> *P: Yeah.*

4 Discussion

In this chapter, I discuss my research questions in light of the study's findings and of previous research and guidance documents in the field of GenAI. Furthermore, I address some of the inherent limitations of the study and propose suggestions for further research on the subject of GenAI in education.

4.1 Pupils' use of GenAI

To address the research question "To which tasks, and to what degree, are pupils using GenAI in the English subject?" I look to the findings presented in Chapters 3.1.1 and 3.1.2, where the first relates to the frequency of use of GenAI and the second to which tasks pupils use GenAI for. As shown in Chapter 3.1.1, 19% of the pupils who responded to the questionnaire use GenAI for tasks in the English subject at least once a month. These results show similarities to the results of Kantar (Zulic, 2023), which found that 20% of adolescents between 15 and 24 years of age use ChatGPT for schoolwork weekly. One cannot compare these results directly, as all results are based on answers from different age groups and worded differently regarding frequency. While my questionnaire used the wording "at least once a week", Kantar's results use the word "weekly". Another difference is that my questionnaire specifically asks about using GenAI in the English subject, while Kantar's asks about using ChatGPT for schoolwork.

By grouping the categories into whether the pupils have used GenAI in the English subject or not, the findings in Chapter 3.1.1 show that 55% report using GenAI in the subject and 45% answer that they have never used GenAI in the English subject. These numbers are similar to those found regarding university students in the Student Survey, where 39% of students report that they have not used AI in their studies (Hauge & Øygarden, 2024). This, in turn, is almost identical to the Norwegian Broadcasting Organization's (NRK) survey, which showed that 59% of 18-29-year-olds had used ChatGPT for school, studies or work in the past year, or inversely that 41% had not used ChatGPT for such purposes (Vik, 2024). The data from OFCOM's report shows that 79% of online teenagers aged 13-17 and 40% of online 7–12-year-olds report using ChatGPT, Snapchat My AI, Midjourney or DALL-E. The OFCOM report did not, however, specify a context of use, and it is impossible to know whether the use reported is school-related or not.

These previously mentioned surveys were not directly aimed at pupils, with the exception of the one from the British Department for Education's Parent, Pupil and Learner Panel survey, where 14% of British pupils years 7-13 report using AI tools (including ChatGPT) for schoolwork. An important note is that the British survey was conducted in April/May 2023, mine was conducted in January/February 2024, and the NRK survey was reported in March 2024. This means that the participants in NRK's and my surveys had had access to AI tools for almost a year longer than the British pupils.

The distinctions between ChatGPT and GenAI are important to note, as they show one of the basic challenges when conducting research in the field of GenAI, namely that ChatGPT, the arguably most prominent GenAI service, is used analogous to the technology as a whole. While I assume that in the surveys Kantar and NRK conducted, they would also be interested in whether some of their survey respondents had used other GenAI services to aid in schoolwork, their choice to ask specifically about ChatGPT limits their finding to one service instead of the technology as a whole. The OFCOM report does almost the same by inquiring about four specific services: ChatGPT, Snapchat MY AI, Midjourney and DALL-E. By doing so, they exclude answers relating to any other GenAI service and have inadvertently surveyed the use of specific products, not the general technology.

Similarly to the aforementioned problem of surveys being overly specific in their description of GenAI, some of the surveys referred to in this section have language that can be construed as too vague to accurately describe the (presumed) technology they seek to explore. One example is the survey from the British Department for Education's Parent, Pupil and Learner Panel, which framed their question as "Have you used AI tools such as ChatGPT in your school or with homework?". Here, they are asking about AI tools (not limited to GenAI), while also providing the respondents with an example; ChatGPT. This is an example of how to simultaneously be overly vague and specific, leaving the survey respondent open to interpret the question in two ways: First, by not limiting the question to GenAI, the use of any AI tool without generative properties would warrant a positive response to this question, such as the likes of autocompleting search engines, voice assistants and grammar checking software. Second, by limiting the question to only involving AI tools such as ChatGPT, the respondent is tasked with establishing whether or not an AI tool is similar to ChatGPT.

The survey that is closest in resemblance to my own in the description of GenAI is the Student Survey, which gives its respondents a definition of AI and specifically states that they are referring to generative AI while also giving examples of services. However, the definition and distinction are only given before the first questions, and in the questions themselves, the term AI is being used on its own, leaving it to the respondent to remember that every time the survey says AI, it means GenAI.

The findings in Chapter 3.1.2 show that the most common use case of GenAI among the pupils participating in my survey is to get difficult words, phrases or texts explained to them, which was chosen by 43% of respondents. This is comparable to the findings in the Student Survey, where the most common use case of AI among students was to "explain subjects, syllabi, concepts, terminology etc.", which was chosen by 67% of the students who had used AI. While the percentage of respondents choosing explaining as an option was higher in the Student Survey, it was still the most frequently picked response. The second most common categories of use in my questionnaire were as a discussion partner for inspiration, brainstorming, etc.; to proofread pupil written texts; to translate text; and "other". All these options were reported by 33% of the participants. These findings also show similarities with the findings in the Student Survey, where around 40% chose both discussion partner and quality assurance of self-written text, being tied for the second-most chosen option in that survey. Where 10% in my survey answered that they have used GenAI to generate new text, 23% in the Student Survey answered the same.

A finding that stands out regarding what the pupils have used GenAI for, is that none report using it for tasks that will be graded. While this is unexpected in light of the ongoing debate on implementing GenAI in schools, where the preconception seems to be that pupils cheating with GenAI for graded assignments, it also matches NRK's survey, which found that only four per cent of 18-29-year-olds stated to have used AI for something they know is prohibited. However, this finding from my survey is partly explained by information gathered in the group interviews. In Chapter 3.2.2, an excerpt is shown in which a pupil gave some insight into the testing practices at the school where I conducted my research. The pupil stated that the tests being conducted at school were handwritten, removing the possibility of pupils using GenAI for sit-down tests. This test condition does not apply to their essays, or other graded work, so it cannot

be seen as the sole reason why none of the pupils report having used GenAI for graded work. However, in Chapter 3.3.4, another pupil points out that they do not do much research anymore, and that digital assignments have become stricter, focusing on knowledge gathered in class or from a textbook. There are surely other factors that could affect this finding, for instance, the sampling for the study. If a pupil had previously used GenAI for a graded assignment with the purpose of cheating, is it not reasonable to believe that the same pupil could have reservations towards participating in a research project concerning GenAI, in fear of being caught? It is also possible that the participants were untruthful, choosing to under-report negatively connoted behaviour, a common phenomenon in surveys.

4.2 Use of generative AI across different levels of achievement

In order to discuss research question two, namely "In what ways, if any, does the use of GenAI differ between pupils of different levels of achievement in the subject?", I will refer to the findings presented in Chapter 3.2.1, and 3.2.2. Regarding the frequency of use, the main finding shows that more pupils expecting higher grades have used GenAI than those expecting lower grades. Within the group that answered that they expect to receive grades 3 and 4 in the English subject, 59% answered that they have never used GenAI in the English subject, compared to those expecting grades five and six, where 33% answered that they have never used GenAI in the English subject. Regarding type of use, the largest disparities shown in the findings in Chapter 3.2.2 were that the use case of GenAI as a discussion partner for inspiration, brainstorming, etc., was reported by 43% of grades five and six, and 14% of grades three and four. The second largest use case was for explaining, picked by 50% of grades five and six, and 29% of grades 3 and 4. 7% of the pupils expecting higher grades had used GenAI to generate new text, while 14% of the pupils expecting lower grades answered the same.

The dissimilarities in the type of use of GenAI between the two groups bring some interesting depth to the discussion on whether GenAI can deepen a figurative digital divide. As mentioned in the Beijing Consensus, the use of AI in education should promote high-quality education and learning for all, and not deepen the digital divide. To further this point UNESCOs "Guidance for generative AI in education and research" (Holmes & Miao, 2023, p. 36), brings up the

possibility of GenAI contributing to a digital divide, meaning that individuals with stronger prerequisites for using digital aids such as GenAI will get further ahead than those who do not. The types of use where the pupils expecting higher grades are more represented are reminiscent of what Holmes and Miao (2023) mention as use cases that do not cause over-reliance, namely to challenge and extend human thinking. As the pupils expecting higher grades are the ones who use GenAI the most, and they have found more beneficial use cases for the technology than those who expect lower grades, I believe there is cause to say that a digital divide is being created among these pupils across the different levels of competence regarding GenAI.

4.3 Pupils' wanted use of GenAI

To discuss the third research question, **"In what manner do pupils believe they should be able to use GenAI in the English subject?",** I will utilize the themes that emerged from the group interviews as presented in Chapters 3.3 and 3.4. Before delving deeper into these themes, it is important to note that while they are presented in the findings chapter in descending order based on the number of coding instances of that theme, that order should not be used as a metric for gauging the level of agreement between pupils on a certain theme. A theme with fewer coding instances might have been brought up fewer times in the interviews as a result of pupils agreeing with one another and choosing not to bring up a theme several times.

The themes found in the interview data regarding the pupils' wanted use of GenAI in the English subject were **inspiration**, **explanation, help, information,** and **improving language.** Themes regarding how the pupils believe GenAI should not be used in the English subject were also found in the interview data: copying**, homework and tests**. These themes of wanted use are reminiscent of the options found in the survey question on type of use, and the amount of coding instances show similarities with the answer distribution in the questionnaire, namely that inspiration and explanation are the most prominent themes in both the survey data on previous use and in the interview data on wanted use of GenAI. This shows that the pupils' wanted use of GenAI in the English subject mostly mirrors the reported use of the pupils that have used GenAI, which might indicate that pupils have already found use cases for the technology and that they want to continue using it in such a manner.

Based on my findings, it seems like many pupils want to be able to use GenAI for inspiration in the English subject. In the excerpts from the group interviews, several specific use cases are mentioned, such as using GenAI for paragraph starters, sentence starters, ideas and guidelines to get them started in the writing process. As such, this suggested use case might be interpreted to mean that these pupils want more detailed outlines and structure for their writing than their teachers are providing them with. Using GenAI to enrich creative processes is discussed in the UDIR competence package for teachers, where using it for giving ideas is listed as an example of good practice. While there are certainly different levels to such use, the possibility of overreliance, as discussed in the previous section, is still an important aspect to consider when discussing getting inspiration from GenAI. The interviewed pupils say that they believe such use is fine as long as only the idea comes from GenAI, and the main body of work is still done by the pupil. One interesting point is made by one of the pupils, who says that tests are more intended to let them show their creativity, but in essays, they are meant to cite their sources. This highlights the point that if they are not assessed on creativity, why should they not be able to use GenAI for inspiration?

Several of the pupils specify homework as a situation where they might want to be able to use GenAI to get inspiration, mentioning that if a pupil is stuck, they can use GenAI to help them get going. The matter of homework is mentioned in connection to using GenAI for explanation and help as well, where pupils believe they should be able to have GenAI explain difficult tasks to them, or help them understand the steps needed to solve a problem. As not all pupils have the same access to help outside of school, using GenAI for those purposes serves as an example of a way in which GenAI can aid in improving equity for all pupils, giving them equal opportunities to perform the tasks asked of them.

The pupils' view on wanting to be able to use GenAI for help, and not necessarily giving them a final answer seems to be shared by many. This is reminiscent of a popular implementation of GenAI in the education sector, namely chatbots that avoid giving straight answers, but rather attempt to function as a tutor. One such chatbot is Khan Academy's "Khanmigo", which, according to their FAQ, does not simply tell the user answers but rather guides them to find the answer themselves (Khan Academy, 2024). Such a use case is reminiscent of one of UNESCO's "Potential but unproven uses" of GenAI to facilitate inquiry or project-based learning, namely

as a Socratic challenger, with the expected outcome being to "Engage learners in dialogue reminiscent of the Socratic questioning of prior knowledge, leading to the discovery of new knowledge or deeper understanding." (Holmes & Miao, 2023). An important note to this point is that while chatbots designed to act as a Socratic challenger will always do so, to achieve the same results from a more generic chatbot like ChatGPT requires the user to ask for it, specifying that they do not want the answer directly. To facilitate such a use case, the teacher would need to instruct their pupils on how to use such specialized chatbots or achieve the same results from more generic chatbots.

Regarding using GenAI to find information, some of the pupils in the group interviews stated that they wanted to be able to use it to enrich a writing process with information. This was suggested in various ways, such as to supplement other information, or as a basis for writing which would need to be edited by the pupil. The participating pupils' sense of academic honesty shone through here, with several pupils stating that they should ask their teacher before using GenAI to get information, and that information gathered from GenAI should be verified using other sources. These attitudes also indicate that the pupils have some understanding of the limitations of GenAI in providing reliable information and that their beliefs on how they should be able to use GenAI are based on these limitations.

The pupils in the group interviews had several notions regarding whether they should be able to use GenAI to improve the language of their texts. They bring up possible use cases such as finding suitable words, rewriting sentences, rewriting whole texts, shortening texts and checking for grammatical mistakes. What stands out across the findings regarding language in Chapter 3.3.5, is the view that while GenAI can be used to improve the language in their texts, the pupils should be taught to use it sustainably. By sustainably, I interpret the pupil to mean a type of use that ensures that the pupils still learn to create text without being reliant on external tools like GenAI. This view is, as mentioned, also shared with UNESCO's guidance on GenAI, which stresses that an overreliance on GenAI can compromise the development of intellectual skills (Holmes & Miao, 2023).

As mentioned, the analysis of the interviews also yielded two themes regarding how pupils believe GenAI should not be used in the English subject, the first of which is related to copying.

When presented with the case of Ola, who copied from GenAI to complete an English assignment, all groups agreed that copying directly from GenAI was cheating, and should not be allowed. The same was also true for a situation where Ola somewhat edited the output from GenAI, where even if Ola cited GenAI as a source, the pupils in the group interviews considered it plagiarism. This strong notion that using text generated by GenAI directly in their work is cheating also serves as context for the survey finding that zero per cent of the pupils asked had used GenAI for graded assignments. The one pupil who compares copying from GenAI to copying from a Wikipedia page, saying it is the same, also underpins the notion that the pupils believe copying from GenAI is cheating. This does, however, possibly show this pupils' lack of understanding of the differences between the two, because while both cases are plagiarism, they are fundamentally different in scope. This is because copying from Wikipedia is easily uncovered by teachers while proving that a pupil has copied from GenAI is very difficult. This is pointed out by another pupil in the same excerpt, saying that you can't really cite ChatGPT, inadvertently pointing out the fact that text generated by such GenAI tools is not searchable or replicable.

The pupils in the three group interviews showed reflections on the negative aspects of unrestricted use of GenAI. They could see possible harmful ways of using the technology, one of which being that pupils using GenAI to do tasks for them could be detrimental to their learning outcomes. The pupils give several examples of this, such as if a pupil always uses GenAI to help them analyse characters, they will never know how to do it themselves. In addition to being harmful to the pupils' learning, the harm that unrestricted use of GenAI could cause outside of school was also mentioned in the group interviews. One pupil brought up the example of a doctor having used ChatGPT throughout their education, pointing out that this would not be ideal for the patient. This high-level view on the effects GenAI could have on learning outcomes is similar to parts of what was proposed in the Beijing Consensus, namely that governments were recommended to support research and analysis on the effects of AI on learning outcomes (UNESCO, 2019), and to the part of UNESCO's guidance, stating that GenAI should not usurp human thinking (Holmes & Miao, 2023).

The second theme regarding how pupils believe they should not be able to use GenAI in the English subject is related to homework and tests. The main finding in this regard is that the pupils believe they should not be able to use GenAI for aspects of homework or tests they are assessed on. One such aspect is grammar, where one pupil suggests that it may be allowed to use GenAI to improve the readability of texts in subjects where they are not assessed on grammar; this should not be the case in subjects where grammar is assessed. In a similar vein, the pupils mention the disparities that are created when this is not enforced. In Chapter 3.4.2 the pupil in one of the excerpts mentions that if some pupils use GenAI for aspects they are assessed on, such use is unfair to the pupils who do not use GenAI. This view can be linked to the previously discussed possibility of GenAI contributing to a "digital divide", but in this instance, a divide would happen between those who choose to use it and those who choose not to. This does not, however, seem to be a challenge among the pupils who participated in this project, as the survey shows that none of the pupils seem to be using GenAI for graded assignments.

4.4 Limitations with the study and thoughts on further research

While implementing a mixed research method allowed me to explore several aspects of pupils' use of GenAI in the English subject, certain aspects were not possible to explore within the confines of the methods chosen. One such aspect was pupils' views on their own prior use of GenAI, which was not possible to discuss in a group setting, as the pupils should not be asked to discuss what might be construed as cheating in front of each other. Another aspect that was not possible to explore based on my research design was the possible differences in how the pupils wanted to be able to use GenAI in the English subject in relation to their competence in the subject. A possible way for further research to address this is to conduct individual interviews, wherein pupils will not have to worry about incriminating themselves in front of their peers and where they could be asked about their grades in the English subject. One strength of group interviews, that the pupils feel safer when outnumbering the interviewer who is unknown to them, would not apply in such a setting, but a skilled interviewer might still be able to make the participating pupils comfortable in an individual setting.

Individual interviews could also uncover what the participating pupils meant by choosing the alternative “other” when asked what they had used GenAI for. The same is also true for any aspect that the researcher could not anticipate regarding pupils’ use of GenAI; in individual interviews, one could uncover such aspects, and probe for more information immediately. In a field where so much is still unknown, and there is little research to base one’s own research on, the ability to inductively uncover new aspects of the subject as a part of the research would be a great advantage.

While my research focuses on the pupils’ use and wanted use of GenAI, this could also be explored in relation to their teachers, and the teachers’ implementation of GenAI in their teaching. An example of such a research project could be to see how the use of GenAI differs between different groups of pupils whose teachers have different rules or approaches to GenAI in their classes. While the current guidelines for teachers suggest that teachers educate their pupils on how to use GenAI in academically sound manners, little is known of how this affects the pupils’ use of GenAI.

Even though there are many GenAI tools explicitly developed for use in education, such as the chatbot Khanmigo mentioned in the earlier discussion, research on the efficacy of such tools is still lacking. Before adopting such tools, it would be helpful for teachers and schools to know if there is a meaningful difference in using such tools over other popular chatbots such as ChatGPT. An example of a research project exploring using such tools could be introducing an education-oriented chatbot to a group of pupils, exploring whether pupils would choose the education-oriented chatbot over other alternatives like ChatGPT.

While the prospect of an “ideal use” of GenAI is hinted at in this thesis, there is still no evidence-based research on what constitutes an ideal use of the technology. Research into the effect that using GenAI has on pupils' learning outcomes is needed before a discussion on pupils' ideal use of GenAI can go any further. However, exploring the effect single variables have on learning outcomes is difficult to isolate in classroom research. A better way to understand the effect GenAI has on learning outcomes, would be to explore its use in longitudinal studies, where the long-term effects of different types of use of GenAI could

become apparent. For the English subject, this could include exploring the long-term effect of the use of GenAI on pupils' development on grammatical competence.

While countries continue to develop their own strategies for GenAI within education, sharing information across country borders by conducting international studies, sharing plans, and constantly evaluating different implementations of the technology is of the utmost importance. As a part of such a process, the translation of guides and competence packages to English would be an important step. This should not be done as with the UDIR competence package discussed earlier in this thesis, kept behind a login wall and only available in Norwegian, where its content is not searchable without access, effectively hindering international researchers from evaluating Norway's work in regulating and training for GenAI technology. As adapting to this technology is important to all countries, schools and teachers, I view the secrecy involved in Norway's adaptation to GenAI as unnecessary, as progress would be faster if all countries were open in their practices regarding the technology, opening them up for scrutiny and research.

While conducting this research, the development and progress in the space of GenAI have been constant, with new tools being released that are more advanced and feature-rich than their predecessors. Such rapid development makes research into GenAI in a school context difficult, as the limits for what GenAI can do are constantly pushed. To bring this point home, these final paragraphs are written based on a recent development in the field of GenAI, which has certain implications for new types of use of GenAI in education. At the time of writing this, May 14th 2024, OpenAI has just released a new foundation model called GPT-4o, which can accept text, audio and image as inputs, even combined, and can generate outputs in any combination of the three as well. The model is stated to respond to audio inputs in 0,32 seconds on average, mimicking human response time, which, combined with video, makes it possible to converse with the model in real time. These advancements bring new use cases for the technology in a school context, which will have to be explored, as they may possibly bring solutions to several of the challenges of using GenAI in education, and even specifically in the English subject. An example of this could be to explore pupils exclusively using GenAI tools orally in the English subject, as opposed to through text, which could mitigate the risk of plagiarism and overreliance on GenAI for writing tasks. If this new model is implemented into education-specific GenAI

applications, UNESCOs suggestion of using GenAI as a Socratic challenger could also be explored in an exclusively oral setting, and its efficacy as a tutor could be evaluated.

Another use case for GenAI in the English subject, which is made possible with the release of GPT-4o, is teaching aspects of oral English. As the model can listen and give feedback with examples, it can show pupils how to pronounce specific words and even give oral examples of sarcasm. As the GenAI tools become even more fluent in more languages, their ability to translate, combined with oral capabilities, will possibly lead to the discovery of even more use cases for GenAI in the English subject. Future research into GenAI in the English subject must take these new use cases into consideration and be open to the possibility that pupils have found new use cases for the technology, which may be unknown to teachers and researchers alike.

4.5 Concluding remarks

The goal of this research project is to answer the thesis question, “In what ways do pupils use GenAI in the English subject, and how do pupils believe it should be used?”. Several ways in which the pupils use and want to use GenAI have been presented and discussed, but as the technology develops as rapidly as it does, these findings must be seen in the context of when and where they were found. As such, their relevance is possibly short-lived. I believe that there are findings, however, that stand out in this research project and are not soon to be irrelevant in the face of new technology. One such finding is that in lieu of clear national rules and regulations, pupils’ reported use of GenAI is mostly in ways which do not constitute cheating, but rather as a means to provide ideas and inspiration, explain and translate difficult text, and improve on texts written by the pupil. Likewise, the findings show that pupils reflect on the use of GenAI in their own work, conscious of its possible negative impacts on themselves and on a societal level.

Works cited

Academy, K. (2024). *Khanmigo for learners*. Khan Academy. Retrieved 16.04 from https://www.khanmigo.ai/learners

Arksey, H., & Knight, P. (1999). Interviewing for Social Scientists. In. SAGE Publications, Ltd. https://doi.org/10.4135/9781849209335

Bigum, C. (2012). Schools and Computers: Tales of a Digital Romance. In L. Rowan & C. Bigum (Eds.), *Transformative Approaches to New Technologies and Student Diversity in Futures Oriented Classrooms: Future Proofing Education* (pp. 15-28). Springer Netherlands. https://doi.org/10.1007/978-94-007-2642-0_2

Bjørndal, C. R. P. (2017). *Det vurderende øyet : observasjon, vurdering og utvikling i pedagogisk praksis* (3. utg. ed.). Gyldendal akademisk.

Brinkmann, S., & Kvale, S. (2018). Doing Interviews. In (Second ed.). SAGE Publications Ltd. https://doi.org/10.4135/9781529716665

Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education* (Eighth edition. ed.). Routledge.

Gleiss, M. S., & Sæther, E. (2021). *Forskningsmetode for lærerstudenter : å utvikle ny kunnskap i forskning og praksis* (1. utgave. ed.). Cappelen Damm akademisk.

Guest, G., MacQueen, K., & Namey, E. (2012). *Applied Thematic Analysis* https://doi.org/10.4135/9781483384436

Hauge, M. S., & Øygarden, K. F. (2024). *Studiebarometeret 2023 – Hovedtendenser* (4-2024). https://www.nokut.no/globalassets/studiebarometeret/2024/hoyere-utdanning/studiebarometeret-2023_hovedtendenser_4-2024.pdf

Holmes, W., & Miao, F. (2023). *Guidance for generative AI in education and research*. UNESCO. https://books.google.no/books?id=cKnYEAAAQBAJ

Johansen, J. I. (2023, August 18). Slik vil ekspertene at KI skal brukes i skolen: – Kunstig intell-igens er mye mer enn juks. *Norwegian Broadcasting Organization*. https://www.nrk.no/vestland/slik-vil-ekspertene-at-ki-skal-brukes-i-skolen_-_-kunstig-intelligens-er-mye-mer-enn-juks-1.16516043

KANTAR. (2023). *Medietrender Ung 2023 - ny rapport om ungdommens medievaner*. Retrieved 12.03 from https://kantar.no/rapporter/medietrender-ung-2023

Krumsvik, R. J. (2024). Everyday Discourse and Research Discourse on Large Language Models. *Nordic Journal of Digital Literacy*, *19*(1), 4-7. https://doi.org/10.18261/njdl.19.1.1

Luo, R., Sun, L., Xia, Y., Qin, T., Zhang, S., Poon, H., & Liu, T.-Y. (2022). *BioGPT: generative pre-trained transformer for biomedical text generation and mining* [bbac409]. [London] :.

Maniaci, M. R., & Rogge, R. D. (2014). Caring about carelessness: Participant inattention and its effects on research. *Journal of Research in Personality*, *48*, 61-83.

Ministry of Education and Research. (2023). *Strategy for digital competence and infrastructure in kindergardens and schools [Strategi for digital kompetanse og infrastruktur i barnehage og skole]*. Retrieved from https://www.regjeringen.no/contentassets/3fc31c3d9df14cc4a91db85d3421501e/no/pdfs/strategi-for-digital-kompetanse-og-infrastruktur.pdf

Munthe, E., Erstad, O., Njå, M. B., Forsström, S., Gilje, Ø., Amdam, S., Moltudal, S., & Hagen, S. B. (2022). *Digitalisering i grunnopplæring; kunnskap, trender og framtidig kunnskapsbehov*. K. f. utdanning. https://www.udir.no/tall-og-forskning/finn-forskning/rapporter/digitalisering-i-grunnopplaringen-bedre-muligheter-for-laring/

Norwegian Directorate for Education and Training. (2024a, January 18 2024). *Kompetansepakke om kunstig intelligens i skolen*. Norwegian Directorate for Education and Training,. https://www.udir.no/kvalitet-og-kompetanse/digitalisering/kompetansepakke-om-kunstig-intelligens-i-skolen/#a194065

Norwegian Directorate for Education and Training. (2024b, January 18 2024). *Kunstig intelligens i skolen*. Norwegian Directorate for Education and Training,. https://bibsys.instructure.com/search/all_courses?design=udir#637

Norwegian Directorate for Education and Training. (2024c, February 29 2024). *Råd om kunstig intelligens i skolen*. Norwegian Directorate for Education and Training,. https://www.udir.no/kvalitet-og-kompetanse/digitalisering/kunstig-intelligens-ki-i-skolen/#a210115

OFCOM. (2023). *Online Nation 2023 Report*. https://www.ofcom.org.uk/__data/assets/pdf_file/0029/272288/online-nation-2023-report.pdf

OpenAI. (2022, November 30). Introducing ChatGPT. *Blog*. https://openai.com/blog/chatgpt

OpenAI. (2024). *Hello GPT-4o*. Retrieved May 14 from https://openai.com/index/hello-gpt-4o/

RM Technology (2023, June 22 2023). *A staggering two thirds of secondary school students use AI to do their school work*. https://www.rm.com/news/2023/artificial-intelligence-in-education

Tashakkori, A., Johnson, B., Teddlie, C., & Tashakkori, A. (2020). *Foundations of mixed methods research : integrating quantitative and qualitative approaches in the social and behavioral sciences* (Second edition. ed.). SAGE Publications, Inc.

The National Committee for Research Ethics in the Social Sciences and the Humanities. (2022). *Guidelines for Research Ethics in the Social Sciences and the Humanities* (L. Auestad, Trans.).

The Open Innovation Team, & British Department for Education. (2024). *Generative AI in education: educator and expert views*. Retrieved from https://assets.publishing.service.gov.uk/media/65b8cd41b5cb6e000d8bb74e/DfE_GenAI_in_education_-_Educator_and_expert_views_report.pdf

UNESCO. (2019). *Beijing Consensus on Artificial Intelligence and Education* International Conference on Artificial Intelligence and Education, Planning Education in the AI Era: Lead the Leap, Beijing, 2019, Beijing. https://unesdoc.unesco.org/ark:/48223/pf0000368303

UNESCO. (2023, 07.09.2023). *UNESCO: Governments must quickly regulate Generative AI in schools* https://www.unesco.org/en/articles/unesco-governments-must-quickly-regulate-generative-ai-schools?hub=83250

University of Oslo. (2023a, February 27 2023). *Autotekst i ny versjon: God og enkel automatisk teksting på norsk*. Retrieved May 6 from https://www.uio.no/tjenester/it/aktuelt/om-it/2023/autotekst.html

University of Oslo. (2023b, February 3, 2023). *Nettskjema-diktafon*. Retrieved May 2, 2024 from https://www.uio.no/tjenester/it/adm-app/nettskjema/hjelp/diktafon.html

Vaughn, S., Schumm, J. S., & Sinagub, J. (1996). Focus Group Interviews in Education and Psychology. In. SAGE Publications, Inc. https://doi.org/10.4135/9781452243641

Vik, I. (2024, March 26 2024). Flertallet av unge bruker KI i skole, studier eller jobb: – Mange bruker det ganske simpelt. *Norwegian Broadcast Organization*. https://www.nrk.no/tromsogfinnmark/flertallet-av-unge-bruker-ki-i-skole_-studier-eller-jobb-1.16815700

Zulic, A. R., Mariann; Granbo, Kristin. (2023, 16.06.2023). 1 av 5 elever bruker ChatGPT til skolearbeid. *NRK*. https://www.nrk.no/kultur/na-sier-1-av-5-norske-elever-at-de-bruker-chatgpt-til-skolearbeid-1.16442993

Appendix 1 - Notification Form

Sikt

Notification Form

Reference number
519228

Which personal data will be processed?

- Voice on audio recordings

Project information

Title
Masterprosjekt - Elevers bruk av generativ kunstig intelligens i engelskfaget

Summary
Formålet med prosjektet er å kartlegge elevers bruk av generativ kunstig intelligens i engelskfaget, samt undersøke hvordan elever mener slike tjenester burde kunne brukes i faget. Med generativ kunstig intelligens mener man tjenester som ChatGPT, Bard, Copilot o.l.

What is the purpose for processing personal data?
Som del av forskningen skal et gjennomføres gruppeintervju med elever. Da jeg er alene i dette prosjektet trenger jeg å ta lydopptak av disse intervjuene slik at det kan transkriberes i ettertid.

External funding
Ikke utfyllt
Type of project
Master's

Contact information, student
Sondre Ingerøyen, sin028@uit.no, tlf: 95978713

Data controller

Institution responsible for the project
UiT Norges Arktiske Universitet / Fakultet for humaniora, samfunnsvitenskap og lærerutdanning / Institutt for lærerutdanning og pedagogikk

Project leader
Hilde Brox, hilde.brox@uit.no, tlf: 77660507

Do multiple institutions share responsibility (joint data controllers)?
No

Sample 1

Describe the sample
Elever i ungdomsskolealder

Describe how you will identify or contact the sample
Samarbeid med lærer på skole. Alle elever i denne lærerens klasser fyller ut et nettskjema, og man lager grupper på seks elever fra hver engelskklasse som skal intervjues

Age group
13 - 16

Are any of these groups included in the sample?
- Vulnerable groups

Which data relating to sample {{i}} will be processed? 1
- Voice on audio recordings

How will data relating to sample 1 be collected?

Group interview

Attachment

Intervjuguide engelsk.docx

Legal basis for processing general personal data
Consent (General Data Protection Regulation art. 6 nr. 1 a)

Who will give consent for children under 16 years?
Parents/guardians

Who will give consent for youths 16-17 years?
Parents/guardians

Online survey

Attachment

Questionnaire GenAI in the eglish subject v1.pdf

Legal basis for processing general personal data
Consent (General Data Protection Regulation art. 6 nr. 1 a)

Who will give consent for children under 16 years?
Parents/guardians

Who will give consent for youths 16-17 years?
Youths

Information for sample 1

Will the sample receive information about the processing of personal data?
Yes

How does the sample receive information about the processing?
Written (on paper or electronically)

Information letter

Information letter GenAI.docx

Third persons

Will the project collect information about third persons?
No

Documentation

How will consent be documented?

- Manually (on paper)

How can consent be withdrawn?
Ved å kontakte ansvarlig på e-post, telefon eller muntlig.

How can data subjects get access to their personal data or have their personal data corrected or deleted?
Ved å kontakte ansvarlig på e-post, telefon eller muntlig. Relevante deler av lydopptak og transkribering kan så gis innsyn i, rettes eller slettes etter ønske.

Total number of data subjects in the project
1-99

Approvals

Will any of the following approvals or permits be obtained?
Ikke utfyllt

Security measures

Will the personal data be stored separately from other data?
Yes

Which technical and practical measures will be used to secure the personal data?

- Restricted access
- Continuous anonymisation
- Encrypted storage
- Multi-factor authentication

Where will the personal data be processed

- ?
- Mobile devices
- Physically isolated hardware

Who has access to the personal data?

- Student (student project)
- Project leader
- Data processor

Which data processor will be processing/have access to the collected personal data?

UiO - nettskjema

Will personal data be transferred to a third country?

No

Closure

Project period

08.01.2024 - 31.07.2024

What happens to the data at the end of the project?

Personal data will be anonymised (deleting or rewriting identifiable data)

Which anonymisation measures will be taken?

- Personally identifiable information will be removed, re-written or categorized
- The identification key will be deleted
- Any sound or video recordings will be deleted

Will the data subjects be identifiable in publications?

No

Additional information

Appendix 2 - Consent form

Are you interested in taking part in the research project

"Pupils' use of generative artificial intelligence in the English subject"?

Purpose of the project

You are invited to participate in a research project whose main purpose is to explore pupils' use of generative artificial intelligence (GenAI) in the English subject. Its main focus is to explore how pupils believe such applications should be used in the subject. By GenAI we mean applications like ChatGPT, Bard, Copiliot etc. This form gives you information about the aims for the project, and what participation involves.

The extent of the research consists of an anonymous questionnaire about the use of GenAI, and three group interviews regarding attitudes toward GenAI.

Which institution is responsible for the research project?

UiT – The Arctic University of Norway is responsible for the project (data controller).

What does participation involve for the pupil?

If the pupil chooses to participate in the project, this will involve the pupil filling in an online questionnaire. It will take approx. 10 minutes. The questionnaire includes questions about the use of GenAI in the English subject. The pupils' answers will be recorded electronically, and are completely anonymous.

If the pupil chooses to participate in the project, they may be asked to participate in a group interview. The interview will take approx. 30 minutes. The interview questions will be on attitudes toward GenAI and the pupils' thoughts on how this technology should be used in the English subject. Audio from the interview will be recorded and transcribed afterwards. The audio recordings will be deleted when the project is finished.

Participation is voluntary

Participation in the project is voluntary. If you choose to participate, you can withdraw your consent at any time without giving a reason. All information about you will then be made anonymous. There will be no negative consequences for you if you choose not to participate or later decide to withdraw.

Your personal privacy – how we will store and use your personal data

We will only use your personal data for the purpose(s) specified here and we will process your personal data in accordance with data protection legislation (the GDPR).

- It is only the student and their supervisor that will have access to managing your personal data that are collected in this project.
- We do not seek to record your name and any other personal data. In the event that such information is mentioned in the interview, it will be redacted, and names will be replaced with fictitious ones. The name of the school will be anonymised, to prevent recognition when the master's thesis is published.
- The interviews will be recorded using the online service Nettskjema-diktafon from the University of Oslo, which encrypts the recordings and stores them securely without needing to download them to personal computers.

- The interviews will simultaneously be recorded using a separate recorder, as a backup if the online service fails. The audio files on the separate recorder will be deleted after all the interviews are transcribed.

What will happen to your personal data at the end of the research project?

The planned end date of the project is 31.07.2024. Any personal information and audio recordings will be deleted at the end of the project.

Your rights

So long as you can be identified in the collected data, you have the right to:

- access the personal data that is being processed about you
- request that your personal data is deleted
- request that incorrect personal data about you is corrected/rectified
- receive a copy of your personal data (data portability), and
- send a complaint to the Norwegian Data Protection Authority regarding the processing of your personal data

What gives us the right to process your personal data?

We will process your personal data based on your consent.

Based on an agreement with *UiT – The Arctic University of Norway*, The Data Protection Services of Sikt – Norwegian Agency for Shared Services in Education and Research has assessed that the processing of personal data in this project meets requirements in data protection legislation.

Where can I find out more?

If you have questions about the project, or want to exercise your rights, contact:

- *UiT-The Arctic University of Norway by student Sondre Ingerøyen, or supervisor Hilde Brox.*

  Sondre Ingerøyen

sin028@uit.no

Hilde Brox

hilde.brox@uit.no

- Our Data Protection Officer:
  Anniken Steinbakk

  personvernombud@uit.no

  Phone: 77 64 69 52

If you have questions about how data protection has been assessed in this project by Sikt, contact:

- email: (personverntjenester@sikt.no) or by telephone: +47 73 98 40 40.

Yours sincerely,

Sondre Ingerøyen	Hilde Brox
(student)	(associate professor/supervisor)

Consent form

I have received and understood information about the project *"Pupils' use of generative artificial intelligence in the English subject"* and have been given the opportunity to ask questions. I give consent for _____ to participate in the project.

(Name of the pupil)

I give consent for the participants' personal data to be processed until the end of the project.

(Signed by parent, date)

(Signed by pupil, date)

Appendix 3 - Interview guide

Interview guide for the group interview

The interview guide is meant to encourage discussion around the use of generative artificial intelligence (GenAI) like ChatGPT. It is worded in such a way as to let pupils discuss the subject without having to discuss their personal use of GenAI.

Before starting the recording:

You can back out at any time, without any consequences. The recordings, however, are anonymous, and your name will never be mentioned anywhere.

Intro:

Hello!

Before anything else: Please do not name names or identify fellow pupils, teachers or parents.

I am conducting this interview because you are the only ones that can help me answer my research question: How does pupils want GenAI (ChatGPT) to be used in the English subject? My intention is that we talk about this without you having to talk about your own use of these tools, or anyone else's. You will be asked directly about your own use in the survey later.

What GenAI is, and the status of rules.

Do you know what ChatGPT is?

Questions:

1. Do you think there should be rules for using GenAI for homework?
    - Do you think the rules for using GenAI should be the same for tests as they are for homework/assessments in the English subject?
    - If you think they should be different, how so?
2. Do you think that using GenAI affects the learning of pupils who use it?
3. Do you think that using GenAI affects the grades of pupils who use it?
4. What do you think can be possible pros or cons of using GenAI in the English subject?
   Possible pros and cons: Time, learning, etc.
5. Is writing texts in the English subject different after GenAI's like ChatGPT became available?
   Possible examples: Handwriting, explicit instruction etc.
    - If yes, in what way?
    - Do you feel that task descriptions are different from what they used to be?
    - Do you feel that tasks are assessed differently from how they used to be?
6. How do you want GenAI to be used in the English subject?

Case:

Ola is working on an English task where he is supposed to write 700 words about the characters from a book he could choose himself. He is expected to write both at school and home. The teacher has not mentioned anything about using GenAI. Ola lets ChatGPT write a paragraph for each of the main characters in the book, which he then edits a bit.

- What do you think of Ola's use of GenAI in this situation?
- How about if the teacher prohibited the use of GenAI?
- How about if Ola did not edit anything from what he copied from ChatGPT?
    - Which means he copied directly.