EdTech and Ethics: Monitoring Suicide Risk in UK Secondary Schools

Jessica Anne Lorimer

Department of Psychiatry, Jesus College

This thesis is submitted in partial fulfilment of the requirements for the degree of DPhil at the Department of Psychiatry, University of Oxford

Supervisors: Professor Ilina Singh, Professor Thomas Douglas

Word Count: 53,148

Acknowledgements

I would like to express my gratitude to my supervisors, Professors Ilina Singh and Tom Douglas, for their invaluable feedback, critical supervision meetings, and compassionate support (both academic and personal) throughout my DPhil journey. I am also grateful for the scholarships that made this research possible, including support by the Medical Research Council through the MRC-DTP Scholarship, as well as the Oxford-Hoffmann Scholarship based at Jesus College.

Thank you to all of the teachers who generously shared their experiences and took the time to participate in this research. I also wish to acknowledge the young people who have lost their lives to suicide, and the families, friends, and school communities who continue to carry their memory. This research is, in part, a tribute to the resilience of those young people and a recognition of the urgent need to support adolescent mental health.

Thank you to everyone I have worked with in the NEUROSEC team at the Department of Psychiatry. To Gabi, Alexandra, David, Katrin, Eddie, Gulamabbas, Arianna, Bessie, and so many others, thank you for being such a consistent and caring presence over the years.

Thank you to my family, whose love and support have inspired this work from the very beginning. To my parents, to Graeme, and to all of my cousins, aunts, uncles, and extended family who have surrounded me throughout this journey. Thank you also to my family-away-from-home, and to every single person in Oxford who I've met over the last eight years who has acted as a study buddy, confidant, and friend. Thank you to Kiran and Paige; to Alice and Shati and everyone from GoodGym; to Amelia, Jyoti, Clara, Maria, and Yvonne; to Tess; to Marisha and Ruth; to Yung and Sruj. To everyone at Oxford Homeless Project for the work they do. To all of the reality TV watching groups (of which there are many). Thank you to Sepi, always.

And finally, to my grandparents. Judy, Richard, Alice, and Donald: thank you for always believing in me, for your unwavering encouragement, and for having my back every step of the way. This thesis is for everyone, but above all, it is for you.

Table of Contents

Acknowledgements 2
Abstract 8
List of Figures 10
List of Tables 11
Abbreviations 12
Notes on Terms 13
Suicide Ideation and Behaviour 13
Suicidal Self-Injurious Behaviour and Non-Suicidal Self-Injurious Behaviour 14
Risk 14
Screening 15
Monitoring 16
Machine Learning 16
EdTech 17
Prevention 18
COVID Statement 20
Chapter 1. Introduction 21
1.1. Aim 22
1.2. Structure 22
1.3. History of Suicide Prediction Technology 23
1.3.1. ‘Push’ and ‘Pull’ Factors 26
1.3.1.1. Repurposed Tools and PREVENT 26
1.3.1.2. Additional Funding and Regulation ‘Pulls’ 29
1.3.1.3. COVID-19 30
1.3.1.4. CAMHS Funding Constraints 31
1.4. State of Play 33
1.4.1. Timeline 34
Chapter 2. Developing the Research Scope 37
2.1. Beneficence 38
2.1.1. EdTech Aligned with the Principle of Beneficence 38

2.1.2. EdTech In Tension with the Principle of Beneficence 39
2.1.2.1.1. Data on Accuracy 40
2.1.2.1.2. Safety Concerns based on Inaccuracy 41
2.2. Introducing Tradeoffs 42
2.2.1. Beneficence, Autonomy, and Privacy 43
2.2.1.1. Definition of Privacy 44
2.2.1.2. Example 1: (Covert) Surveillance 45
2.2.1.3. Example 2: Data Sharing 47
2.2.2. Beneficence vs Justice (and Discrimination) 48
2.3. Responsibility Framework 50
Chapter 3. Methodology 53
3.1. Aims, Objectives and Research Questions 53
3.2. Defining Responsibility 56
3.2.1. Responsibility 56
3.2.2. Shared Responsibility 58
3.3. Empirical Bioethics 59
3.3.1. Choosing A Methodology within Empirical Bioethics 62
3.3.2. Operationalizing My Methodology 63
3.3.3. Veridical, Realist, and Pragmatic Conditions 66
3.4. My Methodology 66
3.4.1. Mapping Review 66
3.4.1.1. What is a mapping review? 67
3.4.2. Systems Analysis 68
3.4.2.1. What is the Systems Approach? 69
3.4.2.2. The Systems Approach and Empirical Bioethics 70
3.4.2.3. Limitations of the Systems Approach 72
3.4.3. Qualitative Interviews 73
3.4.3.1. The Value of Qualitative Interviews 73
3.4.3.2. Interview Development 74
3.5. Conclusion 75

Chapter 4. Mapping EdTech Companies 77
4.1 Methodology - Operational 79
4.1.1. Identifying the Research Question 79
4.1.2. Identifying Relevant Technology 79
4.1.3. Selecting Relevant Technology 80
4.1.4. Collating/Summarising and Reporting Results 83
4.2. Results 85
4.3. Discussion 88
4.3.1. Shared Responsibility 88
4.3.2. Dual Harm, Singular Risk Scores, and Responsibility 89
4.3.3. Moderating, Utility, and Responsibility 91
4.4. Conclusion 94
Chapter 5. Systems Approach: Macro-System Analysis 97
5.1. Application of Systems Theory 97
5.2. Models 98
5.3. Search Strategy 102
5.4. Findings 104
5.4.1. Non-Statutory Advice 104
5.4.2. Professional Standards 105
5.4.3. Statutory Regulation 107
5.5. Analysis 109
5.5.1. Shared Responsibility 110
5.5.2. Hypothesis One: Divergent Roles / Theories of Education 112
5.5.3. Hypothesis Two: Neoliberalism and the Move to Accountability 114
5.6. Conclusion 114
Chapter 6. Interviews 118
6.1. Methodology 120
6.1.1. Recruitment 120
6.1.2. Additional Demographic Information 123
6.1.3. Interview Sessions 124

6.1.4. Data Analysis ... 127
6.2. Results ... 130
6.2.1. Responsible for Identification, Referral, and/or Prevention ... 132
6.2.1.1. Identification ... 133
6.2.1.2. Working Together (Through Referrals) ... 134
6.2.1.3. Working Together (Beyond Referrals, Into Prevention) ... 134
6.2.2. Why are there differences? ... 135
6.2.2.1. Workload and Prioritization ... 137
6.2.2.2. Emotional Availability ... 138
6.2.2.3. Job Title ... 139
6.2.2.4. School Ethos ... 140
6.3. Conclusion ... 141
6.3.1. Limitations ... 142
Chapter 7. Shared Responsibility ... 144
7.1. Whole School and Multi-Agency Approaches ... 146
7.1.1. Differences / Alignment ... 148
7.1.1.1. Exclusion of Key Stakeholders ... 149
7.1.1.2. Lack of Commitment towards Shared Responsibility ... 150
7.1.1.3. Lack of Clear Roles ... 151
7.1.1.4. Blurring of Geographic Boundaries ... 155
7.2. Ethical Challenges ... 156
7.2.1. Medicalisation of the Classroom ... 157
7.2.2. Conflicts of Interest ... 158
7.2.3. Single Data Controller ... 160
7.3. Increasingly Autonomous Systems ... 162
7.3.1. Benefits ... 164
7.3.2. Growing Responsibility Gap ... 165
7.3.3. Complicated Processes of Data Sharing ... 167
7.4. Conclusion ... 168
Chapter 8. Conclusion ... 170

8.1. Contribution to Academic Field ... 170
8.2. Recommendations ... 171
8.2.1. Recommendation 1: Referral, Not Intervention, and Shared, Not Individual ... 172
8.2.2. Recommendation 2: Overcoming Key Barriers to Shared Responsibility ... 173
8.2.3. Recommendation 3: Enhancing Ethical and Transparent Practice in EdTech Governance ... 174
8.2.4. Recommendation 4: Integrating Socio-Political and Historical Contexts into EdTech Evaluation and Policy ... 175
8.3. Limitations ... 177
8.3.1. Limitation 1: The Need for Evaluation ... 177
8.3.2. Limitation 2: Participant Sampling ... 178
8.3.3. Limitation 3: Online Interviewing ... 178
8.3.4. Limitation 4: Limited Systems Analysis ... 179
8.4. Future Research ... 179
8.4.1. Project Extensions ... 179
8.4.1.1. Extension 1: More Participants ... 179
8.4.1.2. Extension 2: Mapping National and School-Level Procedures ... 180
8.4.1.3. Extension 3: Accounting for Increasingly Autonomous Agents ... 181
8.4.2. Substantial Direction for Future Research 1: Trust ... 181
8.4.3. Substantial Direction for Future Research 2: Abolition ... 182
Bibliography ... 184
Appendix ... 212
Appendix One: Available EdTech Software to Monitor for Suicide Risk ... 212
Appendix Two: CUREC Approval ... 218
Appendix Three: CUREC Approved Consent Form ... 219
Appendix Four: Interview Guide ... 223
Appendix Five: Demographic Questionnaire ... 230

Abstract

In line with a national push towards integrating artificial intelligence and technology-based approaches in mental health, and new legislation requiring schools to have monitoring systems in place on all computers, many UK schools have begun to invest in new, relatively low cost, mental health monitoring software. This software uses natural language processing to scan through hundreds of gigabytes of naturally occurring digital data, which are then used as a source of information about students' mental health, for instance, monitoring suicide risk.

It is critical to understand how this new technology may impact the responsibility of different parties (e.g. psychologist, teacher, developer, child) in responding to a student who is considered 'at-risk' of suicide. For example, teachers already have a formalised duty of care and custodianship towards their students' wellbeing. Do these new monitoring programs change the nature of responsibility? This line of enquiry led to the development of my core (primary and secondary) research questions:

1. What responsibilities do teachers and schools have when using technology for suicide prediction, and how do these compare to what their responsibilities should be?
2. Given that teachers do not work in isolation (and that children are embedded within multiple overlapping support systems), how might a model of shared responsibility (involving teachers, clinicians, parents, technology developers, and/or the students themselves) function? How should responsibility be shared (as indicated by legal, policy, and ethical frameworks), and is it being done so in practice?

To answer these questions, I first explore the ethical foundations on which school-based suicide-prediction tools are both challenged and defended (Chapter 2). I then outline a novel bioethics methodology (Chapter 3), focusing on the theme of responsibility and consisting of three key components: a mapping review (Chapter 4); systems analysis (Chapter 5); and qualitative interviews (Chapter 6).

Within Chapter 4 I map the field of 'EdTech for suicide prediction,' analysing nine companies providing these tools in UK schools, and collecting data on the following: technical features of the

tools offered; number of schools subscribed to these tools; available efficacy data; and how responsibility is embedded in the software's functionality and promotional materials.

In Chapter 5, I explore the concepts of individual and shared responsibility in greater depth by using Systems Analysis methodology. This approach enables me to identify the key actors in the field and examine the regulatory, ethical, and best practice documents that inform how teaching and policy communities define individual and shared responsibilities for young people's mental health and the use of EdTech for suicide prediction. These documents include (but are not limited to), *Keeping Children Safe in Education, PREVENT*, and *the English Standard Framework of Teaching*.

In Chapter 6, I introduce a qualitative study exploring teachers' values and preferences regarding EdTech for suicide prediction. This chapter examines how teachers understand their legal and statutory responsibilities for identifying and referring students at risk, while also highlighting variations in teachers' interpretations of responsibility. For example, while some teachers view their role as strictly limited to referrals, others adopt an 'enhanced' approach, engaging in prevention efforts such as Personal, Social, Health and Economic (PSHE) curriculum development or school awareness campaigns. This division is shaped by factors such as workload, emotional availability, job expectations, and school ethos, with implications for both student wellbeing and teacher burnout.

I conclude this DPhil with two chapters dedicated to understanding shared responsibility within the context of EdTech for suicide prediction. Chapter 7 synthesises insights from the mapping review, policy analysis, and empirical research through an empirical bioethics framework. It explores gaps between current multi-agency approaches to suicide prediction and normative ethical ideals of 'shared responsibility,' including the exclusion of key stakeholders, the absence of a shared commitment, unclear role definitions, and the blurring of geographic boundaries. Considering these gaps, Chapter 8 concludes the thesis by offering concrete recommendations for policymakers and other stakeholders on how responsibility 'should' be held (and shared) by teachers, critically reviewing the DPhil as a whole, and suggesting future research directions to support ethical and effective implementation.

List of Figures

Figure 1: Selections from the Impero Website……………………………….………..………….....28
Figure 2: 5 Years of Backlash Against EdTech Tools: Flashpoint Events from the US and UK.....36
Figure 3: Explanation of Methodology……………………………………………………….….….62
Figure 4: Case Study One, Schools Using Ativion …………………………….……………..…..79
Figure 5: Adapted PRISMA Diagram for Identification of EdTech Providers…………….…………83
Figure 6: Description of Data Flow…………..………………………………………….……….87
Figure 7: Adapted Version of Bronfenbrenner's Ecological Theory……………………….………102
Figure 8: Policy Environment and Macro-System Dictating Teachers' Roles and Responsibilities for Suicide Prediction and Mental Health Monitoring in UK Schools…………………….………105
Figure 9: Relevant Sections of the English Framework of Standards (2013)...............................107
Figure 10: Complex Educational Systems Analysis…………………..……………..……..………117
Figure 11: Coding Scheme, What are Teachers' Roles and Responsibilities………..........….……131
Figure 12: Coding Scheme, What are Teachers' Reasons for Responsibility……….………..…...131
Figure 13: Flowchart Illustrating What Contextual Factors Shape Teachers' Responsibility Practices……………………………………………………………………………………...132
Figure 14: Software Companies as Gatekeepers of Student Data……………...…………………162

List of Tables

Table 1: Push and Pull Factors for ML-Based Suicide Prediction……………………………………34
Table 2: Students' Concerns About School Surveillance…………………………………………47
Table 3: Primary and Secondary Research Questions ……………………………………………56
Table 4: Methods Overview…………………………………………………………………………66
Table 5: Keywords for Searching ……………………………………………………………………81
Table 6: Inclusion and Exclusion Criteria …………………………………………………………82
Table 7: Keywords and Search Strategy …………………………………………………………104
Table 8: Inclusion and Exclusion Criteria …………………………………………………………121
Table 9: Abbreviated Participant Summary…………………………………………………………122
Table 10: Condensed Interview Guide……………………………………………………………127
Table 11: Types of Responsibility and Factors Shaping Teachers' Role Perceptions……………133
Table 12: Whole School v Multi-Agency Approach to Responsibility……………………………148

Abbreviations

ACLU - American Civil Liberties Union
AI – Artificial Intelligence
AP – Alternative Provision
CAMHS - Child and Adolescent Mental Health Services
CDT - Centre for Data and Technology
CESA - Complex Educational Systems Analysis
COVID-19 - Coronavirus Disease 2019
C-SSRS - Columbia Suicide Screening Severity Scale
CSS – Columbia Suicide Screen
CYP - Children and Young People
DCS – Director of Child Services
DSL - Designated Safeguarding Lead
EdTech - Education Technology
EFF - Electronic Frontier Foundation
EMHN - Emerging Mental Health Need
GDPR - General Data Protection Regulation
KCSIE - Keeping Children Safe in Education
LTP – Local Transformation Plan
MH - Mental Health
ML - Machine Learning
NHS - National Health Service
NICE – National Institute for Health and Care Excellence
NLP - Natural Language Processing
NSSI - Non-Suicidal Self-Injurious (behaviour)
Ofsted - Office for Standards in Education, Children's Services, and Skills
ONS - Office for National Statistics
PPV - Positive Predictive Value
PSHE - Personal, Social, Health and Economic (education)
PCPCH - Royal College of Paediatrics and Child Health
SML - Supervised Machine Learning
SSI - Suicidal Self-Injurious (behaviour)
TAC – Team Around the Child
UML - Unsupervised Machine Learning
WHO – World Health Organization
YPAG - Young People's Advisory Group

Notes on Terms

Doing interdisciplinary work often means working across fields that lack a shared vocabulary. For instance, while my DPhil is based in Psychiatry, I also draw upon other literature, including Bioethics and Social Data Sciences.

In this short section I define some of my most used terms and create a shared vocabulary for evaluating mental health technology, as well as suicide more generally. Establishing a ‘note on terms’ section at the beginning of the DPhil helps clarify points of consensus and highlight the limitations within and across these fields.

Suicide Ideation and Behaviour

It is important to distinguish between suicidal ideation, planning, and behaviour early in the DPhil, as traditional screening tools, such as the Columbia Suicide Screen (CSS) and Beck’s Scale for Suicide Ideation, explicitly assess differences among these three categories (Harmer et al., 2024).

To measure suicidal ideation, scales typically include the frequency, intensity, and quality of suicidal thoughts (Beck et al., 1979; Harmer et al., 2024), ‘plan,’ refers to whether the adolescent who reports suicidal ideation has formulated a specific/detailed plan (Harmer et al., 2024). A plan would include a method and anticipated outcome (Harmer et al., 2024). The third category, behaviour, includes whether the adolescent acts on their suicidal thoughts.

While clinicians and traditional tools (e.g., the CSS and Beck’s Scale) clearly distinguish between ideation, planning, and behaviour, some predictive technologies conflate these categories and provide adolescents with a generic risk score without differentiating between suicidal ideation and behaviour. This DPhil therefore explicitly acknowledges the distinctions among these three categories, while also engaging critically with technological literature that addresses suicidal ideation and behaviour collectively.

Suicidal Self-Injurious Behaviour and Non-Suicidal Self-Injurious Behaviour

Likewise, the Psychiatry literature also distinguishes between suicidal self-injurious behaviour (SSI) and non-suicidal self-injurious behaviour (NSSI), while the Education Technology (EdTech) itself tends to conflate the two categories (authors who explain the difference between SSI and NSSI include: Grandclerc et al., 2016; Kapur et al., 2013; Orlando et al., 2015).

Explicitly distinguishing between the two behaviours, Orlando et al. (2015) explain that NSSI includes actions such as cutting, burning, or hitting oneself, carried out without suicidal ideation, intent, or planning, while SSI may (or may not) include the above actions, but with added, explicit suicidal ideation, planning, and behaviour (defined more in the previous section). It is critical to note that a large risk of NSSI is that it evolves into suicidal behaviour (Grandclerc et al., 2016). In addition, these behaviours are often comorbid (Grandclerc et al., 2016), and some psychiatrists, such as Kapur et al. (2013) argue that distinguishing between the two behaviours has limited clinical utility.

Risk

It is important to define the term 'risk' at the outset of this DPhil. Psychologists, technology experts, statisticians, and philosophers all use the term 'risk,' though often in distinct and context-specific ways. For instance, prior to the 1960s, individuals were often categorised in binary terms (according to Rose (2010), either 'dangerous' or not), which led to straightforward, often punitive interventions for those deemed inherently 'high-risk' (Castel, 1991; Rose, 2010). This binary framework began to change in the 1970s and 1980s, influenced by mental health advocates who emphasised that individuals' likelihood to display mental health symptoms was shaped by a range of situational and contextual factors (Rose, 2010). As a result, risk began to be viewed less as an inherent trait and more as a probabilistic estimate (Rose, 2010).

For the purposes of this DPhil, risk will be defined as a likelihood score, typically expressed as a numerical odds ratio or probability (Calman & Royston, 1997). In the context of suicide prediction, risk will refer to a probabilistic estimate of how likely suicide is to occur, based on characteristics known or believed to be associated with an increased risk of SSI in adolescents[1]. This assessment will also take into account the protective factors which may reduce the likelihood of suicide. Ultimately, a risk estimation involves balancing these risk and protective elements, while acknowledging the inherent uncertainty and unpredictability involved.[2]

Note: It is important to note that, while psychological assessments typically specify the outcome a risk score is intended to measure (e.g., suicide), this DPhil shows that EdTech-generated risk scores often conflate distinct behaviours (e.g. SSI and NSSI). This can result in overly broad classifications, with students being labelled generically as either 'high risk' or 'low risk.' Furthermore, this DPhil demonstrates that EdTech-generated risk scores often overlook background and protective factors, relying instead on a narrow range of digital data focused on immediate risk.

Prediction

Clinicians use patient-specific information to hypothesise various future health outcomes: from prognosis, diagnosis, and treatment (Efthimiou et al., 2024). Typically, the results of these models are presented as risk scores, which are the scores that quantify the likelihood of specific future outcomes (defined in more detail above). The clinical goal of prediction is early prevention or treatment, i.e. the mitigation of negative outcomes (Efthimiou et al., 2024).

Screening

According to the UK's National Health Service (NHS) website on screening (NHS, 2024, para. 1), "[s]creening is a way of finding out if people have a higher chance of having a health problem, so that early treatment can be offered or information given to help them make informed decisions." Screenings often include a prediction of risk (mentioned above) and occur on a discrete basis (e.g.

---

[1] For example, a systematic review by Nordin et al. (2022) identified several categories of inputs for suicide prediction models, including: demographic information, military characteristics, family history of psychopathology, physical health or illness, treatment history, internalizing and externalizing psychopathology, social factors, prior suicidal behaviors, and cognitive abilities. Other studies have also incorporated social data inputs, such as online behaviour (e.g. on Facebook; Marks (2019)).

[2] To acknowledge the inherent uncertainty and unpredictability involved, statisticians will often include other metrics with a suicide risk prediction, e.g., a sensitivity discrimination score, discrimination metrics, and positive prediction value (Large, 2018).

once a year at the GP or in a school classroom), unlike monitoring, which happens continuously. An example of a screening program is the CSS.

Monitoring

In this DPhil, monitoring is defined in contrast to screening. While screening typically involves a single, discrete assessment of risk (defined above), monitoring refers to ongoing evaluations conducted over time. These ongoing evaluations/risk assessments can happen either continuously (e.g., 24/7) or at regular intervals. In addition, monitoring can be targeted towards specific, at-risk groups (e.g., patients recently discharged from hospital or military veterans; Nelson et al. (2017)); or else more universally (e.g. large groups of the public; Moura et al. (2020)).

Ultimately, screening and monitoring represent two distinct approaches to generating predictions or risk scores.

Machine Learning

While earlier prediction models often relied on limited indicators such as a single high-risk trait or behaviour, contemporary approaches are built from larger, more diverse datasets and often employ complex methodologies. This includes Supervised Machine Learning (SML) or Unsupervised Machine Learning (UML).

In SML, data comes with pre-attached labels, selected by the algorithm designers, that are categorical or continuous. For example, one can either have a diagnosis of major depressive disorder or not (categorical data; Graham et al. (2019)). The same person could also be along the spectrum of depressive symptoms, therefore be located on a continuum (examples also found in Graham et al. (2019)). In SML, data is taken from a variety of sources, including demographic and clinical measures, and this is used to predict someone's best suited category/label (Graham et al., 2019). The algorithm is supervised first as the labels 'teach' the algorithm how to associate specific data points with one of the prescribed labels (Graham et al., 2019). This is done initially with a training dataset (to make sure the model is correct in its classifications) before being used with unlabelled and independent data (Graham et al., 2019).

While in SML, data comes with pre-attached labels and cut-off points (e.g. certain threshold scores which mean a patient has depression), in UML, there are no labels, and therefore the algorithm does not have cut off points already coded into it (Graham et al., 2019). Instead, the algorithm works to associate input features and consider the underlying similarities between data points. An example of UML is when, “neuroimaging biomarkers provide large feature datasets that may hold information regarding unknown subtypes of psychiatric illnesses like schizophrenia” (Graham et al., 2019, p.4).

Most EdTech tools use a combination of Natural Language Processing (NLP) and/or Digital Phenotyping as a specific SML/UML method. NLP focuses on how computers process and analyse unstructured text and human languages (Khurana et al., 2023). It uses the above SML and UML methods and relies on semantic understanding as well as information extraction (Khurana et al., 2023). Digital Phenotyping, on the other hand, is a more complicated process, first coined by Tom Insel and colleagues (Insel, 2017). This includes continuous, passive data collection from sensors (tracking activity/location), voice and speech (exploring prosody/sentiment), and human computer interactions (e.g. typing, scrolling) (Insel, 2017). These data sources are then used as proxies for cognition, mood, and behaviour. For example, changes in activity/location can be an early sign of depression (Insel, 2017; Jacobson et al., 2019) and changes in pronoun use can be an early sign of depression (moving to using first person singular) (Insel, 2017; Jacobson et al., 2019).

EdTech

According to Williamson (2021), Education Technology (otherwise known as EdTech),

> “has become an increasingly capacious category. It designates a huge variety of actors (human and nonhuman), organizations (public, private or multisector), material and technical forms (hardware, software, supporting documents), modes of practice (of teachers, designers, promoters), and framing discourses, as well as being a highly varied field of research, development and critical inquiry.” (p.1)

Williamson’s (2021) work shows that EdTech is used as umbrella term within the literature, encompassing various tools, practices within educational settings. This includes technologies used for a broad spectrum of purposes, including (but not limited to): attendance monitoring, assignment submission, communication with parents, and safeguarding and mental health support. Even within these broad categories, the range of technologies remains extensive. For instance, *EdTech for mental health* can refer to tools used for risk monitoring (across a range of different mental health conditions), as well as those designed for therapeutic intervention, such as chatbot therapists (Chan et al., 2025). In this DPhil, *EdTech for suicide prediction* specifically refers to software employed to identify or monitor suicide risk, thereby excluding intervention-oriented tools such as chatbot therapists. Chapter 4 will provide a more detailed explanation of the features of these tools. However, in general, they operate by keyword searching students' computer activity to identify terms associated with risk.

Prevention

In the context of suicide and schools, prevention refers to the strategies, programs, and practices aimed at reducing the risk of suicidal thoughts and behaviours among students (Large, 2018). According to the World Health Organization (WHO)'s LIVE LIFE implementation guide for suicide prevention (WHO, 2021), prevention efforts can be grouped into four main categories: "Limiting access to the means of suicide; Interacting with the media to promote responsible reporting of suicide; Fostering socio-emotional life skills in adolescents; Early identification, assessment, management, and follow-up of individuals affected by suicidal behaviours" (p. xi).

From this definition, it is clear that prevention serves as an umbrella term encompassing a wide range of interventions. For example, in the UK, the Department for Education (DfE) published new guidance in 2025 that requires suicide prevention to be taught as part of the school curriculum (DfE, 2025). Other preventive approaches may include behaviour management, anti-bullying initiatives, access to counselling, and referral pathways to Child and Adolescent Mental Health Services (CAMHS).

Within this thesis, I focus on a specific aspect of prevention, namely the early identification and prediction of individuals at risk, rather than prevention as a whole. However, while my focus is on the use of EdTech to detect early warning signs, risk indicators, and behavioural patterns that may suggest a student is in distress, I acknowledge that, within schools, prediction is typically undertaken as part of broader suicide prevention initiatives rather than as a standalone effort.

COVID Statement

This DPhil began in 2019, just before the COVID-19 pandemic. The legacy of COVID-19 on teachers, schools, and student mental health has been profound, and has impacted this thesis in several ways. It changed how technology is used in education, with digital tools being adopted more quickly. It also affected how the research was carried out. For example, interviews had to move online, teachers were under increased pressure due to rapidly changing expectations and workloads, and it became harder to recruit teaching staff for the study. Finally, people's views on mental health and technology shifted, with greater awareness of both the benefits and limitations of digital tools used within the classroom. These changes (contextual, practical, and ethical) are explicitly woven throughout this thesis and form a critical backdrop to the research presented.

Chapter 1. Introduction

Suicide is a leading cause of death among British adolescents. According to the Office for National Statistics (ONS, 2024) and the Royal College of Paediatrics and Child Health (RCPCH, 2021), the three top causes of death in the 10-19 age group are: (1) accidents, (2) cancer, and (3) intentional self-harm. Adolescent suicide rates have also increased over the past decade (although there have been some anomalies, e.g. during the COVID-19 pandemic; Pathirathna et al. (2022)). In 2021, suicide rates reached a 20-year high amongst UK adolescents ages 15-19 (6.3 deaths per 100,000), and recent data from the ONS shows that this number is still high, at 5.4 deaths per 100,000 (ONS, 2024).

Because of the severity of these statistics, implementing effective prevention strategies is a top priority for public health professionals (Hughes et al., 2023; Mann et al., 2021). As outlined in the World Health Organization's *LIVE LIFE* guidance (WHO, 2021), prevention encompasses a range of activities, including "limiting access to the means of suicide; interacting with the media to promote responsible reporting of suicide; fostering socio-emotional life skills in adolescents; and the early identification, assessment, management, and follow-up of individuals affected by suicidal behaviours" (WHO, 2021, p. xi). Although the early identification of suicide risk represents only one aspect of prevention, it is widely regarded as a critical component to be embedded within broader suicide prevention measures (e.g., Large, 2018) and thus is the basis of this thesis.

A range of tools are currently in use to identify early signs of suicide, particularly in emergency room, primary care, and school settings (Horowitz et al., 2009). According to Horowitz et al. (2009), schools are one of the most promising settings for early identification and risk screening, as the school is the place where adolescents spend the majority of their time.

Schools use a variety of tools to predict students' suicide risk. These range from established clinical instruments like the Columbia Suicide Screen (CSS) and the Columbia-Suicide Severity Rating Scale (C-SSRS), to newer online tools that leverage digital data for more dynamic and continuous monitoring of students' mental health. Smoothwall is one example of a digital programme implemented in schools to help identify students at risk of suicide.

According to the UK Government's Directory of UK Safety Tech Providers (Department for Business & Trade & Department for Science, 2024), Smoothwall is used by approximately 40% of all schools. Self-titled as a 'Digital Safeguarding Solution,' Smoothwall uses Machine Learning (ML) and inductive methods to analyse individual students' online activity and generate a risk score indicating their likelihood of self-harm or suicide (Smoothwall, 2025).

1.1. Aim

The aim of this DPhil is to investigate the field of new, school-based digital monitoring tools for suicide prediction (such as Smoothwall), using an interdisciplinary methodology grounded in Empirical Bioethics. This project will explore the social, ethical, and political implications of using EdTech to predict suicide within UK schools, with a focus on the theme of 'responsibility.'

This DPhil approaches the theme of responsibility through the lens of role responsibility, i.e. the duties that teachers and other stakeholders hold by virtue of the roles or offices they occupy (Hart, 1968). This includes individual role responsibility, such as the responsibility a teacher, school counsellor, or safeguarding lead has to identify and act on signs of suicide risk, as well as shared role responsibility, which refers to the collective and distributed duties of schools, technology providers, policymakers, and other stakeholders involved in suicide prediction through EdTech tools (these definitions of responsibility are informed by Hart's (1968) work. I provide a full explanation of both individual and shared role responsibility in Chapter 3.2).

1.2. Structure

Chapter 1 sets the stage by tracing the history of suicide prediction and mental health monitoring in two parts. First, the chapter analyses the 'push' and 'pull' factors behind the recent emergence of big data and ML–based interventions for suicide prevention, particularly within UK secondary schools. Second, the chapter examines the basis for the growing public criticism of these tools. The aim of this chapter is to provide the reader with a broad overview of the emerging field of 'EdTech for Suicide Prediction.'

Chapter 2 sharpens the focus of my DPhil, establishing my core research question(s) regarding the use of these software programmes. I draw on Chapter 1's real-world cases to highlight key ethical challenges that have emerged in the field of 'EdTech for Suicide Prediction,' including concerns related to privacy, autonomy, justice[3], and responsibility. Of these, responsibility emerges as the most relevant ethical framing for understanding the use of such technologies. My core (primary and secondary) research questions are as follows:

- *What responsibilities do teachers and schools have when using technology for suicide prediction, and how do these compare to what their responsibilities should be?*
- *Given that teachers do not work in isolation (and that adolescents are embedded within multiple overlapping support systems), how might a model of shared responsibility (involving teachers, clinicians, parents, technology developers, and/or the students themselves) function? How should responsibility be shared (as indicated by legal, policy, and ethical frameworks), and is it being done so in practice?*

Chapter 3 describes the specific methodological approach I designed to address my key research questions, an approach which includes an analysis of the technology itself (Chapter 4), of the school system and legal policies that govern schools (Chapter 5), and interviews with teachers (Chapter 6). Within Chapter 3 I explain how each of these empirical methodologies functions independently, as well as the bioethics strategies I use to integrate their findings in order to address the normative dimension of my research. I draw on a broad range of data sources, in conjunction with philosophical theory, to develop a well-grounded analysis of responsibility in my final two concluding chapters using operational methods described by Ives et al. (2017) and McMillan & Hope (2008) (methodology outlined extensively in Chapter 3).

1.3. History of Suicide Prediction Technology

As outlined above, my introductory chapter presents the history of suicide prediction and monitoring for mental health and suggests reasons why there has been an emergence of big data

---

[3] Mapping onto fundamental bioethics principles (Beauchamp & Childress, 2013; Varkey, 2021).

and ML-based interventions for suicide, particularly within school settings. This section answers the sub-question: *how has the field of suicide prediction developed over the past decade within the UK, and what factors have led to an increasing reliance on school-based big data/ML screening tools?*

Beck, Kovacs, and Weissman developed the first Scale for Suicide Ideation in 1979. This tool quantified the intensity of suicide ideation in order to proactively identify people who were at risk of suicide. Since Beck et al. (1979) developed their scale, other clinical tools have been created, including: the Suicide Behavioural Questionnaire (Osman et al., 2001), Life Orientation Inventory (Range & Lewis, 1992), Reason for Living Inventory (Pirani et al., 2021), and the Columbia Suicide Screening Severity Scale (C-SSRS) (Posner et al., 2008)

According to the NHS Screening Website, "[s]creening is a way of finding out if people are at higher risk of a health problem, so that early treatment can be offered or information given to help them make informed decisions" (NHS, 2024, para. 1). In general, screenings tend to include a prediction of risk and occur on a discrete basis (e.g. once a year). Suicide risk screening programmes, like the C-SSRS, are often used in combination with preventative interventions and evidence-based responses, as part of a more comprehensive suicide strategy (Ayer et al., 2022).

While screening for suicide is common practice in schools, primary care, and emergency departments (Horowitz et al., 2009), recent studies have questioned the effectiveness of traditional screening tools for suicide (Chan et al., 2016; Large et al., 2016; Runeson et al., 2017; Velupillai Sumithra et al., 2019). For example, Large and colleagues (2016) have shown that the C-SSRS has a low Positive Predictive Value (PPV). Across 53 different samples, Large and colleagues (2016) calculated the PPV of suicide risk assessment tools to be 5.5%. This means that only 5.5% of those predicted 'at high risk' will then die by suicide (*within Large and colleagues' 2016 sample). Because of the above cited limitations around the PPV, and clinical efficacy of these tools, organisations such as the National Institute for Health and Care Excellence (NICE) have warned clinicians not to use traditional assessment tools and scales to predict suicide (NICE, 2022).

Academic researchers and clinicians are therefore seeking better ways to prevent suicide. Some argue for moving away from risk assessment altogether and instead focusing on other aspects of suicide prevention. For example, Hawton and colleagues (2022) advocate shifting from risk prediction to a focus on therapeutic assessment, formulation, and risk management. Within schools, Hawton and colleagues' approach would involve teachers investing time in building a therapeutic alliance with students, using that relationship to identify unmet needs, and collaboratively developing care plans. However, this approach is time-intensive, and schools often lack access to trained therapists who can help "establish a therapeutic and empathetic rapport" with students and co-produce safety plans (Hawton et al., 2022, p. 924).

Hawton and colleagues (2022) are still in a minority of researchers in suggesting that the way forward lies in moving away from prediction toward more holistic (yet time-intensive) approaches. Instead, many others take a more positive view of risk prediction and are looking to develop 'better' screening tools (e.g. risk screening tools with higher PPVs). One popular avenue for improvement is with the use of big data and ML.

A hospital emergency department in Spain was one of the first institutions to experiment with data mining and ML techniques for suicide screening in the 1990s (Baca-García et al., 2006). In their study, Baca-García and colleagues (2006) explored whether clinical and demographic data could be analysed using automated methods to detect suicide risk patterns. Their findings suggested that data-driven models could complement traditional clinical judgment by highlighting patients at elevated risk. More specifically, within their results, Baca-Garcia and colleagues (2006) show how traditional statistical techniques classified 72% to 88% of patients correctly, while their 'data-mined' model classified 99% of patients correctly (n = 509 patients).

By 2017, hospitals across the world had created automated suicide risk assessment tools from statistical models. For example, in 2019 Velupillai Sumithra et al. (2019) surveyed six different tools which used ML-techniques (also see: Ilgen et al., 2009; Poulin et al., 2014). These programmes use data-driven approaches, free-text and natural language processing, primarily with patient information and electronic health records, to determine patient health and suicide-related outcomes (Barak-Corren et al., 2017).

The private sector has also been using big data and ML techniques for suicide screening, using social data in a process which Marks (2019, p.98) refers to as '"social suicide prediction." This includes data taken from social media, smartphones, and other sources on the Internet of Things (Marks, 2019). For example, in 2018 Facebook announced (and then quickly retracted) a suicide prediction and screening service using data collected on its platform (Kaste, 2018). Another example of non-medical data being used in suicide prediction is the data collected in schools and on student laptops and/or smartphones. Within this thesis I explore nine products available on the UK market which use non-clinical or non-traditional health data to monitor for suicide risk (Chapter 4).

1.3.1. 'Push' and 'Pull' Factors

By 2016, EdTech programs impacted more than "half a million students and staff in the UK" (Donkersley, 2016, para. 6), and this is continuing to rise. For instance, the Department for Education (DfE)'s 2022-23 Technology in Schools Survey said that 90% of teachers had used technology within their classrooms in the last year (IFF Research, 2023). Smoothwall alone was reported to be in over 40% of UK classrooms, according to the UK Government's 'Directory of UK Safety Tech Sector Providers' (Department for Business & Trade & Department for Science, 2024). While EdTech tools do not all have suicide prediction capabilities, many (like Smoothwall) do, and the number of these tools used in schools is likely to increase over the next decade.

The following section offers multiple hypotheses to explain the growing adoption of EdTech programs for suicide prediction in schools, and suggests that this trend is likely to continue. Contributing factors include: economic demand, increased government funding, the impact of COVID-19, and evolving regulatory frameworks– particularly with the PREVENT duty, as well as Keeping Children Safe in Education (KCSIE).

1.3.1.1. Repurposed Tools and PREVENT

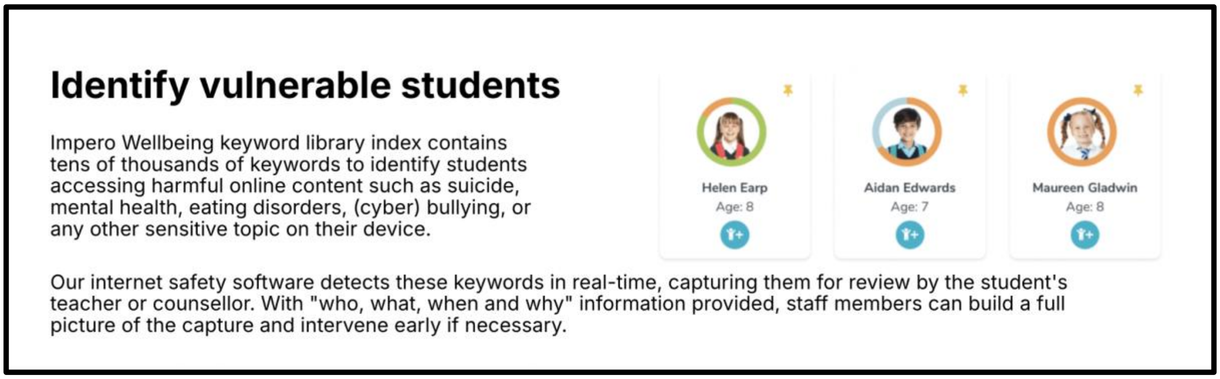
First, it is important to note that the EdTech tools currently used in schools for suicide prediction were not explicitly designed for suicide prediction, nor were they originally introduced for that purpose. Instead, many appear to have been repurposed from other applications, such as classroom monitoring and administrative functions. Notably, one of the primary origins of these tools lies in technologies initially developed to support counter-terrorism efforts under the PREVENT duty[4], and primary evidence suggests that, even today, these tools continue to serve both functions (Figure 1).

To illustrate how closely suicide prediction and PREVENT are intertwined within the tools currently used in schools, I have included selections from the Impero website below (Figure 1), which highlights the multi-functional nature of EdTech programs (e.g. explicitly advertising for safeguarding and suicide, support for the PREVENT duty, and compliance with KCSIE regulations).

[4] Since July 2015, all UK schools have been required to comply with the PREVENT duty under Section 26 of the Counter-Terrorism and Security Act. The PREVENT duty, updated in 2023, mandates that schools must pay "due regard to the need to prevent people from being drawn into terrorism" (Home Office, 2024b, Section 26).

Figure 1

Selections from the Impero Website (Impero, 2017b)



Under the PREVENT duty, the Home Office mandates that UK schools conduct online monitoring and recommends the implementation of comprehensive online risk assessment processes (Home Office, 2024b). It is therefore possible that seemingly unrelated initiatives (e.g. PREVENT), are inadvertently driving the development and deployment of suicide prediction tools, simply because both functions are integrated within the same technological systems used in schools. This is a theory sometimes referred to as ‘Function Creep,’ discussed later in this thesis (Hope, 2019; Koops, 2021; Persson, 2022).

Because EdTech programs often link suicide prediction with anti-terrorism monitoring (as illustrated in the example above), it is important to ask whether an increase in school-based anti-terrorism efforts is behind the rise in suicide prediction initiatives. Although it is unclear whether PREVENT referrals are specifically coming from the use of EdTech programmes, available data suggest that overall, there is a significant and growing number of PREVENT referrals coming from the education sector. According to UK government data, “the Education sector made the highest number of referrals (2,788), accounting for 40% of all referrals this year; this is similar to last year and this is the highest proportion for any source of referral since data was first published in 2015 to 2016” (Home Office, 2024, para. 6).

Whether or not PREVENT is an incidental driver of suicide prediction tools, the rise in school referrals is worth highlighting, because many ‘high risk’ referrals of students from EdTech tools may be originating from systems designed to flag both suicide risk and PREVENT-related concerns.

While this thesis does not aim to conclusively prove a direct link between suicide prediction tools and PREVENT, this section shows that it is unlikely that the rise of EdTech tools is driven solely by educational or clinical needs. Instead, broader policy pressures (like PREVENT) also contribute to their expansion, and must be considered in developing my methodology (Chapter 3) and subsequent ethical analysis (Chapters 4-8).

1.3.1.2. Additional Funding and Regulation ‘Pulls’

Beyond the need for improved and more accurate suicide prediction tools (outlined in Section 1.2) and PREVENT duty guidance (outlined in Section 1.3.1.1), another policy pressure contributing to the increased use of EdTech tools in schools is the growing governmental emphasis on data sharing within the broader contexts of safeguarding and healthcare. For example, by 2018, the

Working Together to Safeguard Children Act specifically mentioned data collection and sharing across public sectors (e.g. schools, police, hospitals) for the purposes of general safeguarding (HM Government, 2023). The growing emphasis on data collection and sharing is likely to drive the introduction of more digital monitoring programmes (such as Impero; Figure 1), and lead to EdTech tools becoming increasingly complex and integrated with broader data-sharing platforms.

In addition to data sharing, the government has also funded initiatives (and promoted regulation) for the purposes of safeguarding and healthcare, further encouraging the development of digital tools for suicide prediction. For instance, in 2021 the UK government pledged £140 million towards Artificial Intelligence (AI) innovation in healthcare (in 2023 this number was lowered to £123 million (NIHR, 2021, 2023)). More recently, in September 2024, KCSIE includes mention of internet and filtering technologies embedded on school devices for the purpose of keeping students safe. KCSIE also includes reference to the DfE's Filtering and Monitoring Standards (Department for Education, 2020), which links internet usage and online data to the safeguarding needs of the school. Finally, in 2023, the UK Government's Suicide Prevention Strategy acknowledged the significance of AI and called for further exploration of its potential benefits and risks in suicide prediction (Department of Health and Social Care, 2023).

It is likely that both regulation and funding directed toward data-sharing and ML programmes are directly contributing to the rise of EdTech tools in UK schools, and in subsequent chapters of this thesis, the implications of additional funding and regulatory 'pulls' will be explored in greater depth, particularly in terms of how they shape the socio-political environments in which these tools are promoted or critiqued (e.g., in Chapter 5).

1.3.1.3. COVID-19

Beyond clinical need and government policies (discussed above), a novel coronavirus (COVID-19) has had a direct impact on the use of student data for suicide monitoring. In 2019/2020 COVID-19 unexpectedly caused a fundamental shift in schooling, making the uptake of, and need for, EdTech programs spike dramatically. By mid-April 2020, schools were closed in 190 countries, and schools resorted to developing remote learning strategies, predominantly taking place online (UNESCO, 2021). COVID-19 necessitated additional online programmes, not only for teaching, but to also promote other school-related activities, including safeguarding, mental health services, and monitoring (West, 2023). This includes the UK Government providing funding for the DfE's EdTech Demonstrator Programme, which was a group of 48 schools committed to providing long-term peer-to-peer instruction on using EdTech in practice (the programme was developed in 2020, but withdrawn in July 2022) (Department for Education, 2022).

While the EdTech demonstrator programme has ended, the impact of this period of remote learning is likely to be long-lasting, due to schools purchasing new technology on subscription, and the significant financial and infrastructural investments made by the government. Many believe that these emergency relief solutions may lead to long-term reform and 'digitally transform' the education sector (Williamson et al., 2020). For example, the UK government proposed a COVID-19 relief act in June 2020, consisting of a £1 billion education 'catch-up plan.' With this in mind, it is clear that COVID-19 and the subsequent increase in funding for digital tools in schools have directly contributed to the rise of EdTech in the UK. While the funding may not have been specifically earmarked for suicide prediction, as previously discussed, many tools serve dual functions. Therefore, it is reasonable to hypothesise that this influx of funding also supported the development (and/or adoption) of tools with the capacity for suicide prediction.

1.3.1.4. CAMHS Funding Constraints

A fourth socio-political factor that may have inadvertently contributed to the rise of EdTech tools in the UK market is the reduction in funding for traditional mental health services and more general governmental policies of austerity. According to the Local Government Association (2023, box 3, para. 4), "Councils have seen a £770m real terms reduction in [public health] funding between 2014/15 and 2020/21 – a fall of almost a quarter (22.3 per cent) per person."

This lack of funds means that schools often rely more heavily on alternative forms of mental health support to make up for the lack of traditional mental health services. Indeed, 80% of primary heads and 72% of secondary heads say their schools changed the way they meet students' social and emotional needs due to financial constraints (Ofsted, 2020).

Digital tools have emerged as key alternatives to traditional mental health services and are often used to address gaps in provision. These include a wide range of prevention-oriented interventions under the umbrella of digital mental health, such as early risk prediction, therapeutic support, and crisis intervention. Their growing appeal lies partly in their potential to reduce costs. For example, Coote et al. (2024)'s economic modelling and evaluation report provides growing evidence that (some) digital mental health tools can deliver cost savings for the NHS. This includes preliminary evidence of cost savings coming from online mental health platforms such as Kooth and Healios, which have been adopted by NHS trusts to deliver web-based counselling and video consultations with therapists (Cootes et al., 2024). As Coote et al. (2024, p.5) note, "the early economic model demonstrates a potential cost saving to the UK NHS of £236.15 per person with an EMHN [Emerging Mental Health Need] over a 1-year time horizon. This increases to £246.54 when the NHS + crime sector perspective is taken." (p. 5). While this example focuses on prevention in the context of virtual therapy, a range of cost-reducing preventive tools could also include those designed for prediction, such as the EdTech approaches explored in this thesis.

While the clinical efficacy of digital mental health tools will be examined later in this thesis (e.g. in Chapter 2), it is evident that such tools are increasingly being considered in UK schools, largely because they are viewed as financially viable alternatives to traditional mental health services.

Overall, Section 1.3 traced the history of suicide prediction tools, showing how technologies initially developed for clinical use gradually became available in classrooms and commercial contexts. It also began to connect social, political, and historical factors (Table 1) to explain the motivation behind the tools' development and subsequent implementation in schools. Specifically, Section 1.3 has shown that a combination of (a) historically ineffective suicide prediction tools, combined with (b) new financial and legislative incentives for digital innovation (particularly within the school context), (c) a socio-economic climate of austerity, and (d) the COVID-19 pandemic, has led to the rise of digital suicide prediction tools, both within the UK and internationally (Table 1).

Table 1
Push and Pull Factors for ML-Based Suicide Prediction

Push	**Pull**
Traditional tools less effective and not recommended by NICE guidelines	New financial incentives from government
Lack of funding for traditional mental health services (e.g. funding cuts to CAMHS and local councils)	New legislation promoting growth and technological innovation
Lack of funding in schools more generally	New legislation promoting increased securitisation and monitoring of school-aged students (e.g. PREVENT).
	COVID-19 and a need for more robust digital services during the pandemic

1.4. State of Play

While Section 1.3 focused on the socio-political environment that contributed to the adoption and development of EdTech tools, Section 1.4. shifts to focus on what happened *after* EdTech tools were introduced in schools, examining how the past decade has shaped their current use and public perception. This section will introduce several key political and social flashpoints from the past five years that highlight the challenges of using EdTech for suicide prediction and help to explain both the public backlash and the decision by many schools to cancel their contracts with EdTech providers.

Like Section 1.3, Section 1.4 aims to connect social, political, and historical factors to the emergence, adoption, and criticism, of EdTech tools in UK classrooms. The goal of Section 1.4. is to introduce the current 'state of play,' which will inform and help refine my overall research questions (grounded in empirical bioethics, and presented in Chapter 2 and 3).

1.4.1. Timeline

Over the past five years, there has been significant backlash against the use of EdTech tools for suicide prediction, with education communities in both the USA and the UK raising concerns and, in some cases, cancelling their contracts with EdTech providers (Figure 2). Specifically, schools have ended their use of these tools in response to pressure from policymakers and advocacy groups operating at local, national, and international levels.

While not conclusive, Figure 2 below presents a visual timeline of key dates and cases illustrating this backlash, helping to introduce, contextualise, and in many ways, complicate, the current 'state of play' for EdTech in the UK.

It is important to note that, although this DPhil focuses on evaluating the use of EdTech tools in the UK, Section 1.4 also covers international examples, including letters written by US Democratic senators critiquing the use of EdTech in US schools (e.g., point 1 in Figure 2). Including international examples in this introductory chapter is essential for two key reasons: (1) most of the EdTech tools used in the UK are developed and used globally, and (2) the socio-political debate emerging from the US context offers valuable insight for the bioethical analysis I conduct later in this thesis (including the development of my research questions in Chapter 2).

Figure 2

5 Years of Backlash Against EdTech Tools: Flashpoint Events from the USA and UK

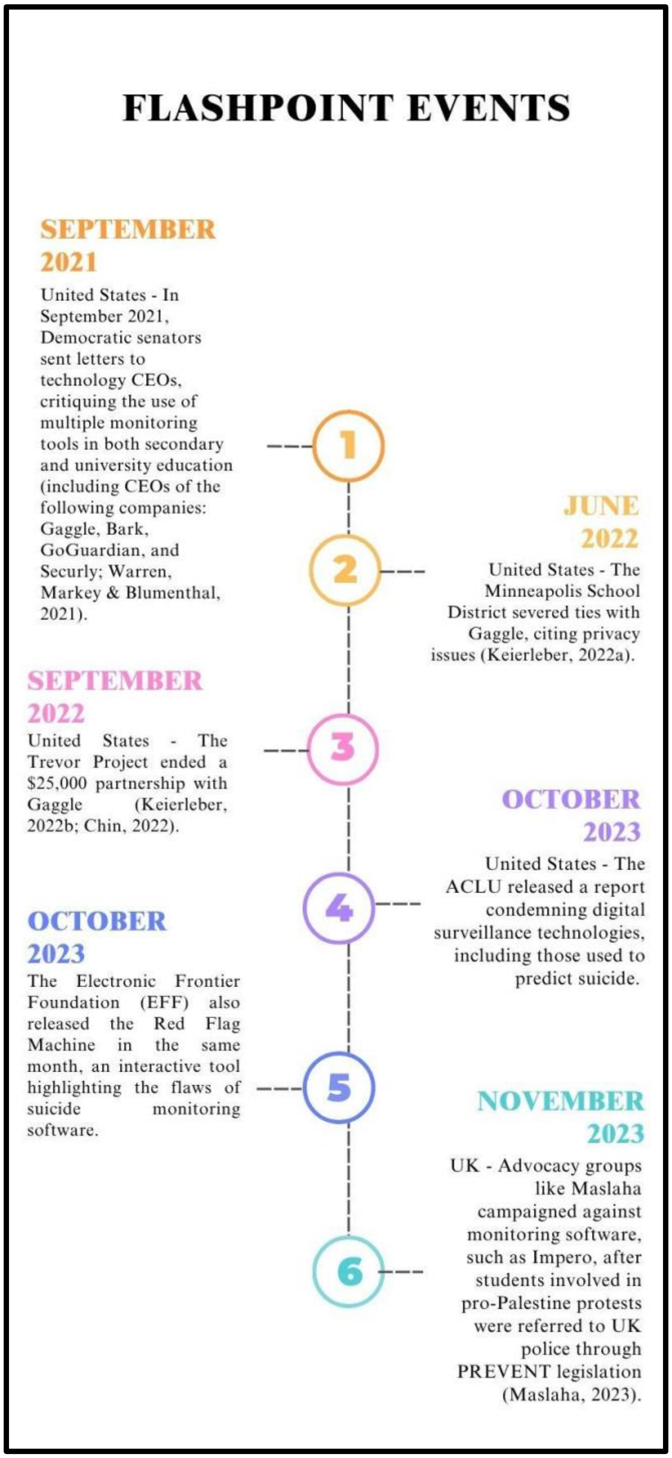


Figure 2 shows that, over the past five years, a growing backlash against EdTech tools has taken shape. In September 2021, US Democratic senators sent letters to major tech CEOs questioning the use of monitoring tools in schools and universities (Warren & Markey, 2022). In June 2022, the Minneapolis Public School District (US) ended its contract with Gaggle (Keierleber, 2022). A few months later, in September 2022, charity organization The Trevor Project severed ties with Gaggle after pushback from advocates over potential harm to LGBTQ+ youth (Chin, 2022).

Next, in October 2023, the American Civil Liberties Union (ACLU) released a report condemning school surveillance technologies (ACLU, 2023), and during the same month, the Education Frontier Foundation (EFF) issued warnings about the ethical risks of predictive EdTech tools (Maass et al., 2023). Finally, in November 2023, UK advocacy group Maslaha launched a campaign against Impero, a classroom monitoring tool, prompting some schools to reconsider or cancel their contracts (Maslaha, 2023).

While Figure 2 presents only a small selection of exemplar cases, it highlights a striking disparity. On one hand, as discussed in Section 1.3, there is a growing demand for suicide prediction tools, driven in part by increased funding, demand from the UK government, and need from organisations such as CAMHS. On the other hand, Figure 2 illustrates how, over the past five years, EdTech tools designed for suicide prediction have come under mounting public criticism and, in some cases, have been removed from schools altogether. This raises a critical question: on what grounds are these tools being challenged? What are the underlying concerns expressed by groups such as the EFF and ACLU, as well as by school districts and national policymakers?

Chapter 2 will develop an analysis to address these questions by examining the socio-political history of EdTech programmes through an ethical framework. It will critically engage with the academic literature around clinical utility, and the benefits of these technologies for mental health prediction, while also exploring ethical concerns such as privacy, autonomy, discrimination, and, most critically, the theme of responsibility.

Chapter 2. Developing the Research Scope

I ended Chapter 1 by describing ‘EdTech for suicide prediction’ as a field full of tension, shaped by growing technological and government demand, as well as by mounting criticism, with schools both adopting and discarding these tools at rapid rates.

Chapter 1 was primarily descriptive and did not explore the underlying cause(s) of this tension. Chapter 2 explores the ethical foundations on which these tools are both challenged and defended by drawing on Beauchamp & Childress (2013)’s four principles approach.

In Section 2.1, I study the ethical arguments in favour of using digital suicide prediction tools, focusing particularly on the principle of beneficence. I begin by defining beneficence and its two main dimensions: clinical utility and positive beneficence (Beauchamp & Childress, 2001). I use examples from both peer-reviewed studies and grey literature, including online materials, an earlier PhD project (Shelton, 2022), and a scoping report by the think-tank RAND Corporation (Ayers et al., 2024), to examine how EdTech tools for suicide prediction demonstrate, or fail to demonstrate, clinical utility and positive beneficence. This includes data on clinical utility, analytic validity, and clinical validity. I conclude Section 2.1 by noting that current data on the use of EdTech in this context is inconclusive. Therefore, substantially more research is needed before a clear argument for beneficence can be convincingly made.

In Section 2.2, I explore the broader ethical critiques and trade-offs highlighted in the ethics literature. These are often contrasted with the principle of beneficence discussed in Section 2.1. To do this, I use materials from both peer-reviewed research and grey literature, including civil society reports and news articles, and focus on two key ethical critiques: autonomy and justice. This section provides a brief overview of the ethical dilemmas arising from both the tools themselves, as well as the socio-political climate in which they are used.

In Section 2.3, I explain that, rather than conducting an in-depth analysis of each ethical principle to study EdTech for suicide prediction (Beauchamp & Childress, 2001), this DPhil instead adopts responsibility as the central ethical framework. I offer a brief rationale for my approach, noting that by framing my project around responsibility (both individual and shared), I align myself with both government policy (e.g. ‘Keeping Children Safe in Education’; KSCIE) and a sub-set of ethical literature that uses frameworks of responsibility to navigate competing ethical concerns (Turoldo, 2009). Ultimately, responsibility underpins both my research question and methodology and represents my unique contribution to the field.

2.1. Beneficence

This section considers the ethical implications of digital suicide prediction tools deployed in UK classrooms and addresses the following question(s): *Does the use of these tools by schools and teachers aligns with the principle of beneficence?* This question is answered before moving to Section 2.2, which asks whether any potential benefits justify the associated risks (such as privacy violations). This two-step approach aligns with comparable analyses of beneficence in the context of AI ethics and suicide prediction more broadly (Floridi et al., 2018; Halsband & Heinrichs, 2022).

According to Beauchamp and Childress (2002, p.165), the principle of beneficence means that “agents must take positive steps to help others, not merely refrain from harmful acts.” Furthermore, Beauchamp and Childress (2002, p.165) state that there are “two principles of beneficence: positive beneficence and utility. Positive beneficence requires agents to provide benefits. Utility requires that agents balance benefits and drawbacks to produce best overall results.” In this section (2.1), I evaluate whether the actions of agents (specifically schools and teacher) align with their duty of beneficence when using EdTech for suicide prediction.

2.1.1. EdTech Aligned with the Principle of Beneficence

This DPhil began by highlighting the severity of suicide statistics among young people (ONS, 2024; RCPCH, 2021), along with the current lack of cohesive support systems for at-risk students, as evidenced by the low positive predictive value (PPV) of existing tools (Large et al., 2016) and the impact of CAMHS budget cuts on available provisions (Local Government Association, 2023; Ofsted, 2020). In my introduction, I explored how factors such as high suicide rates and limited funding for CAMHS help drive the political and social momentum behind the development and deployment of new digital mental health tools.

If beneficence involves schools and teachers actively taking steps to help others rather than merely avoiding harm (Beauchamp & Childress, 2002, p. 165), then the implementation of suicide prediction tools in schools initially appears consistent with this principle. Schools can introduce digital tools in environments where young people are regularly present, which allows them to detect heightened suicide risk even when individuals have not directly expressed suicidal thoughts (Roy et al., 2020). More effective prediction may in turn allow teachers to provide better mental health care (either themselves, or referring students more effectively to external resources), thus aligning with their duty of beneficence. Additionally, as noted by Roy and colleagues (2020), Sueki (2015), and D'Hotman & Loh (2020), online platforms present uniquely beneficial opportunities for reaching young people due to the availability of highly personal data on students' personal devices.

Considering the limitations of conventional suicide prediction tools, and the severity of youth suicide, there is a compelling case for saying that schools and teachers are acting in-line with their duty of beneficence to the case of EdTech for Suicide Prediction, provided the effectiveness of these tools can be demonstrated (D'Hotman et al., 2020), and provided that the use of the tool is accompanied by appropriate follow-up intervention (the latter key to 'proving' beneficence). However, a closer examination of the literature raises ongoing questions about the clinical utility of these tools.

2.1.2. EdTech In Tension with the Principle of Beneficence

As mentioned previously, Beauchamp and Childress (2002) describe beneficence as encompassing both (a) teachers' duty to help others and (b) their duty to refrain from harmful acts. From a virtue-ethics perspective, Pellegrino and Thomasma (1981) also emphasize the importance of the moral character of the individual. They define beneficence as the moral obligation and virtue (typically of the clinician, but in this context of the teacher) to act for the good of the individual and promote their well-being.

Applied to teachers' use of EdTech for suicide prediction, the perspectives of Beauchamp and Childress (2002) and Pellegrino and Thomasma (1981) converge on a common concern: teachers will fail to uphold the principle of beneficence if they use tools that cause harm or do not support students' well-being. Accordingly, in this section I explore two main claims: (1) the EdTech tools in question are inaccurate (and I explain how such inaccuracy *may* lead to harm), and (2) the tools do not automatically translate into successful interventions or effective links to services (which again, *may* lead to harm).

2.1.2.1. Inaccuracy

With regards to inaccuracy, importantly, the harm does not arise merely from the fact that technologies are imperfect; rather, when suicide-prediction tools are insufficiently accurate, the recommendations or interventions they produce *may* fail to support students effectively. Such failures, whether through false positives or false negatives, can lead to a lack of utility, inappropriate responses, misallocated resources, and ultimately ineffective or even harmful care.

2.1.2.1.1. Data on Accuracy

A 2023 report from the think-tank RAND Corporation titled, '*Artificial intelligence-based student activity monitoring for suicide risk'* (Ayers et al., 2023), found that although several US-based EdTech companies have issued statements claiming their AI algorithms are useful for identifying students at risk of suicide (Bason, 2021; Madhusudan, 2021; Patterson, 2021; Shinde, 2021), there is a lack of transparency regarding the accuracy of these tools. In particular, Ayer and colleagues (2023) write that there are no studies reporting on the types of data being monitored, the accuracy of suicide-prediction EdTech tools, their positive predictive values, or any other related performance metrics. This is true both in the US and UK.

In addition, even when expanding beyond classroom-used tools to broader ML models for suicide prediction, the lack of transparency in reporting persists, and many models in fact perform poorly. Specifically, research indicates that ML models, much like traditional approaches, continue to have a low predictive performance (e.g. as discussed in Chapter 1, section 1.3). Kessler et al. (2020) explored a recent meta-analysis which explored 17 studies of ML-suicide prediction and concluded that the PPV averaged to be 0.004.

As mentioned earlier, what ultimately matters are the tools' utility in preventing suicide (rather than accuracy in itself), particularly in relation to the school's and teachers' duty of beneficence.

However, as outlined below, several safety concerns have emerged in the ethics literature, stemming directly from accuracy.

2.1.2.1.2. Safety Concerns based on Inaccuracy

There are significant safety risks which stem from the inaccuracy of AI-based suicide predictions (D'Hotman et al., 2020; Marks, 2019). The risks associated with false negatives are the most obvious with regard to safety (failing to identify someone at immediate risk of suicide); however, there are also safety risks in false positives. For example, a patient labelled as 'high risk' may receive more restrictive treatment and may be subject to social and/or self-stigma (Marks, 2019). In addition, patients deemed at 'high risk' could be subject to police interventions. Police interventions may have unexpected safety consequences, for example, police presence could escalate what was already a tense situation and lead to unintended or dangerous outcomes, including an increase in violence (Marks, 2019)[5].

As mentioned previously, there is insufficient information regarding the PPV of EdTech tools used for suicide prediction. Therefore, any hypotheses made by researchers, clinicians, or developers, about safety risks are based on broader ML literature. A thorough evaluation would require not only more robust data on PPV, but also comparative analyses across specific populations, as well as an appreciation that it is the tool's utility in predicting and preventing suicide (rather than its accuracy alone) that determines whether teachers and schools are acting with beneficence when using these tools.

2.1.2.1. Concerns around Effective Triage / Linking with Services

Another challenge to the principle of beneficence arises from broader questions: even if a test yields a true positive result, does the use of EdTech for suicide prediction truly lead to improved mental health outcomes for students? What happens after a student is flagged as at-risk (i.e., what

---

[5] For example, in October 2019, BuzzFeed reported on the use of Social Sentinel, an EdTech tool implemented by schools to detect suicide risk and potential school shootings (Thuy Vo & Aldhous, 2019). The article's authors raised concerns about the validity of the tool's findings, particularly highlighting the prevalence of false positives. Specifically, Thuy Vo and Aldhous (2019) demonstrated how students' social media posts 'demonstrating suicide risk', including tweets, were often saved and sent to schools without proper assessment and were frequently inaccurate, leading to unnecessary alarm or misidentification of risks.

are the intervention pathways, and whether or not they are successful)? Do these tools provide meaningful and actionable guidance for treatment?

The issue of effective triage remains under-researched, largely due to the limited data shared by the companies developing these technologies. However, a US-based doctoral thesis presents some findings in this area (Shelton, 2022). This study reviewed 533 cases flagged by Gaggle and assessed whether school teams took subsequent action. 93% of cases were deemed low risk, nearly half (48%) were not followed up by school teams, and formal suicide or risk assessments were conducted in only about 2.5–2.6% of cases (Shelton, 2022).

The low levels of follow-up reported by Shelton (2022) are not necessarily problematic, nor do they imply that EdTech tools lack the ability to support student mental health. Gaggle was designed to flag cases where further assessment should be considered, not only those where it is definitively required. Therefore, these statistics may be consistent with the tools functioning as intended. However, without comparative case studies (e.g. comparing support and triage processes in schools with and without EdTech) and/or more detailed, long-term monitoring of intervention data, including rates of false negatives and false positives, and more detailed rates for specific groups of students, it is not possible to make a definitive claim about beneficence or utility (either for or against).

2.2. Introducing Tradeoffs

Beyond beneficence, what other ethical principles are at play in the context of EdTech for suicide prediction? How are these principles defined, and what kinds of ethical trade-offs, if any, do teachers (and other agents) face when using these tools to address student suicide in their classrooms?

In the following section, I examine additional ethical principles outlined by Beauchamp and Childress by highlighting some of the ethical trade-offs emerging in the use of EdTech for suicide prediction. Specifically, Section 2.2 will explore two key tensions: between beneficence and autonomy (with a particular focus on privacy), and between beneficence and justice (with a focus on non-discrimination).

2.2.1. Beneficence, Autonomy, and Privacy

According to Beauchamp & Childress (2001) autonomy is defined as follows:

> Personal autonomy is, at minimum, self-rule that is free from both controlling interference by others and from limitations, such as inadequate understanding, that prevent meaningful choice. The autonomous individual acts freely in accordance with a self-chosen plan, analogous to the way an independent government manages its territories and sets its policies. A person of diminished autonomy, by contrast, is in some respect controlled by others or incapable of deliberating or acting on the basis of his or her desires and plans. (p.58)

There are various ways in which respect for autonomy and beneficence may be traded off in the context of using EdTech to predict students' suicide risk in UK classrooms. Importantly, because these are minors, the ethical landscape is more complex: children and young people possess developing autonomy, and the balance between respecting that autonomy and acting beneficently often differs from that applied to adults (Alderson, 2007). In addition, whether child or adult, McKernan et al. (2018) argue that the ethical principles of beneficence and respect for individual autonomy are in conflict when someone is deemed at 'immediate' risk for suicide. According to McKernan and colleagues (2018), someone who is deemed 'high' or 'immediate' risk by an EdTech tool may receive a wellness check, which may later bring about involuntary hospitalisation. This may save a student's life, but impede autonomy (McKernan et al., 2018).

Typically, in the context of involuntary hospitalization, ethicists argue for the importance of balancing an individual's lack autonomy with the principle of beneficence (e.g. in McKernan et al. (2018). However, the fact that minors have only partial or developing autonomy makes this balance especially delicate (especially in the case of a young person who is not able to be autonomous due to their age *and* presence of mental illness). And if EdTech companies can offer no, or only limited, demonstrated clinical benefits (as illustrated in Section 2.1), the justification for overriding or constraining a student's developing autonomy is significantly weakened. In such cases, concerns about autonomy become more pronounced and ethically troubling.

Beyond involuntary hospitalization, privacy is a key ethical issue rooted in autonomy. What follows is a brief introduction to the definition of privacy (Section 2.2.1.1), then descriptions of two privacy violations observed within the context of EdTech for suicide prediction (Section 2.2.1.2).

2.2.1.1. Definition of Privacy

Providing a single, clear definition of privacy is challenging, as scholars from various disciplines, including law, bioethics, and philosophy, continue to debate its precise meaning[6]. However, within bioethics, and the literature around suicide prediction tools, privacy has often been framed as a bioethical principle centred on individual autonomy. As Milton (2019) notes, "Privacy has been understood as a bioethical concept whose focus is on personal choosing or the right to control access to self" (p.106).

In the context of EdTech for suicide prediction, "access to self" (Milton, 2019, p. 106) can be understood as a young person's right to control how their digital data in particular, is collected and used. For example, according to Gomes de Andrade et al. (2018), privacy within the context of AI-based suicide prediction may involve control over the data collected to train the algorithm, data used during the actual risk assessment process, and access to information such as the risk score itself, rates of false positives and false negatives, and any post-intervention monitoring. Regan and Jesse (2019), who also emphasise the prioritisation of privacy within ethical analyses, note that within the use of EdTech overall, there are additional nested concerns, including "information privacy; anonymity; surveillance; autonomy; non-discrimination; and ownership of information" (p. 167).

Peterson (2016, p.962) argues that privacy represents a "central tension" in the current implementation of EdTech in schools, and many of the critiques of EdTech introduced in Chapter

[6] According to Nass et al. (2009) the reason why there are so many ways of interpreting privacy in the literature is because privacy is highly context specific. Nass (2001, p. 18) explains, privacy "acquires a different meaning depending on the stated reasons for the information being gathered, the intentions of the parties involved, as well as the politics, convention and cultural expectations."

1 (e.g., Figure 2) are rooted in privacy concerns.[7] What follows are two examples of ethical trade-offs between privacy and beneficence within the context of EdTech used for suicide prediction: surveillance and un-authorized data sharing.

2.2.1.2. Example 1: (Covert) Surveillance

According to the UK's Code for Crown Prosecutors (Crown Prosecution Service, 2021),

> Surveillance is defined by section 48(2) of RIPA as including monitoring, observing, or listening to persons, their movements, conversations or other activities and communications.
>
> Surveillance is covert if, and only if, it is carried out in a manner calculated to ensure that any persons who are subject to the surveillance are unaware that it is or may be taking place (section 26(9)(a)).
>
> (Crown Prosecution Service, 2021, Chapter 6, para. 1-2)

Surveillance, particularly covert surveillance, has become a major privacy concern with the rise of EdTech in classrooms. Public attention to the rise of covert surveillance within the classroom grew significantly in 2009 when the Lower Merion School District in the US "remotely activated its school-issued laptop webcams to capture 56,000 pictures of students outside of school, including in their bedrooms" (Marlow, 2025, para. 1). Although this case did not involve suicide monitoring specifically, it sparked lasting concerns about the misuse of surveillance technology in educational settings, and these concerns have persisted in recent years. For example, a US study of 502 students ages 14-18 showed that young people are concerned about unauthorized and/or covert surveillance in their schools (ACLU, 2023, Table 2).

---

[7] For example, in the US context, Senators Warren, Markey, and Blumenthal (2021) have argued that the ethical trade-off between beneficence and privacy is not justifiable, as the associated privacy risks (including eroded trust, increased harm, exacerbated racial inequalities) are too great. According to Warren, Markey and Blumenthal (2021, p.6), "It is crucial that the tools school districts select will keep students safe while also protecting their privacy, and that they do not exacerbate racial inequities and other unintended harms." Academics and policy professionals are not the only ones connecting the lack of privacy safeguards in EdTech to eroded trust, increasing harm, or racial inequality. In 2023, the ACLU published a research report titled 'Digital Dystopia: The Danger in Buying What the EdTech Surveillance Industry is Selling,' which similarly linked inadequate privacy protections to these broader issues. The think tank, RAND Corporation also reported the same, in their 2023 report titled '*Artificial Intelligence-Based Student Activity Monitoring for Suicide Risk*' (Ayer et al., 2023).

Table 2

Students' Concerns About School Surveillance (ACLU, 2023, p. 21)

Students' Concerns About School Surveillance

I always feel like I'm being watched	**32%**
How it could be used to discipline me or my friends	**27%**
What your school and companies they contract with do with the data (such as sell it, analyze it, etc.)	**26%**
How it limits what resources I feel I can access online	**24%**
Could be shared with law enforcement	**22%**
Could be used against me in the future by a college or an employer	**21%**
Could be used to identify students seeking reproductive health care (such as contraception or abortion care)	**21%**
Could be used to identify students seeking gender-affirming care (such as transgender students seeking hormones)	**18%**
Could be used against immigrant students, especially those who are undocumented	**18%**
How it limits what I say online	**17%**
Could be used to "out" LGBTQIA+ students	**13%**
I have no concerns regarding surveillance in my school	**27%**

Source: YouGov. School Surveillance, fielded October 20-26, 2022. Commissioned by ACLU

As shown in Table 2, US students express concerns not only about privacy and surveillance but also about potential disproportionate surveillance of vulnerable groups, including those seeking reproductive or gender-affirming care, as well as immigrant and LGBTQ+ students (ACLU, 2023).

In addition to privacy being a concern in its own right, the use of monitoring technologies in educational settings has been found to raise significant mental-health risks. Research by Semour, McNicoll, and Koeing (2024), suggests that the awareness of being monitored can increase vigilance, elevate anxiety, and influence neural responses associated with stress and diminished well-being. In the context of school-based suicide-prediction tools, this means that the very experience of data-monitoring can itself be triggering for some students, particularly those already experiencing mental-health difficulties. When tools intended to protect students simultaneously generate heightened anxiety, feelings of constant scrutiny, or disproportionate impacts on already marginalised groups, the ethical justification for their use becomes more tenuous.

While the research discussed in this section has predominiately focused on the US context, disproportionate surveillance is also emerging as a concern in the UK. Criticism has arisen over concerns that Impero Education Pro software (which is used to predict suicide risk) is disproportionately surveilling students from Muslim communities, raising serious concerns about privacy and the potential for racially biased profiling (Hooper, 2015).

Increasingly, critics argue that the surveillance of students using EdTech tools is often disproportionate to its intended (beneficial) purpose and frequently occurs without meaningful oversight (Peterson, 2016; Regan & Jesse, 2019). Furthermore, even when these technologies are promoted as tools for suicide prevention, they often result in surveillance that extends far beyond the scope of that purpose. And, while such practices may not constitute a legal privacy violation, especially if students or parents technically consent through school-based ‘Internet Use’ consent agreement, scholars increasingly agree that this may represent a situation where the principles of beneficence and autonomy (through privacy) are in significant tension.

2.2.1.3. Example 2: Data Sharing

Returning to Milton’s original definition of privacy, Milton (2019) notes that “Privacy has been understood as a bioethical concept whose focus is on personal choosing or the right to control access to self” (p. 106). In the context of EdTech for suicide prediction, the right to choose whether to share one’s data (and with whom) is a fundamental aspect of privacy.

On the one hand, sharing data can sometimes benefit the student. Sharing information between schools and Child and Adolescent Mental Health Services (CAMHS) could reduce siloing and potentially save lives (Department of Health & Social Care et al., 2021). However, concerns arise because data is often shared beyond these intended purposes, frequently ending up in the hands of third parties such as parents[8], advertisers, or journalists.

Bryan & Lurye (2025) from the Seattle Times and Associated Press, while writing a piece on AI tools, accidentally gained access to nearly 3,500 student records during a public records request related to Vancouver Public School District's use of Gaggle (a prominent EdTech tool used to predict suicide risk in North America). According to Bryan & Lurye (2025, para. 8), “Vancouver school staff and anyone else with links to the files could read everything. Firewalls or passwords didn’t protect the documents, and student names were not redacted, which cybersecurity experts warned was a massive security risk.”

The case at the Vancouver Public School District highlights two key issues: (1) privacy was compromised due to a lack of consistent data anonymisation, and (2) privacy was further infringed through unintended data sharing.

Ultimately, Section 2.2.1. demonstrates that while monitoring tools may help predict suicide risk and support students who might otherwise be overlooked, the current use of these tools in schools also raises serious privacy concerns (particularly regarding surveillance and data sharing), which can lead to unintended and significant consequences for both schools and students.

2.2.2. Beneficence vs Justice (and Discrimination)

According to Hosseinabadi-Farahani et al. (2021, p.1), “justice in health means the lack of systematic and potentially resolvable differences in one or more aspects of health in a population and economic, social and geographical subgroup.” Cookson (2015, p.99), in reference to how

---

[8] Academic researchers have questioned whether legal guardians should have the right to access their child’s algorithm-captured clinical profile, an ethical issue shared in the debate on privacy and electronic health records (Meredith, McCarthy, & Hemsley, 2018).

NICE conceptualises justice, identifies three substantive principles: "(1) cost-effectiveness, (2) non-discrimination, and (3) priority to the worse off."

In both definitions listed above, the ethical principle of justice is closely linked to (and often contrasted with) discrimination. In the context of health care, discrimination is defined as "a lack of provision, incomplete provision, or different provision of health care to an individual or group of individuals because of their individual and social characteristics" (Hosseinabadi-Farahani et al., 2021, p.1).

Justice, and by association, non-discrimination, are key ethical principles to consider with the use of EdTech for suicide prediction. For example, while it may be established that certain demographic groups are at higher risk of suicide (e.g. LGBTQ+ groups; Haas et al., 2010), should these groups be labelled (and monitored) more explicitly? Doing so may help target interventions (beneficence), but it also risks reinforcing stigma, discrimination, or over-surveillance (harm).

In 2022 Keierleber raised concerns about the keyword libraries on which EdTech programmes are based. Keierleber (2022) criticised these libraries of labelling words related to sexual orientation, including 'gay' and 'lesbian,' as harmful and risky content and outing LGBTQ+ students. In her work, Keierleber (2022) found more than three dozen incident reports in which Gaggle, a programme used predominantly in the United States, flagged sexuality-related terms such as 'gay' and 'lesbian' as indicators for a student being at 'high risk.'

Keierlber's example not only constitutes discrimination, as defined by Hosseinabadi-Farahani et al. (2021) as "different provision of health care to an individual or group of individuals because of their individual and social characteristics" (p. 1), but also has the potential to cause unintended harm. For instance, Desai-Hunt (2021) describes a case in Minneapolis where a student was outed after school technology flagged LGBTQ+ keywords in their writing, prompting staff to contact the student's parents. Researchers and lawyers have reflected on the harm of 'outing' young people to their parents without consent (Schafer, 2015), and therefore incidents like the one in Minnesota outlined above (Desai-Hunt, 2021) raise serious ethical concerns involving autonomy (privacy, data and sharing), beneficence, and justice.

While this section focused on LGBTQ+ students as an example, other groups have also faced discrimination. For instance, Muslim students in the UK have experienced over-surveillance as a legacy of the PREVENT duty (as discussed in the introduction and Section 2.2.1; Zempi & Tripli (2023)), and Hooper (2015), Persson (2022), and Maslaha (2023) highlight that this could potentially stigmatize students or result in punitive or disciplinary actions. Black and Hispanic students in the US have similarly been subjected to disproportionate scrutiny, as well as those coming from lower-income areas (Ayers et al., 2023)[9].

2.3. Responsibility Framework

Section 2.2 examined Beauchamp and Childress's ethical principles in relation to the use of EdTech for suicide prediction in secondary schools. In particular, Section 2.2 introduced tensions surrounding autonomy (including issues of privacy and data sharing), beneficence, and justice. It is important to note, however, that critics within the bioethics community, including Turoldo (2009) and Kass (2001), have pointed out significant limitations of Beauchamp and Childress's original four principles approach in the context of public health research. Turoldo (2009) and Kass (2001) argue that the four principles framework is overly focused on the individual (particularly, individual autonomy)[10], and tends to frame ethical issues in rigid oppositional terms or 'trade-offs', such as autonomy versus beneficence, and/or the values of the patient vs doctor.

Therefore, while the four-principles approach remains influential and is frequently used in the literature, particularly as a heuristic for structuring ethical analysis, its individualistic orientation renders it insufficient for the present topic. In contexts such as school-based suicide prediction (which involve institutional responsibilities and multi-actor decision-making), the framework struggles to capture the ethically salient features of practice.

[9] According to interviews conducted by Ayers et al. (2023, p. 32), "Black and Hispanic students, students who come from, you know, live in low, low-income areas . . . they are more likely to be dependent on a school-issued device and therefore not have like the means to essentially opt out of this tracking."

[10] Kass (2001, p.1776) notes that "codes of medical and research ethics generally give high priority to individual autonomy, a priority that cannot be assumed to be appropriate for public health practice."

More specifically, rooted in individual patient vs doctor dynamics, researchers such as Turoldo (2009) argue that the principle of individual autonomy is particularly ill-suited for addressing broader, community-level concerns typical of public health, e.g. preventative measures such as the use of EdTech for suicide prediction (Turoldo, 2009). In response, several alternatives have been proposed. First, relational autonomy theories emphasise that individuals are socially embedded, and that autonomy is not exercised through isolated self-sufficiency but is developed and sustained through relationships (Gómez-Virseda, de Maeseneer, & Gastmans, 2019). Second, although not used within this thesis, virtue ethics could add a complementary lens that focuses on the character traits and dispositions expected of teachers, school leaders, and EdTech developers, such as practical wisdom, trustworthiness, attentiveness, and compassion, which are not easily captured within a principles-based framework but are central to understanding ethical practice in educational settings (e.g. Radden (2006)'s work on Virtue Ethics and Professional Ethics). And third, extending beyond the meso-level focus of relational autonomy to the macro-level, scholars such as Schicktanz and Schweda (2012) and Turoldo (2009) advocate for a responsibility-oriented bioethics that moves "beyond the early boundaries of this discipline [bioethics]" (Turoldo, 2009, p. 1197).

According to Schicktanz, Schweda, Turoldo and others, responsibility-oriented bioethics provides a broad and flexible framework, helping researchers navigate more complex and macro-level ethical tensions seen in public health. Turoldo (2009) argues that responsibility is a useful concept because it encompasses complex systems thinking and includes social responsibility (Ahola-Launonen, 2015) and corporate responsibility, such as the ethical obligations of technology developers involved in suicide prediction (Broer, 2022). This is critical because EdTech companies themselves are emerging as new parties within the safeguarding process, and with this emergence come questions about who within the company might hold responsibility (e.g., CEOs, developers), how these actors should behave, the extent to which they should (and realistically can) be held responsible, and how they ought to operate within a wider network of professionals and institutions.

In addition, these examples also show that, while theory shows there should be a clear distinction between teachers, EdTech companies, and those who are mental health trained or designated safeguarding leads, in practice, these boundaries often become blurred. This lack of clarity introduces further ambiguity regarding responsibility, making it difficult for schools to understand where accountability lies when an automated system becomes involved in a safeguarding process.

A framework that accounts for both social and corporate responsibility is also highly relevant in the broader context of school mental health, where the notion that 'everyone is responsible' is embedded in UK statutory and non-statutory guidance - most notably KCSIE, which had been discussed in greater detail in Chapter 1 (and will be expanded upon in Chapter 5).

Aligning with these arguments, this DPhil proposes a responsibility-oriented framework for systematically investigating how ethical tensions unfold in real-world educational settings. It shifts away from a rigid autonomy-versus-beneficence paradigm toward a more collective and flexible ethical approach that recognises shared and corporate responsibilities among individuals, professional groups, and society (Turoldo, 2009). Chapter 3 moves to outline my methodological approach and clarify the specific definition of responsibility adopted - namely, antecedent responsibility, as opposed to responsibility framed as blame or accountability[11], or responsibility as a virtue.

[11] Turoldo (2009, p.1197), "I stress the difference between retrospective and prospective responsibility, showing that only the latter is a good candidate for public health ethics."

Chapter 3. Methodology

The key goal of my proposed research is to create a rigorous evidence base that helps to broaden the academic literature on the ethics of EdTech for Suicide Prediction, with a specific focus on the concept of responsibility. Following the method of empirical bioethics, this DPhil examines the theme of responsibility within the context of EdTech for suicide prediction by integrating empirical data with ethical analysis. This approach allows me to draw normative conclusions regarding the responsibility of teachers within this context.

To describe this methodology, Chapter 3 includes four key sections:

- **3.1:** Outlining my Primary and Secondary Research Questions: This first section outlines my primary and secondary research questions, focused on the theme of responsibility.
- **3.2:** Defining Responsibility: This section begins to outline definitions for both individual and shared responsibility, drawing on role-responsibility as a conceptual foundation.
- **3.3:** Introducing Empirical Bioethics: Within this section I define empirical bioethics and explain how this methodology can be used to bridge the gap between teachers' real-life experiences and normative theory, thereby answering my primary research question. I introduce Empirical Bioethics by using definitions provided by Ives et al. (2017). I also review several existing operational approaches that have been used to integrate empirical and normative work in the literature (e.g. McMillan & Hope, 2008).
- **3.4:** Mapping My Approach: In the final section, I outline how I will apply empirical bioethics throughout the remainder of the thesis to address my core research questions. Within this section I describe my process in-depth, evaluating each of my methodologies individually, as well as my approach to integration.

Ultimately, I hope that by documenting this process, other researchers may be prompted to consider similarly interdisciplinary approaches to mental health and ethics research.

3.1. Aims, Objectives and Research Questions

While it is well established that student mental health is supported by a broad 'ecological system' of individuals in society beyond just teachers (including doctors, parents, social services, peers, and other stakeholders; Jackman et al., 2022), this thesis begins by focusing specifically on teachers, and teachers' responsibilities. This decision is informed by both practical and conceptual considerations. Practically, the scope of a DPhil project requires limiting the number of interviews and groups that can be studied in depth. Conceptually, teachers offer a particularly compelling starting point, as they occupy a unique position in students' daily lives, often acting as first responders to emerging mental health concerns (Gunawardena et al., 2024). It is important to recognise, however, that teachers are not a monolithic group: different types of teachers hold different responsibilities, and roles such as safeguarding leads or pastoral staff carry distinct expectations in relation to student wellbeing. Accordingly, this thesis also explores how responsibility may vary within the teacher population (for example, depending on role, training, or designated pastoral duties). In addition, while the empirical sections of this thesis focus on teachers, and teacher responsibility, other stages of the DPhil project, including the conceptual analysis completed in Chapter 7, expand to include a model of shared responsibility between teachers and other stakeholders.

Table 3 presents my primary and secondary research questions in greater detail.

Table 3

Primary and Secondary Research Questions

Primary Research Question	• What responsibilities do teachers and schools have when using technology for suicide prediction, and how do these compare to what their responsibilities should be?
Secondary Research Question	• Given that teachers do not work in isolation (and that adolescents are embedded within multiple overlapping support systems), how might a model of shared responsibility (involving teachers, clinicians, parents, technology developers, and/or the students themselves) function? How should responsibility be shared (as indicated by legal, policy, and ethical frameworks), and is it being done so in practice?
Additional Research Questions (*by chapter*)	**Chapter 1:** • How has the field of suicide prediction developed over the past decade within the UK, and what factors have led to an increasing reliance on school-based big data/ML screening tools? • What ethical cases have emerged in the public sphere, regarding EdTech for suicide prediction? **Chapter 2:** • What ethical dilemmas are currently identified in the academic literature (regarding the use of EdTech for suicide prediction in schools)? What are the research gaps? **Chapter 4:** • How is EdTech currently being used for suicide prediction? • How do EdTech companies describe the role of the teacher and the school in their products' use? **Chapter 5:** • What is the context of using these tools within the larger school system, and how does teacher responsibility work within real world, complex environments? • What responsibilities do teachers and the school have for suicide prediction within the context of UK legal and statutory systems? **Chapter 6:** • What are the values and preferences of teachers regarding the use of new educational technologies to monitor young peoples' immediate risk of suicide online? **Chapter 7:** • How should responsibility be shared (as indicated by legal, policy, and ethical frameworks), and is it being done so in practice?

3.2. Defining Responsibility

Before outlining the operational methodology of this DPhil, and explaining how I answer the research questions listed in the previous section, it is important to first provide a framework for the key concepts of responsibility (3.3.1) and shared responsibility (3.3.2), as these are the foundational lens through which this entire research project is built.

3.2.1. Responsibility

Defining responsibility is challenging, because, despite the popularity of the term 'responsibility,' it is used differently across and within disciplines. As such, its definition is at times unclear and subject to disagreement (Schalock, 1998). For instance, researchers in psychology often define responsibility to be an individual, fixed personality trait (e.g. Bierhoff, 2002), while philosophers see responsibility differently (Hart, 1968).

In his quest to find "the welter of distinguishable senses of the word 'responsibility' and its grammatical cognates" (Hart, 1968, p.211), Hart found there to be four basic types of responsibility within the field of philosophy. These are role-responsibility, causal-responsibility, liability-responsibility, and capacity-responsibility. In this thesis I use a framework of responsibility-as-role, which is often used interchangeably with prospective responsibility or 'duty,' and examine a framework of role responsibility developed specifically within the context of a school/education system.

I chose to use the responsibility-as-role model for three reasons. First, because it is established within educational philosophy, with those such as Lunenberg et al. (2014) arguing that, by taking on the role of a teacher a person is attributed with a specific role-related task, in the same way someone would in any other job. Second, understanding role-responsibility and the duties assigned to a specific group is essential for determining potential liability or accountability (Meier et al., 2025a, 2025b). Finally, role-responsibility is the most directly action-guiding form of responsibility, making it especially relevant for offering practical guidance to schools and other stakeholders (Ryan et al., 2023).

Role-responsibility is often used synonymously with prospective [12] responsibility, or duty. According to Hart, the most notable feature of role responsibility is holding an office. Hart (1968) argues that role responsibility is derived from a person's place in a social organisation or institution, and the duties attached to said role. These duties could include providing for others or fulfilling the organisation's purpose/goals. With this in mind, the overarching, explicit framework of responsibility used in this thesis is:

> **Responsibility**: *The duties teachers (and other parties) have by virtue of the offices they hold.*

What does it mean to hold an office, and do the roles come pre-specified? Hart argued that not only does public institutional responsibility focus on jobs explicitly assigned with the role, but also tasks assigned by agreement or otherwise. Hart goes on to say that this office may be private (for example, based in the institution of the family) or public (for example, government or teaching). Hardimon (1994) also suggests a third sub-category – "non-institutional" role responsibilities which are attached by the reason of being human.

It is also important to distinguish between two types of roles teachers occupy: their professional role (e.g. practice-based obligations such as safeguarding, pastoral care, and pedagogical judgment), and their legal role (e.g. statutory duties, mandated reporting requirements, and compliance with regulatory frameworks). There may be overlap between the two (discussed further in Chapter 5). Responsibility also varies within the profession itself, as teachers holding specific positions such as designated safeguarding leads, carry additional role-specific duties that shape what is required of them in practice. In this thesis, I attend to both dimensions, recognising that the responsibilities attached to the professional practice of teaching may not always align neatly with the duties imposed by law, and that role-specific responsibilities may further

[12] This can be seen in comparison to retrospective, casual responsibility, which refers to the theory that if a person is responsible for something, they were the cause of it or can rightly be held accountable for it.

complicate this picture. Yet all of these are crucial for assessing how teachers ought to act when engaging with EdTech.

As this thesis begins by focusing specifically on teachers, using the definition above it is clear that, to understand the role-responsibility of teachers, I must look at their authority and duties within the student's environment. The office of a teacher comes with a number of legal and professional standards, as well as unwritten cultural expectations. For instance, using this definition, being responsible as a teacher in the UK would include holding Qualified Teaching Status (QTS), and being employed by a school (Department for Education, 2011). Throughout this thesis, and using the framework of role responsibility, I will more closely examine normative work, i.e. what ethics research suggest the roles and responsibilities of teachers ought to be (e.g., by exploring both legal and de facto authority through a scoping review of the literature, Chapter 5). I will then invite teachers to respond to these perspectives, allowing for a refinement of how responsibility actually manifests in their lived experience (this methodology is discussed further later in this Chapter/Chapter 3).

3.2.2. Shared Responsibility

If the overarching definition of responsibility within this thesis aligns with Hart (1968) and is "the duties teachers (and other parties) have by virtue of the offices they hold," then within this thesis, shared responsibility will refer to collective and distributed roles of diverse stakeholders in achieving a common goal.

> **Shared Responsibility:** *The collective and distributed duties stakeholders have to achieve a common goal, by virtue of the offices they hold.*

It is important to clarify that, by using the term "collective and distributed," I assume that collective responsibilities are not only held by the group of relevant stakeholders as a whole, but are also distributed among individuals within the group. Each person, therefore, ultimately bears individual responsibility. This distinction is crucial, as this understanding of shared responsibility requires active engagement from all stakeholders to prevent the dilution (or even complete abandonment) of the sense that they have a duty to act. This phenomenon is commonly referred to as the ‘problem of many hands’ (van de Poel et al., 2012).

Multiple researchers have conceptualized what collective, or shared, responsibility looks like across relevant parties (Lee & Loeb, 2000; Schalock, 1998). The definition listed above implies that responsibility is shared across all participants, each of whom has clearly defined roles and an active, collaborative role in decision-making processes. This definition is shared by Kon et al. (2016), who, in their study on shared decision making in Intensive Care Units, define shared responsibility as: “a collaborative process that allows patients, or their surrogates, and clinicians to make health care decisions together, taking into account the best scientific evidence available, as well as the patient’s values, goals, and preferences” (p.2).

The two definitions of individual and shared responsibility (both based on Hart’s framework of role-responsibility) will serve as the starting point for my analysis of teacher responsibility in relation to EdTech. However, these definitions may change as new methods and data are introduced in this thesis. Although these frameworks may evolve with new information, by clarifying the definitions of responsibility and shared responsibility early in this Chapter, I am now able to outline the methodology I will use to collect and integrate both normative and empirical work. I begin this process by introducing the academic discipline through which I approach this problem: empirical bioethics.

3.3. Empirical Bioethics

Within this DPhil project, I aim to explore the ethics of using EdTech for suicide prediction, focusing specifically on responsibility (through both individual and shared responsibility lenses), with the goal of making concrete recommendations about how responsibility ought to be shared within the current education system. I also will suggest ways in which the current process of sharing responsibility may not live up to these normative standards (if applicable).

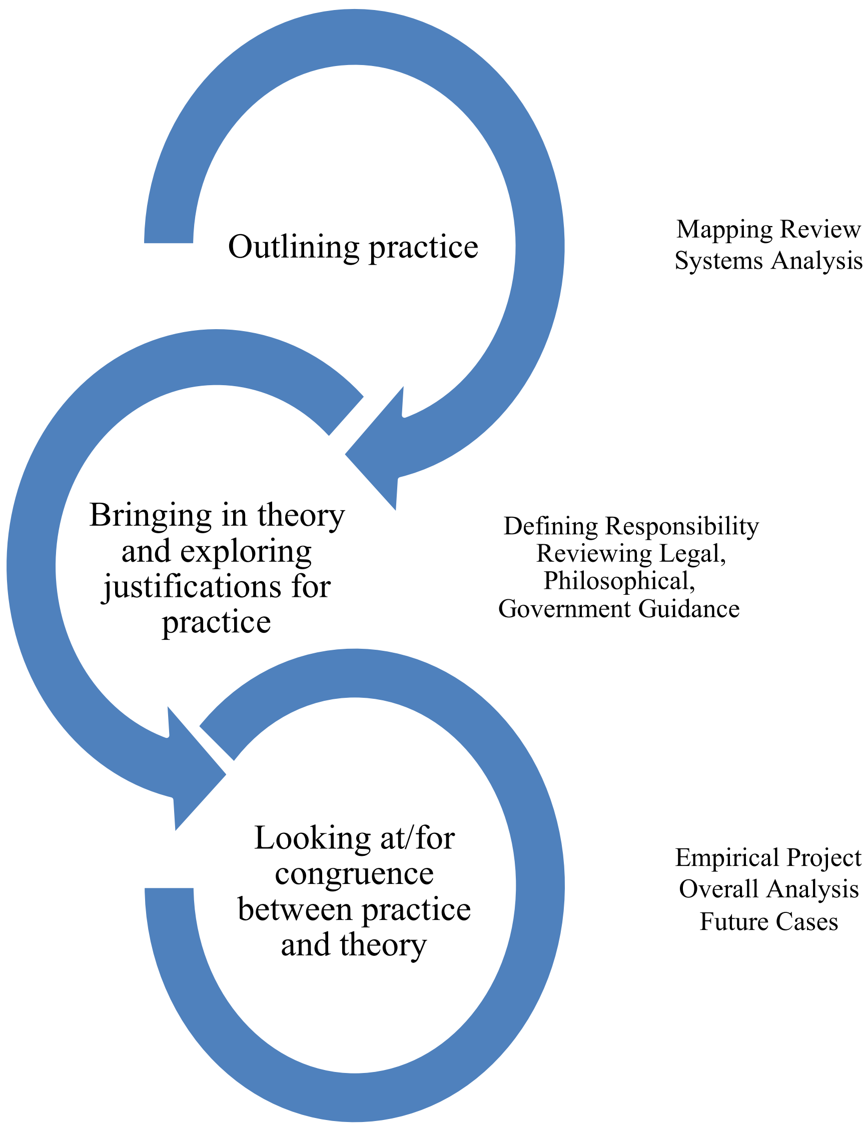
To do this meaningfully, I cannot simply ask teachers whether they feel responsible for their students' mental health without grounding the inquiry in a normative philosophical account of what responsibility is. Such findings alone would fail to engage with the ideal forms of responsibility, namely, what ought to be done (Solomon, 2005). At the same time, a purely theoretical approach would be inadequate, as it may overlook the ethical issues teachers experience organically within school settings (Solomon, 2005). For this DPhil project, my aim is to develop normative standards that can guide the appropriate distribution of responsibility within the current education system, and to examine whether these standards align with existing policies, practices, and expectations in schools. Put differently, the goal is to make recommendations about how responsibility *should* be shared and to identify any gaps between these recommendations and the processes currently in place. To achieve this, I propose using an empirical bioethics approach to bridge the empirical–normative gap. This methodology allows me to combine normative analysis with empirical research to explore how philosophical theories of responsibility are understood, enacted, and distributed in practice.

Empirical bioethics (Hunt & Carnevale, 2011; McMillan & Hope, 2008; Solomon, 2005), integrates the moral experiences of stakeholders with philosophical theory to inform normative conclusions and recommendations for future practice. McMillan and Hope (2008), in particular, emphasise that empirical research can uncover the ethical issues that arise organically within real-world settings, particularly within the context of healthcare and medicine. For these reasons, I chose to use empirical bioethics as my methodological tradition.

There are a number of operational ways to perform empirical bioethics research. Within this chapter, I first outline the types of methodologies available (3.2.2.), before identifying the methodologies I will use within the DPhil thesis. This includes (1) testing ethical theories and arguments by examining their implications in specific cases and looking for congruence, and (2) exploring (and comparing) how responsibility is distributed in comparable domains (Figure 3).

<u>Figure 3</u>

Explanation of Methodology



Ultimately, by using empirical bioethics methodologies to answer the primary and secondary research questions found in Table 3, this DPhil can start uncovering the lived ethical tensions within educational environments. These empirical findings can then be compared with normative accounts of responsibility to explore where practice aligns with, or diverges from, theoretical ideals (e.g. Chapter 7). A detailed example of how this will be carried out is provided in the following sections.

3.3.1. Choosing A Methodology within Empirical Bioethics

Most researchers agree that the process of empirical bioethics should be that of empirical work occurring alongside normative investigation (rather than a linear process with empirical work coming first). For example, Ives (2014) argues that normative analysis can be conducted alongside and embedded within the process of data collection.

However, researchers differ in their approach to conducting empirical bioethics, as there are several operational ways to conduct normative analysis alongside data collection. There are differences in data gathering strategies and methods of integration (McMillan & Hope, 2008). In their work, McMillan and Hope (2008, pp.17–18) identify six ways in which empirical and ethical approaches can be combined. Some of these approaches align with the methodology used in my thesis, while others do not.

More specifically, my thesis follows what McMillan and Hope (2008, p.19) refer to as cases where "empirical studies directly engage with ethical concepts." McMillan and Hope (2008) note that qualitative research studies often fall under this category, particularly qualitative studies which are used to challenge or refine existing ethical analyses. This may involve identifying key issues or beliefs that warrant empirical investigation (e.g. legal or ethical frameworks surrounding who ought to be held responsible for suicide prediction), and then conducting empirical studies that then cast doubt on (or affirm) the adequacy of current conceptual frameworks (such as qualitative, semi-structured interviews), and feeding those findings back into further ethical reflection. This is an iterative process, and maps onto the process illustrated in Figure 3.

This research project's approach also aligns with another model McMillan and Hope (2008) describe in which "ethical analysis identifies key empirical questions" (p.18). For instance, McMillan and Hope (2008) reference a study exploring what justifies a doctor to breach confidentiality: an initial ethical analysis was used to determine which empirical facts might matter, followed by empirical work to investigate those facts, and finally a return to ethical analysis informed by the data. This type of approach, where "ethical analysis identifies key empirical questions" specifically mirrors the structure of the initial stages (e.g. Chapter 4) of this project.

Ultimately, following McMillan and Hope's two models listed above, my goal is to develop an empirical study that directly engages with the ethical concept of responsibility, including (but not limited to) having the ethical analysis identify key empirical questions, which are later explored and integrated. To do this, I begin by exploring the different conceptions of responsibility found in legal and policy frameworks, therefore expanding on the initial definitions offered in this chapter, and then use this analysis to inform the design of empirical interviews. Combining the ethical and empirical within this way, I am able to investigate how teachers' values and conceptions of responsibility are interpreted and negotiated. A full description of this methodology is found in the next section (3.3.2).

3.3.2. Operationalizing My Methodology

Taking into account the above, the following section outlines my methodological approach (first as a whole before describing and justifying each individual piece of work).

1. First, I conduct a comprehensive mapping review to assess the 'state of affairs' regarding technology adoption in the United Kingdom (Chapter 4) and explore how (and whether) the normative concepts of responsibility and shared responsibility are expressed within technology manuals, advertisements, and/or any other promotional material. Is responsibility understood in a way that aligns with the definitions outlined earlier in this chapter? If not, how does the reported understanding of teacher responsibility differ from these normative definitions, and what do these differences suggest about how teachers ultimately bear responsibility for using EdTech in UK secondary schools?
2. Second, I undertake an analysis of policy and law to gain insights into the integration of suicide-prediction technology within schools from a macro-level perspective (Chapter 5). This includes searches within the fields of legal studies, government, and education, and a subsequent thematic analysis. By analysing ethical and legal reports of responsibility (both individual and shared), I can identify the key issues that my subsequent empirical study will address (e.g. a type of empirical bioethics methodology described in McMillan and Hope (2008) above).
3. In the third stage of research, I complete in-depth teacher interviews to understand how responsibility was conceived by educators, and what (if any) were the gaps between theory and practice (Chapters 6 and 7).

4. Ultimately, within the final chapters of this thesis, I integrate theoretical work with policy analysis and empirical bioethics, answering my main research question through the lens of shared responsibility (Chapter 7), and allowing me to make normative recommendations for policy makers, educators, and other researchers (Chapters 8).

The following table explains, in additional detail, the purpose, strengths, and weakness of the core methodologies: mapping review, systems review, and empirical research.

Table 4
Methods Overview

Type	Purpose/Rationale	Strengths and Weaknesses
Mapping Review	To outline the technology available for suicide prediction in UK schools and have a clear understanding of how responsibility is described within the tools themselves.	One key strength of a mapping review is that it is usually more rapid than a traditional systematic review (Gough et al., 2012). In addition, according to Campbell et al. (2023), the mapping review is also useful for fields where no review has been made (i.e., for an initial review, to pave the way for future work; Grant & Booth, 2009). However, according to Campbell and colleagues (2023), the primary limitation of a mapping review is that it is less rigorous compared to a systematic review. This leads to potential bias.
Systems Analysis	To outline how EdTech is used within the school and have a clear understanding of which actors might play a role within a young person's school system, and how responsibility is defined and theorized for these actors, both in ethics and in statutory and non-statutory guidance.	According to Kalafat (Institute of Medicine (US) Committee on Pathophysiology and Prevention of Adolescent and Adult Suicide, 2001) the strengths of Systems Theory / Systems Analysis, particularly within suicide research, is that the methodology can help consider the interactions among multiple stakeholders. This may lead to a deeper understanding of shared responsibility. In addition, according to Kalafat (Institute of Medicine (US) Committee on Pathophysiology and Prevention of Adolescent and Adult Suicide, 2001), by examining the system as a whole, using Systems Analysis, researchers may be able to identify underlying, systemic issues causing suicide. However, schools are inherently complex, and each school system may have different boundaries, interactions, and dynamics. As such, one of the limitations of this methodology is that a full systems analysis would be difficult to implement (Miser & Quade, 1985).This may inadvertently lead to this DPhil project missing some stakeholders (or links between stakeholders). Thus, this thesis relies on simplifying the system and focusing on a macro-level analysis, in order to ensure the analysis is manageable.
Qualitative Interviews	To gain a clear overview of how teachers interpret responsibility, working in practice, within their own school contexts. To assess if there are any gaps between theory and practice.	According to McMillan and Hope (2008), qualitative interviews are among the most widely used empirical methods in empirical bioethics. A key strength of this approach is that it is grounded in real-world data and enables deep exploration of ethical dilemmas, including participants' motivations, reasoning, and values (McMillan & Hope, 2008). This depth allows for a richer understanding of complex moral issues (McMillan & Hope, 2008). However, a noted limitation is the potential for subjective interpretation and the relatively small sample sizes, which may reduce generalisability and introduce bias (Lim, 2024).

3.3.3. Veridical, Realist, and Pragmatic Conditions

While I have attempted in the table above to justify my use of each methodology, not all empirical work relating to ethics can be classified as empirical bioethics. Instead, there are particular methods of distinguishing empirical bioethics from research grounded primarily in psychology or other cognate disciplines. Ives et al. (2017) propose a useful framework for determining whether a project qualifies as empirical bioethics. According to their framework, a study must meet at least one of the following three conditions to be classified as empirical bioethics:

(1) the veridical condition, which according to Ives and colleagues (2017) is:

> The research process attempts to ensure that the ethical issue being researched is genuine and authentic; framed in terms of the way it is experienced and negotiated in practice by moral actors, rather than constructed in abstract by a moral theorist. (p.3)

(2) the realist condition, which according to Ives and colleagues (2017) is:

> The research process attempts to ensure that the analysis is attendant to the circumstances in which moral actors find themselves, and pays due consideration to factors that may constrain or limit the actions or choices available to actors. (p.3)

(3) the pragmatic condition, which according to Ives and colleagues (2017) is:

> The research process attempts to generate conclusions/solutions to normative problems that are sufficiently respectful to, and engage sufficiently seriously with, the concerns. (P.3)

The next section will describe each section of my DPhil methodology in detail - both in its own right, and in relation to the broader Empirical Bioethics framework. This includes examining how each component of the methodology satisfies the criteria outlined by Ives and colleagues (2017) above, and how the process of integrating empirical with normative will take place.

3.4. My Methodology

3.4.1. Mapping Review

A mapping review was the necessary first step of my doctoral project. While a subsequent systematic review may be required, the field of EdTech for Suicide Prediction is still in its infancy and experiencing a surge of rapid growth. Because of this growth, a mapping review is essential – it allows for the rapid survey, catalogue, and evaluation of new EdTech companies. In addition, this methodology is flexible and has strong theoretical and practical precedence. A mapping review also helps meet the first of Ives and colleagues' (2017) conditions for empirical bioethics: the veridical condition. This is because it helps provide a genuine insight into what technology is being used in practice in UK schools.

3.4.1.1. What is a mapping review?

A mapping review is a methodology developed by the Evidence for Policy and Practice Information and Coordinating Centre (Grant & Booth, 2009). According to Campbell et al. (2023) a mapping review is sometimes called a mapping study, scoping review, or scoping study. Mapping reviews are often confused with systematic reviews, meta-analyses, and/or literature reviews. However, unlike systematic reviews, mapping reviews very rarely require rigorous quantitative analysis. Instead, mapping reviews are unique in their ability to help researchers to rapidly contextualise a field for subsequent literature reviews (Grant & Booth, 2009). Rather than finding all articles on a topic (what would happen within a systematic review), the goal of a mapping review is to find a good sample for eventual contextualization (Petersen et al., 2015).

Mapping reviews are commonly used to outline new trends in the use of technology for healthcare (Asan & Choudhury, 2021), as well as to summarise current ethical debates within healthcare and AI (Morley & Floridi, 2024). Dai & Ke (2022) constructed a systematic mapping review of the educational applications of AI, while Lorenz et al. (2019) mapped technology on the dementia care pathway. The latter included a literature review, as well as user and expert reports.

Beyond its use in healthcare research, this methodology is also used to map trends in technology and outline the availability of tools 'in the real world.' This can occur in industry, government, and the third-sector. For example, a recent attempt to map Humanitarian Technology, conducted by the third-sector organisation Access Now (Access Now, 2023). The Department for Education has also mapped the broader Education Technology market, using glass.ai and commercial trackers such as Crunchbase to identify companies working in this space (Aston et al., 2022). This included a quantitative analysis of technology available based on products, services, and applications.[13]

While there are definite benefits to using mapping reviews (such as their flexibility to outline rapidly growing fields), there are some limitations. Petersen and colleagues (2015) criticise mapping reviews because the search requirements are less rigorous compared to systematic reviews. As the purpose of a mapping review is to illustrate research trends (rather than evaluate), researchers are not required to consider *all* evidence, or *all* papers, published for it to be successful.

While a subsequent systematic review may be required, an initial mapping review was preferable for the sake of this DPhil project as it could be done rapidly, is flexible, and has strong theoretical and practical precedence. Furthermore, the mapping review also provides genuine insight into what technology is used in UK schools, which is integral to evaluating how the tools are used, by whom, and the ethical issues that arise in the wake of their use. This includes a focused review of how responsibility is described within the technology itself, *as compared to* the normative accounts listed earlier in this chapter. As Ives et al. (2017) argue, it is essential to ensure that the ethical issue under investigation (in this case, teacher responsibility) is genuine, rather than abstract or artificially constructed (meeting the Veridical Condition). Examining the technology directly is therefore a key starting point for establishing the authenticity of any future normative claims.

3.4.2. Systems Analysis

The next methodology I use is Systems Analysis. Within this thesis, I describe Systems Approaches broadly while placing particular focus on how the education macro-system (especially the policy environment) shapes teachers' use of EdTech for suicide prediction. The Systems Analysis approach has been used in education research (Vanderstraeten, 2023), as well as psychology (Bronfenbrenner, 1979). The following section explores: (1) the definition of the systems approach; (2) how I plan to integrate this approach into the development of normative findings; and (3) the potential limitations I anticipate in adopting this strategy.

---

[13] The Department for Education's report (Aston et al., 2022), included products for digital learning, class aids, school management, and hardware). The report did not include a specific analysis of "technology for mental health," "technology for wellbeing," or "technology for safety."

3.4.2.1. What is the Systems Approach?

Originally developed by Von Bertalanffy (a biologist working in the mid-20th century), the Systems Approach emphasises the importance of studying the system as a whole rather than in its isolated parts (Von Bertalanffy, 1968). Otherwise known as General Systems Theory (GST), or Systems Thinking, the Systems Approach is interdisciplinary (and transdisciplinary) in nature, looking holistically at how a system's constituent parts relate to one another and work together (Meadows & Wright, 2015).

A system is more than just the sum of its parts, so in addition to exploring all of a system's constituent parts, the Systems Approach also looks at how context, structure, and function, further relate to and create what we see as 'the School (/School Mental Health) System'.

Systems Theory underpins much of current psychology, and is used in development, evaluation, and implementation sciences. For instance, Bronfenbrenner (1979) looks at child development within the context of five different interrelated systems: a microsystem, mesosystem, exosystem, macrosystem, and chronosystem. For example, how parents might be interacting with their child (micro-system); how teachers and parents may interact with one another (meso-system); or how an intervention may relate to a school's mental health policy, specific laws, and national policies (macro-systems). On the other hand, coming from implementation sciences, Systems Theory also provides suggestions on how a singular intervention could be more efficient by considering (and utilising) the child's full ecological context, including mapping the relationships between technology developers, teachers, parents, and other responsible actors, and exploring the policy landscapes in which the school is situated.

In 2001, Kalafat put forward a rationale for using the Systems Approach for suicide prevention in schools (Institute of Medicine (US) Committee on Pathophysiology and Prevention of Adolescent and Adult Suicide, 2001). Systems Theory is now a popular approach to both evaluating and integrating mental health interventions within the school, even beyond suicide[14].

---

[14] Bradshaw et al. (2009) considered the importance of school context for school bullying campaigns; Schuelka & Engsig (2022) used the Systems Approach to study inclusive education and practices; Jennings (2021) used the approach to explore pupils' experiences of mental health in primary schools.

Kalafat [15] (Institute of Medicine (US) Committee on Pathophysiology and Prevention of Adolescent and Adult Suicide, 2001) argues that a Systems Approach has utility within the context of suicide and suicide prevention because:

1. *A Systems Approach allows us to consider what contributes to suicide*. The premise of this argument is that suicide is a complex behavioural issue, and more importantly, rooted in community (Institute of Medicine (US) Committee on Pathophysiology and Prevention of Adolescent and Adult Suicide, 2001). According to this perspective, community factors contributing to suicide include breakdowns in communication, antagonistic intergroup relationships, time constraints in care provision, and unclear divisions of responsibility between teachers, parents, students, and other staff. By using a Systems Approach, Kalafat (Institute of Medicine (US) Committee on Pathophysiology and Prevention of Adolescent and Adult Suicide, 2001) argues that schools may address these factors, which ultimately may reduce suicidality.
2. *A Systems Approach allows us to consider the system's needs*. Kalafat (Institute of Medicine (US) Committee on Pathophysiology and Prevention of Adolescent and Adult Suicide, 2001) argues that, not only can a Systems Approach explain what causes increased suicide rates in schools, it may also help suggest efficient solutions. For instance, there may be scenarios where an intervention has been introduced, and yet clinicians with training in suicide and youth suicide prevention are scarce. Such a situation highlights the need for comprehensive training among other staff members including teachers and auxiliary personnel (e.g. school bus drivers, cleaning staff), and even among students.

3.4.2.2. The Systems Approach and Empirical Bioethics

The Systems Approach, most typically applied within empirical research to evaluate the utility and efficacy of specific interventions, is not only effective in psychology (Bronfenbrenner, 1979), education (Vanderstraeten, 2023), and suicide research (Institute of Medicine (US) Committee on

[15] Within this paper Kalafat defined the Systems Approach, its rationale, and two case studies which used the Systems Approach as part of Suicide Prevention interventions. (Institute of Medicine (US) Committee on Pathophysiology and Prevention of Adolescent and Adult Suicide, 2001)

Pathophysiology and Prevention of Adolescent and Adult Suicide, 2001), but is also a critical and, indeed, necessary tool in empirical bioethics. Specifically, it offers a valuable framework for examining the broader paradigm of responsibility, and particularly shared responsibility, which relies on an understanding of which actors are playing a role within the context of using EdTech for suicide prediction.

I am not alone in arguing that systems analysis and systems thinking are essential for addressing normative, philosophical questions and for shedding light on ethical concepts such as responsibility. Cognate disciplines (particularly AI ethics) have already embraced systems thinking as a foundational approach. For instance, in AI ethics, this perspective is often described as a "socio-technological approach to technological ethics" (Green, 2021, p.209) or as a form of "relational ethics" (Birhane, 2021). Relational ethicists, including those listed above, highlight how ethical responsibility emerges through complexity, context, and interdependence - principles that are equally relevant in empirical bioethics, and in this specific context, relevant to exploring how teachers use EdTech for suicide prediction.

1. First, complexity: drawing on Birhane's (2021) model, a relational ethics framework enables attention to the diverse experiences within a school and to how students may be differently affected depending on their situatedness. This supports the development of responsive, child-centred interventions.
2. Second, context: consistent with longstanding arguments in Psychology and Education, this framework acknowledges that individuals are embedded within broader social, institutional, and power structures, all of which shape how technologies are experienced (Birhane, 2021).
3. Finally, interdependence: the framework recognises the mutual dependence of individuals and communities, including how technologies are both shaped by and integrated into these networks. This perspective also highlights that deploying such technologies has ethical implications that extend across the communities in which they operate (Birhane, 2021).

Therefore, in my Systems chapter (Chapter 5), I will not only use policy and law to map the various actors and institutions within the student's broader ecosystem (following traditional social science models, e.g. Bronfenbrenner's macro-system), but I also will iteratively compare how responsibility is distributed in practice to how philosophers, such as Hart, argue it ought to be shared (Chapter 5, 7). By evaluating the differences between multi-agency approaches as prescribed in statutory law and non-statutory guidance, with normative ideals of responsibility, I am able to go beyond analysing the EdTech programmes themselves and also consider how responsibility is (and ought to be) distributed (or shared) between responsible actors (including teachers, parents, and social services).

Ultimately, Systems Theory (and by extension, relational ethics) stipulates that, in order to consider the ethics of new technology such as EdTech for Suicide Prediction, or make a final determination on how responsibility ought to be distributed, researchers must first understand the context within which this technology is developed, and then address relational dynamics and power structures inherent in the technological systems. Conducting an iterative comparison of how responsibility is distributed in practice and how philosophers, such as Hart, argue it ought to be shared, offers a valuable lens for ethical analysis, and meets the standard of empirical bioethics I seek to achieve within this DPhil (Chapter 7).

3.4.2.3. Limitations of the Systems Approach

Despite its strengths, the methodology has limitations. In particular, it omits individual factors such as genetic predisposition, which may influence suicide risk. Researchers such as Brent and Mann (2005) suggest that genetic factors (including heritable traits like impulsivity and psychiatric vulnerability) contribute to suicidal behaviour. Although an investigation of genetics lies outside the scope of this paper, it is important to note that a systems approach carries the risk of de-centring or overlooking individual-level influences such as genetic factors. In addition, while the theoretical benefit to the methodology is that I am able to map a system, it is undeniable that the school is an inherently complex place. Therefore, a full systems analysis is difficult to implement, both in general (Miser & Quade, 1985), but particularly within the context of the DPhil. To account for this second limitation, I will limit my analysis to a few specific features of the system, specifically exploring the macro-system, and the statutory and non-statutory guidance which is in place to steer teaching communities. While this approach has the potential to overlook relevant stakeholders and micro- and meso-level details, it does ensure the analysis is manageable, which is important for this DPhil, which is time-limited.

A full operational explanation of this methodology can be found in Chapter 5, which explores how the Systems Approach can be used to look at how EdTech programmes are integrated into schools, looking at various systems approach models proposed by psychology and education scholars, and illustrating how the models have been used in practice.

3.4.3. Qualitative Interviews

The third methodology I employ in this thesis (following a mapping review and systems analysis) is qualitative, semi-structured interviews. During the course of this DPhil I interviewed teachers to gain a deeper understanding of their perspectives on the roles and responsibilities of various stakeholders in the use of EdTech for suicide prediction (Chapter 6). This approach was used in conjunction with normative accounts of responsibility, adopting a reflexive methodology approach: a well-established approach within empirical bioethics, and one which I explore further in this section.

3.4.3.1. The Value of Qualitative Interviews

There are two main benefits of including qualitative interviews in this piece of empirical bioethics work. First, as McMillan and Hope (2008) note, qualitative interviews are an important method within empirical bioethics, as they can be used to (a) identify key issues or beliefs that warrant further empirical investigation, and (b) challenge the adequacy of existing conceptual frameworks. In this DPhil research project, interviews are used to provide insight into teachers' values and how these relate to broader ethics concepts, including responsibility (individual and shared). Specifically, the interviews will help identify how responsibility is created, allocated, and enacted in practice.

Second, interviews can also help meet the pragmatic condition outlined in Ives and colleagues' (2007) framework. Specifically, by grounding this understanding in the lived experiences of teachers and exploring the nuances of how responsibility is conceived and developed, I am able to generate a potential solution to the normative question of *who* should be responsible and *how* responsibility ought to be shared which is rooted in real-life experiences, therefore pragmatic and more likely to be adopted by schools and policy professionals.

3.4.3.2. Interview Development

Within this project, interviews are used to provide thicker descriptions of normative issues, including responsibility. Developing ‘thick’ descriptions, by concentrating study on a very specific topic, is an advantage of qualitative interviews that is widely acknowledged across various methodological traditions (Geertz, 1973).

Developing interviews can be approached in a variety of ways, and interview methodologies can draw from various disciplinary traditions. For example, interviewing in empirical bioethics research has been conducted in collaboration with sociologists (Kingori, 2013) and anthropologists using approaches such as 'ethno-immersion' (Parsons et al., 2024).

In this project, interviews are conducted using psychological and psychiatry methodologies and include Grounded Theory as a way of rooting the analysis (Noble & Mitchell, 2016). Specifically, while interviewing teachers, I conduct my analysis by comparing what teachers think and experience in real school contexts (gathered empirically through interviews) with normative understandings and definitions of responsibility. I will outline the full interview methodology in Chapter 5, however, broadly speaking, this comparison is carried out through a reflexive process.

What is reflexive bioethics? Ives (2014) suggests reflexive bioethics as a method for pragmatic and interdisciplinary bioethics. According to Ives (2014), reflective bioethics works by a researcher (1) identifying a moral problem, (2) inquiring into the problem from a ‘disciplinary naïve’ standpoint, and then, (3) using a process of reflexive balancing to integrate the findings. Ives (2014, p.311) continues to say that the process of reflexive balancing can be considered similar to the method of deviant case analysis, where “the theorist forms a hypothesis and then actively searches for disconfirming data, which either results in the hypothesis being changed, or the apparently disconfirming data being explained and shown not to be inconsistent with the hypothesis after all.” According to Ives (2008) this system works:

> By challenging them [the participants], by pointing out inconsistencies, by bringing up counterfactual cases and thought experiments, we hone in on what is really important to them and the fundamental values on which their arguments and opinions are based - and the picture is very often different to how it was at the start (p. 4)

Within my qualitative interviewing section (Chapter 6), I will ensure that normative theory informs the way the interview is carried out. One way I will do this is by prompting teachers to respond to specific normative claims and reflect on different models of responsibility, assessing whether they agree or disagree with the ways in which teacher responsibility has been laid out previously by philosophers (as well as policymakers and lawyers). Teachers' responses, gathered through these semi-structured interviews, will then help refine my normative analysis by providing context on how teachers think and act in real-life situations, as well as their values and preferences for future practice (e.g. in future scenarios where the types of suicide prediction technology used in schools changes or adapts, explored in Chapter 7). This may include the teachers' values and preferences that are overlooked in existing normative frameworks, and as such can then be tested through additional work (e.g. policy analysis, further interviews, or additional normative inquiry). Ultimately, the methodology described in this section reflects an iterative process, where empirical data and ethical theory continually inform one another. A full description of my qualitative interviewing methodology (including approach to data analysis and normative reflexivity) is provided in Chapter 6.

3.5. Conclusion

In this chapter, I have described a methodological approach that is consistent with answering my key (primary and secondary) research questions.

With Hart's definition of responsibility in mind (outlined earlier in this chapter), it is clear that, to understand the role-responsibility of teachers, I must look at their authority and duties within the student's environment. Within this thesis, I choose to use an empirical bioethics approach to do so, and in subsequent chapters it will be clear how I conducted research about teacher's responsibilities for their students from philosophical, educational, legal, and technological points of view, using a collection of methodologies (mapping review, systems analysis, and qualitative interviews).

The subsequent chapters, Chapters 4, 5, and 6, enact my research methodology (mapping review, systems approach, and qualitative interviews). Each chapter will further explain and operationalise the method employed, including an exploration of the procedures, data collection techniques, and analytical frameworks. Then, in Chapter 7, I will synthesise the key findings and insights, drawing together the central themes that have emerged throughout the analysis, specifically focusing this analysis through the lens of shared responsibility. Chapter 7 will integrate the results, reflect on their broader implications, and consider how they align with or challenge existing literature, using reflexive techniques introduced above. I will conclude this DPhil by highlighting the contributions of this research project to the field and suggest directions for future inquiry (Chapter 8).

Chapter 4. Mapping EdTech Companies

In line with UK statutory guidance[16] all UK schools are required to use some form of monitoring or filtering software on their students' digital devices. These tools are typically developed by proprietary companies which deliver monitoring software for school management and administration.

As demonstrated in Chapter 1, these tools were originally put into schools to prevent students from accessing inappropriate and extremist material (Department for Education, 2020). However, due to a national push towards integrating artificial intelligence and technology-based approaches in mental health (Department of Health & Social Care et al., 2021; Department of Health and Social Care, 2023), many of these tools now also include 'wellbeing' or 'mental health' packages.

One example of this is Imepro's "well:being" program, introduced in Chapter 1, which, according to their website, uses ML to generate student risk scores, including factors such as suicide, attendance, bereavement, bullying, anxiety, friendship, communication, attitude to learning, and sexualized behaviour (Impero, 2017b, 2017a, 2018a, 2018b, 2024). Impero was re-named as Ativion in late 2024, and therefore names are used interchangeably throughout this thesis. Impero/Ativion was a prime candidate for suicide prediction as it is already integrated into schools for classroom and network management (and therefore the health and wellbeing 'add-on' was easy to purchase). A case study of a secondary school using Impero/Ativion for suicide prediction is described below. This case study is based on available material, as well as the experiences of a private secondary school in the Greater London area, reported by a safeguarding lead at the 2019 ACAMH Judy Dunn National Conference.

[16] Specifically, KCSIE and PREVENT. According to KCSIE, schools are required to: "ensure appropriate filtering and monitoring systems are in place" (Department for Education, 2024, p.39). According to PREVENT, schools are required to: "ensure children are safe from terrorist and extremist material when accessing the internet in school, including by establishing appropriate levels of filtering" (Home Office, 2024b, p.16).

Figure 4: Case Study One, Schools Using Ativion

Case Study One: *School Using Ativion*

School X has already been using the proprietary software Ativion for network administration. A representative from Ativion then informed school administrators that they have a new add-on: one which safeguards students and can provide suicide risk prediction. Its use will be covered by the general IT privacy agreement which parents signed earlier in the school year. School X signs up for this software. Later in the week, on a Thursday evening at 8pm, the school's IT Administrator and principal receive a message to say Student X is deemed high risk of suicide. When they click the link, they discover that Student X has googled "why does someone commit suicide" on their personal home computer. The principal calls the student's mother.

Schools have been using tools such as Ativion for the past decade, both for their original purposes (classroom management and preventing access to inappropriate material), and for wellbeing and mental-health monitoring (as illustrated in the case above). It is important to note that, because of this dual purpose, students and parents are often unaware of the extent to which such tools are used, as their operation is typically described only in general terms through students' signing of an "Internet Use Agreement." This raises questions about the validity of consent. The use of these tools also introduces further ethical concerns (addressed later in this thesis), including how their deployment may be shaped by a school's resources. For example, many private schools issue students with school-owned laptops that continue to be monitored at home, creating disparities in how surveillance is experienced across school types.

Despite the growing number of tools available, and the emerging discussion of their ethical implications, as of March 2025, there has been no mapping review or outline of these technologies – in academia or in the public domain[17]. Therefore, in this chapter I review the digital tools used for monitoring suicide risk in schools, answering the question: *how is EdTech currently being used for suicide prediction?*

In line with my primary research question, I then focus my analysis through the lens of responsibility. This includes examining how responsibility is enacted within the development,

[17] There has, however, been a review of available EdTech more generally, conducted by the Department for Education (Aston et al., 2022)

access, and use of the technology, e.g. by collecting data on the roles and responsibilities discussed in relation to teachers, parents, students, and the EdTech company themselves, as well as any other party listed. Ultimately, I believe that, by mapping available technologies, and analysing their associated informational materials, I am generating new insights about responsibility that can be explored through future normative work (a key way in which empirical and normative are linked, e.g. as per McMillan & Hope, 2008).

4.1 Methodology - Operational

According to Arksey & O'Malley (2005), there are five stages to a mapping review:

1. Identifying the research question
2. Identifying relevant studies
3. Study selection
4. Charting the data
5. Collating, summarising, and reporting results

Within this DPhil project, these steps were carried out in the following way.

4.1.1. Identifying the Research Question

A core research question for this chapter is: *How is EdTech currently being used for suicide prediction?* A related sub-question is: *How is responsibility conceptualized within these tools, either in the software itself or in any promotional materials I can find?* This includes examining who each programme identifies as a responsible party, and whether the role of each party is explained in any depth (both individually, and/or in any shared capacity).

4.1.2. Identifying Relevant Technology

I developed a list of known EdTech tools used for predicting and monitoring suicide risk. Throughout the DPhil I periodically updated this table upon searching academic databases

(PsycInfo, Medline), industry trackers (Crunchbase)[18], and grey literature. The following is a non-exhaustive list of keywords used for searching (Table 5).

Table 5

Keywords for Searching

Keyword	**School**	**Technology**	**Suicide**
Synonyms	School* College Student Class*	Techn* App Ed-Tech Software Monitoring Device Management	Suicid* Self-harm SSI Mental Health Wellbeing

To enhance the search strategy listed above, I also met with an EdTech specialist, attended EdTech Conferences to meet with company representatives, e.g. the Bett conference in 2020), and invoked public consultation through social media (e.g. Twitter). Finally, I cross-referenced this list with the UK Government's Directory of Safety Tech Providers (Department for Business & Trade & Department for Science, 2024).

4.1.3. Selecting Relevant Technology

While I chose to only include tools that were used in the UK, many of these tools are used in multiple countries - including the USA. There are several other tools that are used only in the USA, including Gaggle and Social Sentinel. I chose to exclude these tools from the current analysis because a primary aim of this thesis is to link the technology with the political and social climate of the UK (e.g. within the context of the PREVENT duty, updated KCSIE, and the newly introduced Online Safety Bill). Inclusion of US-centric tools would skew this data. However, US-centric case studies may be used as a method of later analysis or comparison.

Searching was focused on tools that are being used in the UK, for universal monitoring programs, and with literature in English. Searching also focused on tools which used ML to monitor 'naturally'

[18] Crunchbase has previously been used by the Department for Education to track EdTech companies more generally, and consider new or emerging technology (e.g. those that are still raising capital) (Aston et al., 2022)

occurring student digital behaviour. This may include a students' search history, emails, and/or other elements of their digital footprint. Exclusion criteria are those that are only used by parents (e.g. Bosco), non-education staff (e.g. those solely used by hospitals and/or Children and Young People's Mental Health Services - CAMHS), those only used with small-subsets of students (e.g. those in a specific after-school club, or Varsity Monitor which is used by sports teams), or those only used with a clinical population (aka targeted screening system). In addition, programs were excluded if they only consisted of self-report data (e.g. AS Tracking by STEER, which consists of online psychological tests taken twice a year, or Tootoot, which allows for confidential student reporting on issues such as mental health). Tools were excluded if they did not include ML elements based on real-time digital behaviour (e.g. iSAMS, iNEQE, CPOMS or SCR Tracker were all based on school and social services records). Finally, tools were also excluded if they are no longer in use (e.g. GEOCOP). A full inclusion and exclusion table can be found in Table 6 below. Further, while this is not a formal systematic review, an adapted PRISMA diagram is included on the next page to illustrate the process of reviewing (and including / excluding) EdTech companies (Page et al., 2021; Figure 5).

Table 6

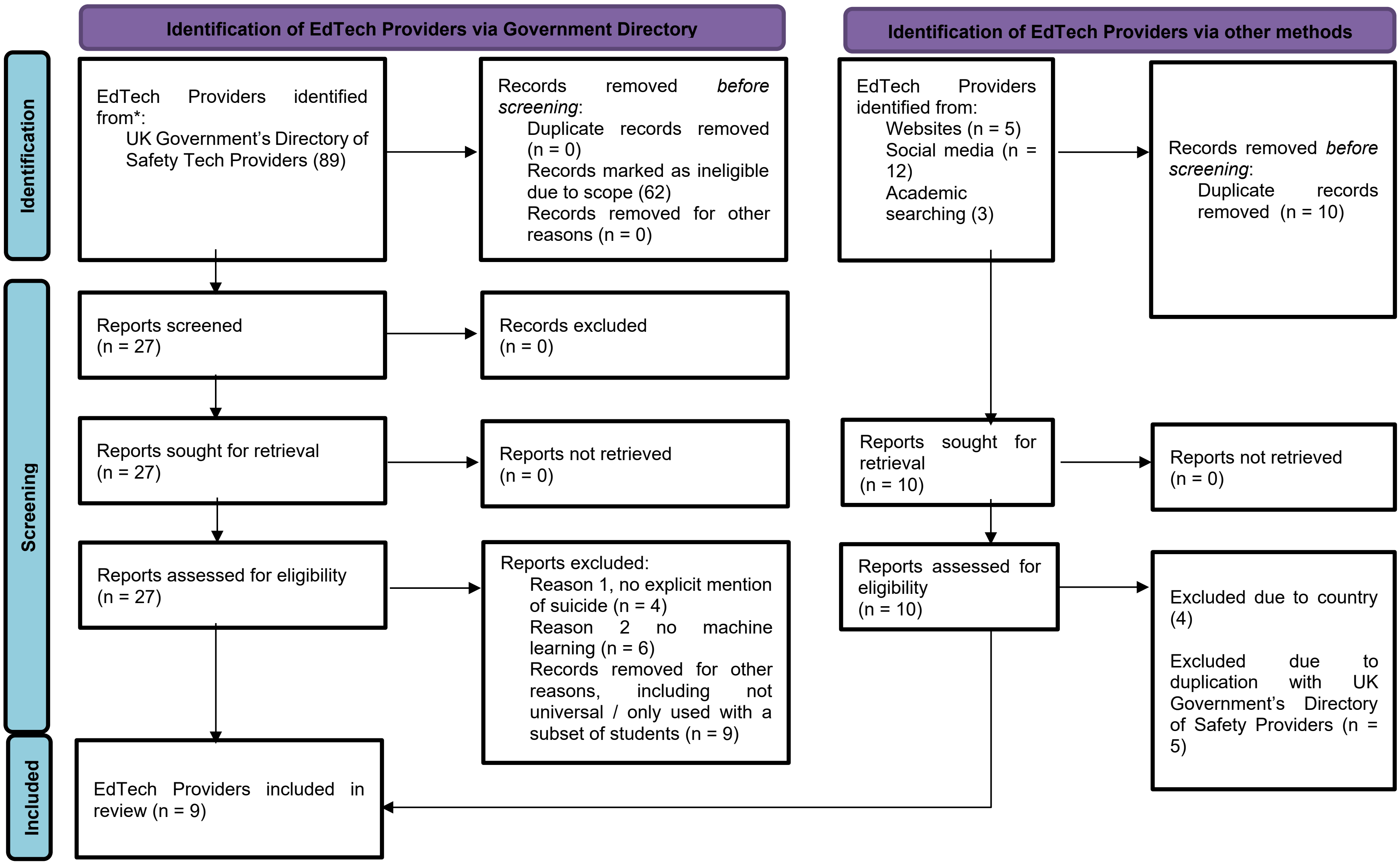
Inclusion and Exclusion Criteria

Inclusion Criteria	**Exclusion Criteria**
● Tools used in the UK ● Tools used by secondary schools ● Universal monitoring programmes ● Tools that integrated Machine Learning and monitoring of student digital behaviour ● Literature in English	● Tools only used by parents or non-educational staff ● Tools only used by primary schools ● Tools used with small subsets of pupils ● Tools only used clinically ● Tools that only consist of self-reported data ● Tools that do not include ML elements ● Tools that are no longer in use

Using the search strategy and the inclusion and exclusion criteria outlined above (and process outlined in Figure 5, below), a total of nine software programmes were identified. These are: Ativon/Impero; Smoothwall; LightSpeed Systems; Securly; NetSupport DNA; NetSweeper; Senso.cloud; Securus; and Schools Broadband.

Figure 5

Adapted PRISMA Diagram for Identification of EdTech Providers (template sourced by Page et al., 2021)



4.1.4. Collating/Summarising and Reporting Results

After identifying my final list of nine software programmes, in order to answer this chapter's secondary research question, "*how is EdTech currently being used for suicide prediction?*" it was imperative to gather data on two main topics: how these programs work with regard to their technology/ML; and how these programs work in regards to school practice (e.g. who interacts with them, how are they integrated into a school mental-health strategy, and how many schools use them/what types of schools). Any information that was deemed relevant to answering those questions was extracted and included. In addition, I collected any information that could help me address my primary research question on responsibility. Specifically, if any of the materials provided details on *who* was using the tools (e.g. teachers, students, parents, and others) and *how* the software defined the roles of each of these groups (if it was mentioned at all), then this was extracted.

Data extraction was completed using white papers, websites, promotional material, and information from informal interviews with company representatives, and added into a standardised data extraction table. In addition, I collected data from the software manual and decision tree of the relevant educational technology, if available. Finally, while I attempted to gather information from school safeguarding policies to understand how schools explicitly used these technologies, this proved extremely difficult. As discussed further in the limitations section, the absence of this information restricted my ability to determine key aspects such as how schools used the technology in practice, procurement processes, costs, add-ons, and the timing of schools' responses to alerts.

I developed categories for my data collection and analysis using an interactive process. Specifically, after collecting data on an initial three EdTech companies, I then proceeded to determine my categories for data summarising and reporting. After a process of categorization and re-categorization based on need and data available, I held a series of informal consultations with a school IT administrator and a representative of an EdTech company to finalise my twelve main categories, which can be seen in Appendix 1. This consultation process also allowed me to verify findings, as well as draw insight into why some data was difficult to capture (e.g. number of schools using each software program).

The final data sits in Appendix 1, and includes the following categories: 1. name of programme, (2) use, (3) materials screened (4) scanning risk of what?, (5) scanning technique, (6) presence of a human moderator, (7) interventions suggested, (8) compliance, (9) parental role, (10) teacher role, (11) student role, and (12) whether the tool was approved by professional bodies.

4.2. Results

In total, nine software packages were evaluated in this review. These are: Ativion (/Impero), Securly, NetSupportDNA, LightSpeed Systems, NetSweeper, Senso.cloud, Securus Software, Smoothwall, and Schools Broadband. All tools were developed by private companies, and the majority of the software programs were developed by companies which were already working within schools (e.g. in administration or governance).

Organisations typically combined filtering and monitoring capacities, for example excluding access to specific websites and keyword searches, as well as capturing data on students' real-time computer use. This included activity online (e.g. search history) and offline (e.g. written on Microsoft Word). All of the software programs included a form of screen-capture (e.g. saved screenshots and screen-recordings when a risk-alert was triggered), and three of the programs also allowed teachers to constantly monitor students' activity (Smoothwall, Lightspeed, and Ativion/Impero).

Some of the technology was based on connection to the school intranet or school device, whereas others could be downloaded on a "Bring Your Own Device" policy. The latter could be programmed with different options. For example, NetSupport DNA allowed parents to either keep restrictions on or turn them off when the device was used at home and in the evenings. If parents opted to keep restrictions on, then the programme continued to collect student data within the student's home.

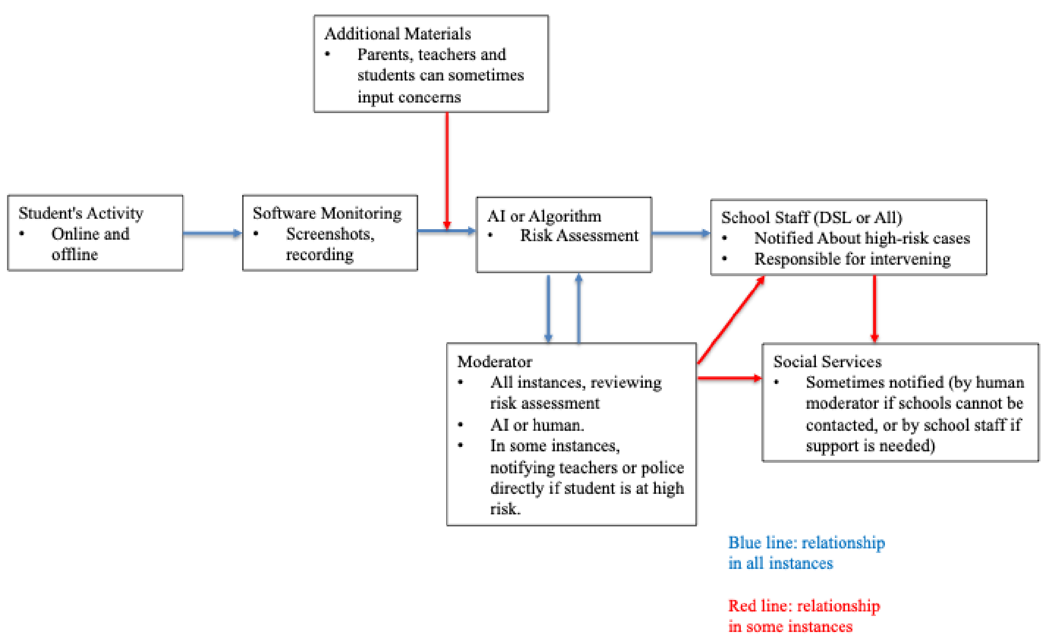
With regard to classification: most organisations used the Internet Watch Foundation (IWF)'s Keyword Library to categorise 'risk' (n = 6). It is important to note that the IWF is a UK-based charity dedicated to identifying and removing child sexual abuse imagery online. As a result, no software only looked at suicide risk. Instead, these programmes monitored 'risk' as a broad spectrum and included: cyberbullying, grooming, radicalization, violence, child pornography, self-harm, suicide, and racism.

Despite this wide spread of behaviours monitored, there was a high conflation of what constitutes 'risk' between various subtypes (e.g. bullying and suicide were both simply labelled as 'high risk behaviours' by all nine companies), and I found no promotional materials (for any of the programmes) that discussed how (and whether) risk would be subcategorized. For example, there was no mention of how violence-to-self and violence-to-others would be distinguished by their programmes, or how the software would respond to, and classify, a student searching for a gun or knife. Would searching for a knife be classified (by the ML tool) as general risk, risk-to-self, or risk-to-others?

In terms of triaging results, programmes filter content using either company-employed human moderators or an AI program before sending an alert to teachers or the school's Designated Safeguarding Lead (DSL). For example, as advertised by partner schools (IBS, 2025, para. 3), Smoothwall is able to auto-moderate content "in real time," and NetSweeper (2025, para 2) explains how, by "Using AI technology, it [Netsweeper] detects problematic user activity, then alerts administrators to verify and evaluate potential risks." In contrast, Securly uses live moderators, who work as a triage system - flagging activity, determining what type of action it needs, and then subsequently contacting the school (Securly, 2019). A map of how data typically flows within the identified EdTech programmes can be found below (Figure 6).

Figure 6

Description of Data Flow



As seen in Figure 6, the tool collects digital data, and this data goes through either (a) a company-employed human moderator, (b) an algorithmic risk assessment, or (c) both, and then it is typically the school's DSL who is responsible for acting on a high-risk assessment. It is important to note that, little information is publicly available about the company-employed human moderators (how they operate, and what training they may have). This process is in alignment with KCSIE guidance (Department for Education, 2020). For example, according to Smoothwall (2023, para. 1) and the Department for Education (2020, p.28),

> DSLs now have a responsibility for "understanding the filtering and monitoring systems and processes in place" as part of their remit.

However, the DSL is not the only one with access to the software systems. For example, according to Smoothwall (2023, para. 1) and the Department for Education (2020, p.32),

> Governing bodies should ensure that all staff undergo safeguarding and child protection training. It should give them "an understanding of the expectations, applicable roles and responsibilities in relation to filtering and monitoring"

While KCSIE makes it clear that a DSL's role is to understand which filtering and monitoring systems are in place, software companies differ on how they operationalise 'teacher responsibility' or a 'teacher's duties,' and it was clear that teachers had varying degrees of control and engagement with the software. For example, five programmes allowed teachers to constantly monitor their students' screens in the classroom, and four programmes explicitly allowed teachers to add personalised reports if they were concerned about individual students.

In addition to discussing teachers, six programmes (Lightspeed, Smoothwall, Securly, NetSupport, Netsweeper, and Ativion) also referred to parents, and parental input. However, they differ in the extent to which parents are able to monitor their children. For example, Ativion lets parents track and monitor devices including browsing history and location, Securly has a 'parents' portal' which schools can either activate or not, to either let parents have access to all activity, home activity or just educational activity. Lightspeed sends parents concentrated, weekly reports on how their children use the internet (including time spent and access to blocked sites).

Beyond teachers and parents, no students were able to directly see their internet use; however, five allowed for students to input data directly into the platform, although in a much more limited capacity. For example, Securly's TipLine. Smoothwall also had a spot for regular input, called the '60 second pulse' which was a student wellness survey conducted weekly (see Appendix One for more details).

Finally, in terms of shared responsibility with other services, only one company (Smoothwall) reported that if a situation is urgent and the company was unable to get a hold of a teacher, they will directly call emergency services.

4.3. Discussion

The following section discusses key findings from this mapping review and examines how the technological features and results relate to the definitions of individual and shared responsibility outlined in Chapter 3.

4.3.1. Shared Responsibility

According to Kon et al. (2016)'s original definition of shared responsibility, included in Chapter 4, an 'ideal' framework of responsibility (and shared responsibility) includes the active engagement of *all* relevant stakeholders in decision-making processes. However, this mapping review shows that, while multiple parties may have given some level of access to these EdTech monitoring systems, many of these avenues for input exist in name only (e.g. parents receiving updates but not being able to access real-time data; or students not being able to access their own data at all). Therefore, according to the materials examined in this chapter, it is evident that shared responsibility in its 'ideal' form is rarely embedded meaningfully into the software design.

In addition, the review is critical in identifying additional individuals within these companies who hold a level of responsibility for predicting suicide risk: the company in-house human moderators. Discussed later in this chapter (section 4.3.3.), the introduction of a human moderator brings about questions on how these actors should behave, the extent to which they should (and realistically can) be held responsible for student suicide prediction, and how they ought to operate within a wider network of professionals and institutions (including the school, teachers, and parents).

Finally, it is also important to recognise that the software functions can vary significantly depending on the school context (and this was explained in multiple documents surveyed). This means that the same monitoring tool may be implemented very differently across different schools. As such, to get a full understanding of how these programmes navigate responsibility and shared responsibility, I must explore how they are integrated into specific school ecosystems (addressed in following chapters).

Ultimately, this mapping review explored how teacher responsibility (and, by extension, shared responsibility) is explicitly addressed in EdTech materials, and how these representations may diverge from the normative framework of responsibility originally outlined in Chapter 3. However, it is important to also address that responsibility is also coded more implicitly into the technology itself. The next section shows how certain technological features (e.g. the use of singular risk scores, particularly focusing on the conflation of risk related to harm-to-self and harm-to-others) could impact teachers, and diverge from the 'ideal' definition of teacher responsibility.

4.3.2. Dual Harm, Singular Risk Scores, and Responsibility

One of the key findings from this mapping review was that the reviewed tools often predict risk for multiple issues simultaneously (e.g., cyberbullying and radicalization and self-harm and suicide) but do not separate these categories, instead provide only a broad determination of whether a student is 'at risk' or not. What are the implications of constructing a singular risk score? Is the lack of specificity a cause for concern, and if so, why? And, more specifically, what are the clinical and ethical implications of collapsing risk into one score, particularly when harm-to-self with harm-to-others is conflated, and how does this conflation impact our evaluation of teacher responsibility?

First, it is important to note that there is a clinical precedent for this conflation, with significant literature surrounding the concept of 'dual-harms.' Dual-harm refers to the theory that self-harm and aggression towards others are significantly associated and share some of the same risk factors (Shafti et al., 2023). Shafti and colleagues (2023) recently conducted a critical appraisal of the dual-harms literature to explore and evaluate this claim.

In their review, Shafti and colleagues (2023, p.1) found that dual harm does exist, and that it results from “the interaction of psychological risk factors that are associated with self-harm and aggression.” Their findings are aligned with Plutchik et al. (1989)’s model of countervailing forces, and Boxer (2010) theory of ‘high loading’; of risk across a person’s life, which leads to the prevalence of both self-harm and others-directed-aggression. It is also aligned with the findings of Steinhoff et al. (2023) who explored dual harm in children and young adults more specifically by conducting a longitudinal study, and found that “Between 13 and 17, 7.2% of adolescents reported dual-harm (self-harm only: 16.2%; other-harm only: 13.3%)” (p.3995).

Despite these findings, within the same meta-analysis, Shafti et al. (2023), acknowledged limitations in current studies around dual harm, including concerns with limited sample sizes. Furthermore, there may be ethical and legal concerns of linking self-harm and harm-to-others within risk assessments, despite initial evidence suggesting a relationship (Boxer, 2010; Plutchik et al., 1989; Shafti et al., 2023; Steinhoff et al., 2023). These ethical and legal concerns map onto the ethical principles outlined in Chapter 2.

For example, the concept of ‘dual harms’, where behaviours like violence, bullying, and self-harm are conflated, has significant implications not only for teachers’ sense of responsibility but also for how they collaborate with external agencies such as the NHS and the police. As these different forms of harm are often intertwined, teachers must navigate complex decisions (and legal frameworks) about when and how to intervene, requiring more nuanced approaches to student welfare. This can blur the boundaries of their role, compelling teachers to share the burden of responsibility with healthcare providers, law enforcement, and social services.

This becomes most evident when considering which types of interventions are suggested on the basis of a high risk score. Since the responses for individuals who pose a risk to themselves differ from those who pose a risk to others, conflating the two can lead to inappropriate interventions. For instance, individuals assessed as being at ‘high risk’ could face police involvement, although this can often intensify an already tense situation, leading to unintended safety risks for those at risk of suicide and self-harm (Marks, 2019).

This is particularly relevant in the present context because the introduction of this technology into schools is closely tied to the UK's PREVENT agenda, which is explicitly concerned with identifying early signs of radicalisation or harm to others. As a result, even when the technology is deployed with a primary focus on suicide prevention, its design, underlying risk prediction logic, and institutional governance remain shaped by a dual-use orientation. This history makes it especially important to remain attentive to how interpretations of risk may slide between concerns about self-directed harm and concerns about violence or extremism, reproducing the very conflation evidenced in the dual-harms literature.

These examples highlight the need for clearer communication, stronger partnerships, and more defined roles among all parties involved in safeguarding students, so that each can contribute effectively without overstepping or underestimating their respective duties. This shift also underscores the need for a deeper examination of whether singular risk scores genuinely influence teachers' responsibility (and/or how they share responsibility with other parties, including parents, social services, and the EdTech companies themselves).

Ultimately, the possibility of singular risk scores impacting teacher responsibility is an emerging idea that is not directly addressed in the existing normative literature. As McMillan and Hope (2008) note, one of the key purposes of empirical bioethics is to generate new ideas or hypotheses that can inform future research. Therefore, in this context, the notion that dual harms may influence conceptions of responsibility is a theme that will be explored further in the following chapter.

4.3.3. Moderating, Utility, and Responsibility

Beyond singular risk scores, a second technological feature that may impact teachers (and further distance these cases from the 'ideal' definition of teacher responsibility discussed in Chapter 3) is the distinction between risk moderation conducted by ML or AI versus that done by a human moderator.

More specifically, some of the tools listed used human moderators to triage cases, while other tools rely on ML. Beyond introducing a new individual that may be responsible for predicting suicide (a moderator), this next section will explore whether there are any clinical differences between the two forms of moderators, and if so, the ethical implications of this difference, particularly when it comes to teachers and teacher responsibility.

All of the tools listed rely on moderation to determine which cases are severe enough to pass forward to school officials or emergency services, and which may be considered ‘low risk’ or false positives (for example, a student googling a music video or typing “why do people kill themselves” for an academic assignment).

To Gorwa et al. (2020), the quantity of data screened is too high for human moderation to be feasible. As such, they automated triage systems, particularly within ‘wellness technology’ or digital tools for mental health. For instance, Milne et al. (2019) put forward an argument for an automated triage system, largely based on timeliness and efficacy. In their example, an automated triage system significantly reduced the time taken to respond to peer-support messages.

However, while a ML/AI moderator may be faster, Milne and colleagues (2019) also found some limitations to its use. First, programmes using ML/AI did not reduce the time taken to respond to high crisis messages (although this might be because human moderators were already more likely to respond quickly to high priority content), and second, there was no evaluation of the impact of AI/ML-based mistakes.

The question of ‘mistakes made’ is critical when determining how schools moderate online spaces, particularly when those spaces focus on mental health and child safety. Therefore, to other researchers (e.g. Marks, 2019), the speed at which ML moderation can be completed is not worth the safety trade-off.

According to Marks (2019), ‘mistakes made’ are safety risks, and safety risks stem from the inaccuracy of AI-based suicide predictions. Within his work, Marks (2019) illustrates that there are two types of risks associated with the inaccuracy of AI-based suicide prediction: risks associated with false negatives, which relate to safety (failing to identify someone at immediate risk of suicide), while the risks associated with false positives relate to stigma, and the provision of incorrect interventions. A patient being labelled as ‘high risk’ may receive more restrictive treatment and stigma (Marks, 2019). Additionally, patients deemed at ‘high risk’ could be subject to police interventions, which may have unexpected safety consequences (included added risk of violence), as responses may escalate already tense situations (Marks, 2019).

In addition to the trade-off between safety and speed, there are other ethical concerns about using AI and ML for triaging. For example, large-scale data collection and current practices of suicide prediction may leave those deemed at-risk vulnerable to invasions of privacy. Ultimately, from training the ML algorithm, to its use in production, interventions, and post-intervention monitoring, there are concerns at all stages of ML-based suicide prediction (Gomes de Andrade et al., 2018).

Based on the above, it seems critical for a human to be responsible for moderating EdTech for suicide prediction. And yet, human moderators are also fallible, and prone to similar mistakes and biases. For example, as discussed by Milne and colleagues (2019), human moderation is often slower[19] .

The fact that different types of moderators (human and AI) are used for different EdTech programmes is a key insight of this chapter, with significant implications for responsibility. This is true both in a descriptive/practical sense (e.g., determining whether a teacher is engaging in moderation) and in an ethical sense (e.g., whether they *should* be responsible for moderating).

This section hypothesises that a critical component of responsibility may lie in the moderator’s actual ability to be clinically effective (e.g. most arguments around moderating with AI vs human are based on this principle). This is a key insight emerging from this chapter and raises an important research question for future analysis: When examining responsibility through the lens of law and ethics (particularly UK statutory law on regarding shared responsibility and ‘working together’, e.g. outlined in KSCIE), is each party’s ability to contribute meaningfully taken into account when responsibility is distributed? Should it be? These questions will be explored further in the next

[19] While not addressed in this chapter, a third, alternative to moderation may be an “integrated moderation approach.” an integrated approach. For example, McCosker et al. (2023) reviewed three different digital mental health services, and models of integrated approaches. They suggest that, even within human moderated programmes, there are reactive and adaptive processes. They suggest that an integrated “adaptive logic of care” may be most effective when it comes to monitoring digital mental health services.

chapter(s), particularly within my analysis of teacher interviews (Chapter 6). It also raises important questions about who within the company can be (and should be) considered a responsible actor. In cases involving human moderation, responsibility might plausibly fall to the in-house moderator. But in systems that rely on AI moderation, should responsibility instead (or additionally), rest with the software developers who built the system?

4.4. Conclusion

This chapter maps the field of EdTech for Suicide Prediction, and suggests how trends in this field (e.g. what data is collected, how risk scores are constructed, and the process of moderation), may start to shape my definition of responsibility: what teachers' responsibilities are, what they look like in the context of the types of technology available, and how this might connect to what we consider teachers' responsibilities ought to be.

However, in terms of the limitations of this project: while this is the first mapping review to be completed in this field, it is by no means complete, and there remains significant work to be done in this field before a full evaluation of EdTech companies can be completed. First, a lack of transparency exists regarding how many schools use this technology, as no company has publicly disclosed such information. Second, there is a limited amount of data assessing the efficacy of these technologies. Not only do we, as an academic community, not know which schools are using these programmes, we also do not know if they work, and what interventions are used as a result of a positive score. Therefore, all discussion about the efficacy of these models, and of human versus AI moderators, is hypothetical.

Third, this review relied solely on publicly available data. A subsequent school-level survey would provide a far clearer understanding of how these technologies are actually used in practice. Such a survey could help clarify, for example, how schools respond to monitoring alerts and the timing of those responses (e.g., whether schools respond only during school hours or also after hours). An analysis of school safeguarding policies would also contribute to this understanding, but this was beyond the scope of the present thesis.

In addition, more information about cost and procurement practices would be beneficial. This includes understanding how much schools pay for this software, what is covered, and which schools use particular add-on features. This information is not readily publicly available, and from an online search, I found only limited examples. For example: Impero Classroom Management was reported on an IT user forum to cost approximately £3.30 per user for a school with around 700 devices, whereas NetSupport Pro reportedly has a starting price of around £10.00 per licence with no subscription model (EduGeekNet, 2019). Pricing is difficult to calculate because costs vary depending on specific requirements, add-ons, and the scale of the organisation. Pricing models also differ, with both subscription-based and one-off purchase options (e.g. Impero vs NetSupport, above). Much more research (ideally at the individual school level) is needed to understand the financial implications of using EdTech for suicide prediction, including how much of a school's budget is dedicated to these systems.

Building on the need for greater clarity around cost and procurement from the school perspective, it is also important to consider the financial structures and commercial practices of the companies providing these data-monitoring systems. Beyond simply understanding what schools pay, Freedom of Information (FOI) requests could reveal how these companies generate profit, particularly in relation to the buying and selling of data. Assuming the information collected is treated not as medical data but as social data, this raises crucial questions about where the data may be sold, how it circulates within wider data markets, and what responsibilities these EdTech companies have regarding the protection and ethical use of such information. These questions connect directly to broader ethical debates in the literature (originally outlined in Chapter 2), particularly around privacy, consent, commercialisation, and the governance of data-driven technologies in education.

Finally, a limitation of this study is that the field is rapidly evolving, as evidenced by the rebranding of Impero to Ativion within the course of writing this chapter. Any data collected became outdated quickly, thus this document required continuous review throughout the DPhil (and will require regular updates beyond the scope of this project).

Despite the above limitations, this Chapter provided an important foundation for my doctoral project by developing my research questions further, suggesting hypotheses to be tested in later chapters, and demonstrating how normative and empirical approaches might initially be integrated. The next chapter continues my exploration of teacher responsibility, integrating my technological analysis with a systems analysis, diving deeper into how these tools work in practice.

Chapter 5. Systems Approach: Macro-System Analysis

In the previous chapter I mapped the software landscape of UK schools, asking "what programmes are being used by UK schools to monitor for student suicide risk?" Within that chapter I listed the software programmes available and then explored the core features/elements of each. This included how each programme navigated moderation, risk assessment, and triaging, as well as how each programme both explicitly discussed responsibility, and implicitly codified responsibility into their software (e.g. through processes of moderation and risk assessment via singular risk scores).

In this next chapter I contextualise the use of these tools within the larger school system, and explore how they work (and by extension, how responsibilities are assigned), within real world, complex environments. To do so, I use a Systems framework (introduced previously in Chapter 3), and focus on a macro-system analysis (including policy, law, statutory, and non-statutory guidance). This chapter ultimately enables me to return to Hart's original definitions of responsibility and shared responsibility, introduced in Chapter 3, and to refine or expand these definitions in ways that are grounded in the macro-systems in which both students and technology are embedded.

5.1. Application of Systems Theory

While, in his 2001 work, Kalafat urges suicide researchers to use Systems Theory within their work, he provides no clear method of application (Institute of Medicine (US) Committee on Pathophysiology and Prevention of Adolescent and Adult Suicide, 2001). School mental health researchers also differ tremendously when applying Systems Theory. This is, in part, due to the difficulty of the task, as it is incredibly difficult to map a system as unpredictable as a school. According to Schuelka & Engsig (2022) it is difficult to map the school system because,

> An educational system is a dynamic space where elements interdependently interact in unpredictable ways; in which new patterns and new phenomenon may emerge; and in

> which elements may adapt based on changes to the system; and where elements themselves may be shaped by their own actions or the shifting dynamics of the system itself. (p.5).

Therefore, the first section of this chapter is dedicated to describing the differences in how the academic literature applies Systems Theory to the school and determining a preferred approach (one most suited to this doctoral project). In order to justify this choice, I analyse the features of each Systems approach and how they may help (or hinder) the analysis of EdTech technologies. I will use this approach to further unpack the nuances of teacher responsibility and examine both how responsibility (individual and collective) is enacted in schools and how it is framed by policy and regulatory environments (key components of a child's macro-systems), with the aim of expanding the ethical analyses introduced in the previous chapter.

5.2. Models

The next section explores various Systems Approach models proposed by Psychology and Education scholars, illustrating how the models have been used in practice to map dynamic (and unpredictable) spaces such as schools.

Following Bertalanffy's original definition (1968), a system is defined as a collection of interconnected elements operating within a network. Within Public Policy and Education research, the term Systems Approach is used very pragmatically to map the school in its entirety. However, researchers differ in which features of a school they include in their maps/models.

For instance, in the 1950s, the think tank, RAND Corporation, suggested that a school system includes actors, tools, and outputs (Digby, 1989). Applying the RAND corporation's model to the case of 'School Suicide Prediction' may look (very briefly) like the following:

1. Actors may include teachers, learning support workers, nurses, parents, software companies, students;
2. Tools may include digital tools and "traditional" mental health programmes;
3. Outputs may include monitoring results, and intervention(s) suggested.

By using the RAND Corporation's model, implementation scientists would use these features (actors, tools, and outputs), to determine how a singular intervention (such as a suicide prediction programme) should be incorporated within a school, as well as to evaluate whether or not said intervention was successful (and why). In many ways, a model built on the actor-tools-output framework could be built directly using the data presented in the previous chapter. Chapter 4 already provided an initial mapping of actors explicitly named by Tech companies. To extend this, I would simply need to identify additional relevant actors (perhaps by consulting school staff directories or relevant legislation) and attempt to trace the programmes' outputs. As discussed previously, however, it is difficult to get efficacy measurements from EdTech companies (including rates of false positives and false negatives), as these proprietary EdTech programmes do not often share this data with the public.

While the actor-tool-output formula was used extensively to map out school systems in Public Policy research during the 1960s and 1970s, its use is now largely discouraged for being simplistic. Namely, the model above is argued to oversimplify complex phenomena by breaking down the system into isolated components (Vembye & Jensen, 2018), and fails to consider additional influences (such as the school environment and the broader 'meso-system,' e.g., school ethos), which undoubtedly shape how technology is used.

To many scholars, reductionist critiques apply predominantly to the simplistic models being used at the time, rather than the Systems Approach as a whole. Therefore, to these critics, the limitation of reductionism could (and should) be addressed through an ongoing refinement of models, and furthering researchers' understanding of how educational actors are situationally embedded within broader social institutions (Honig, 2006; Opfer & Pedder, 2011)

Moving away from the actor-tool-output model, education researchers now utilise an 'embodied approach' and Bronfenbrenner's Ecological Theory (1979) to map schools and schooling systems. Bronfenbrenner's Ecological Theory (1979) is foundational in both education and psychology, and looks at child development within the context of five different interrelated systems: a microsystem, mesosystem, exosystem, macrosystem, and chronosystem.

Using Bronfenbrenner's Ecological Theory as a conceptual basis, Moore et al. (2021) argue that a schooling system map should not only include actors, tools, and output, but also consider the interconnectedness of private industries other than the school, and their place within broader education policies (e.g. state and federal level), thereby situating individual interventions within broader School Systems. According to this perspective, no singular intervention can be considered without understanding the school context in which it works, and no school can be considered without considering the larger policy or economic context. In this way, the concept of 'shared responsibility' becomes integral to my analysis.

Applying Bronfenbrenner's Ecological Theory and Moore and colleagues' (2021) model to the case of School Suicide Prediction may, therefore, also include the presence of private EdTech companies, as well as public policies and safeguarding initiatives such as KCSIE or PREVENT. In addition, by highlighting the mesosystem, Bronfenbrenner's Ecological Theory may also account for individual schools' ethos around mental health - for example, its approach to wellbeing education, promotion of positive mental health, behavioural standards, and anti-bullying practices, which all play a role in preventing suicide (WHO, 2021).

Acknowledging individual school-based factors, as well as the socio-cultural-historical complexity in suicide prediction, as well as EdTech, can further deepen this thesis and analysis, as it helps draw upon literatures around surveillance, race, disability, and law.

Figure 7

Adapted Version of Bronfenbrenner's Ecological Theory (Bronfenbrenner, 1979)

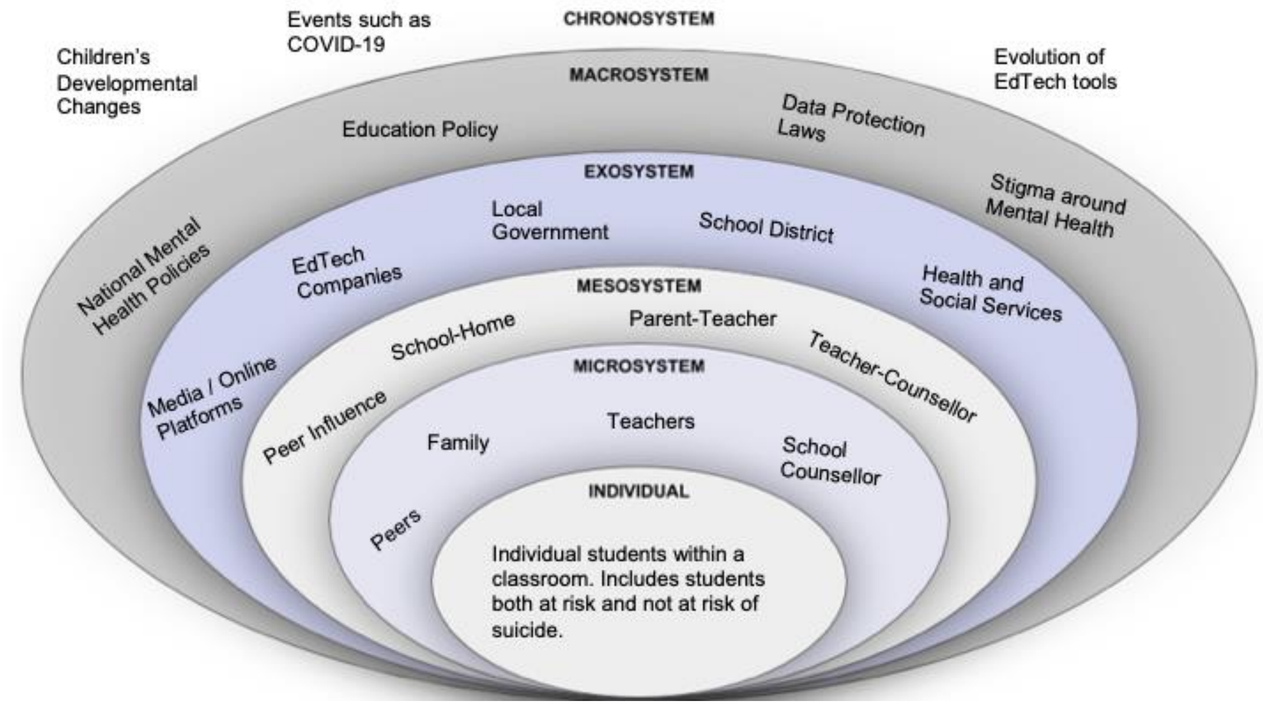


Figure 7 above represents an initial application of Bronfenbrenner's ecological model, illustrating how micro-, meso-, exo-, macro-, and chrono-level systems may influence suicide prediction for individual students within a classroom context. For instance, within this model I identify the role of frontline figures in children's lives, including: teachers who often observe behavioural cues and either input data into the predictive system or respond to alerts, and school counsellors and psychologists who are central to initiating interventions and providing direct psychological support.

Beyond school influences (e.g. this could include the school or school districts' ethos around mental health), this model also highlights the student's exosystem, including the role of EdTech companies and media or online platforms as critical actors. For instance, as illustrated in Chapter Four, EdTech developers shape the design of their tools, and as discussed in Chapters 1, 2, and 4, these design choices are influenced by financial and economic factors. These include shifts in funding (such as those resulting from COVID-19 stimulus packages (a chronosystem factor) and the broader influx of investment that has driven the evolution of EdTech.

While Figure 7 offers a valuable foundation for understanding the multifaceted and systems-based nature of student mental health within educational settings, within this thesis it is important to go deeper into the broader ecological complexity that underpins the use and impact of suicide prediction technologies in schools. In particular, the macro-system warrants deeper attention. The macro-system encompasses the cultural, political, and institutional structures within which all other systems are embedded (Bronfenbrenner, 1979). National education standards, mental health policy, and legal frameworks governing safeguarding, all play critical role in shaping how EdTech programmes are developed, deployed, and interpreted, as well as how responsibility is determined (and shared).

Therefore, Chapter 5 investigates the macro-system of EdTech for Suicide Prediction, specifically the UK educational and legal systems and how each engages with the use of EdTech for suicide prediction, focusing on the theme of responsibility. This includes exploring, by law and professional teaching standards (e.g. indicators of macro-level 'community-based' expectations of care), who is expected to use these technologies and what responsibilities are assigned to different actors.

Importantly, other levels of the model (e.g. micro-and meso-level teacher–student relationships, as well as the school's ethos) can only be understood through empirical research. Therefore, these systems will continue to be addressed in a separate chapter (Chapter 6), which draws on qualitative interviews to explore shared dynamics (e.g. teacher-teacher; teacher-parent; teacher-EdTech provider), and shared responsibility in more depth.

5.3. Search Strategy

To analyse the school system at the macro-level, with a focus on answering my key research questions around teachers' roles and responsibilities for using suicide prediction tools, I use consistent keywords across multiple databases to gather available policies within education and policy/law. A combination of search terms combining truncated and non-truncated terms was used to create this search protocol, as shown in Table 7. The search protocol also included a broader search, whereby only three of the categories needed to be met. In this instance, both the school and responsibility categories had to be met, while either technology or suicide was optional. For example, relevant literature around the ethics of suicide risk prediction (but not EdTech) or EdTech ethics (without suicide risk prediction) were both included.

Table 7

Keywords and Search Strategy

Keyword	School	Technology	Suicide	Risk	Ethics and Responsibility
Synonyms	School* College Student	Techn* App Ed-Tech Software Monitoring Device	Suicid* Self-harm SSI	Predict* Risk Universal Screening	Responsibl* Duty Role (also broadened out to ethics more generally)

I used different resources for each academic field, inspired by the Libguides made available by the Bodleian Library[20]. This included, but was not limited to, the following databases[21]:

1. Westlaw and LexisNexis for law
2. PsychINFO, Proquest, Web of Science, and SCOPUS for information within the fields of Psychology and Policy
3. The Department for Education website for specific policy documents, professional standards documents, and teacher training manuals

Documents were only included if they referred to secondary school teachers in the United Kingdom, with relevance to either mental health or monitoring technology. Documents also had to be either non-statutory advice, given by the government, professional standards, given by teaching associations, or law. Documents were excluded if they focused exclusively on primary or tertiary students, made no reference to mental health or monitoring technology, or were unrelated to the roles and responsibilities of teachers, such as those addressing only the perspectives of governors or parents. In addition, documents were excluded if they were not governing the UK

[20] https://libguides.bodleian.ox.ac.uk/

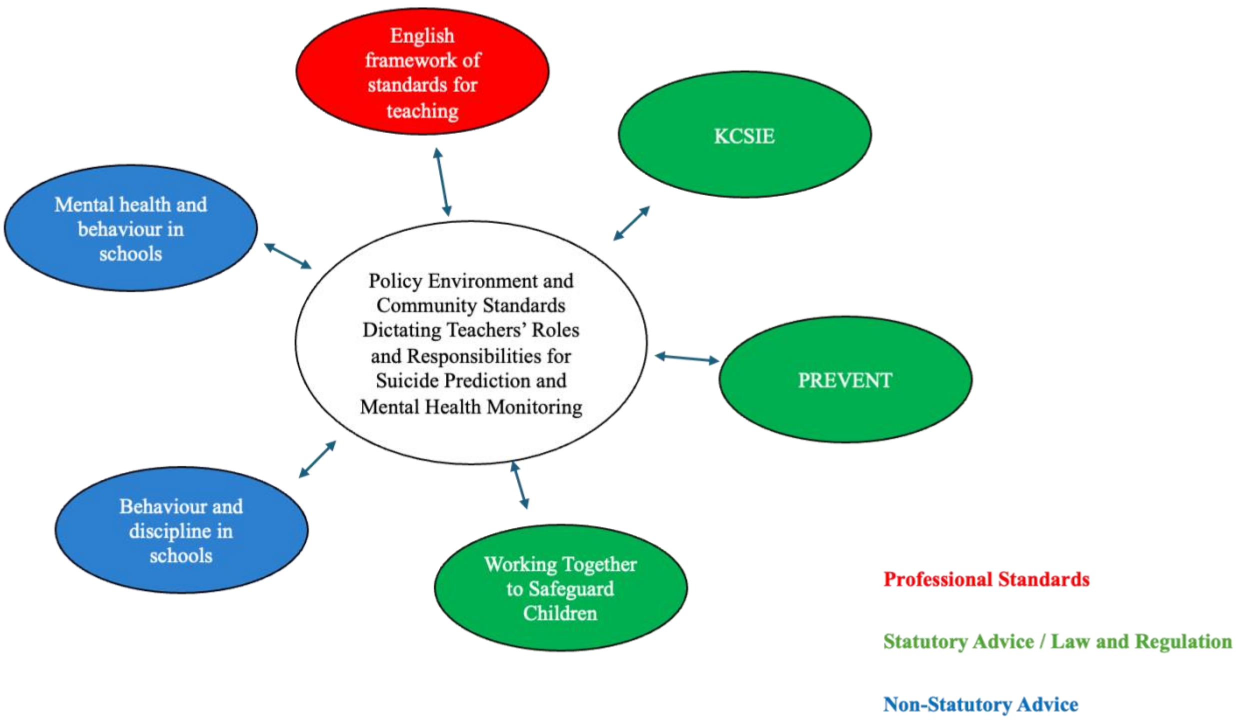
[21] These are only briefly discussed in the ToS, with a more detailed explanation of the specific searching strategies used with each resource available in the thesis.

education system on a macro-level (for example, blog posts by individuals were excluded, or documents focusing on non-UK countries).

This search strategy led to the identification and analysis of documents across three categories: non-statutory advice, professional standards, and law/regulation, described in Figure 8 below.

Figure 8

Policy Environment and Macro-System Dictating Teachers' Roles and Responsibilities for Suicide Prediction and Mental Health Monitoring in UK Schools



5.4. Findings

As illustrated in Figure 8 above, within this review six core documents were identified, where the responsibilities of a teacher for suicide prediction and mental health monitoring are outlined by non-statutory advice, professional standards, and law. These are discussed in greater detail below.

5.4.1. Non-Statutory Advice

In 2018 the Department for Education (DfE) published non-statutory advice about mental health in schools that is relevant to this thesis (Department for Education, 2018). This document, titled 'Mental health and behaviour in schools,' "sets out schools' roles and responsibilities in relation to mental health and behaviour, within their existing duties" (p. 3), and is linked to another document published by DfE in 2016, titled 'Behaviour and Discipline in Schools' (Department for Education, 2016). Key points in the former document affirm that the school plays a key role in supporting good mental health, that any strategy should be part of a whole-school approach, and that school staff should not act as mental health experts or work to diagnose any students (although the school should still have systems and processes for referral). This non-statutory advice is consistently backed up by statutory guidance, including KCSIE. It is important to note that schools are not required to hold any official mental health policy. Instead, on the basis of the documents analysed, holding an official mental health policy is only recommended.

While these documents are not statutory, they are very comprehensive and put an increasing emphasis on early intervention. This emphasis can be (and often is) used to support the introduction of EdTech monitoring tools. Key elements to early intervention, according to the 2018 report 'Mental Health and Behaviour in Schools' (Department for Education, 2018), include: prevention, identification, early support, and access to specialist support (p.6). In addition, this document outlines some tools that the Department believes could support identification and measurement, including: (1) effective use of data (suggesting the Strengths and Difficulties Questionnaire), and (2) an effective pastoral system (Department for Education, 2018, p.16).

Finally, the DfE's 'Mental Health and Behaviour in Schools' (Department for Education, 2018) also promotes collaborations with other agencies (e.g., other schools or specialist Children and Young People's Mental Health Services - CYPMHS), through developing a Local Transformation Plan (LTP) and going through the Director of Children's Services (DCS). The document promotes collective responsibility further by discussing working with parents and carers, as well as local authority Alternative Provision (AP). This is an important feature of this document, as it insists that teachers and school staff members do not take over medical roles, but rather enhance the work of professionals and take on a linking or referral role.

5.4.2. Professional Standards

The English framework of standards for teachers outlines a teacher's role in relation to their students' health and wellbeing (Department for Education, 2011). According to the DfE (2011), the English framework of standards for teachers outlines the minimum requirements of what it means to be a teacher in the UK, and these are the standards used to assess trainees and all teachers who are working towards, or have achieved, qualified teacher status (QTS). The national guidance, as used today, was first introduced in July 2011, before undergoing slight updates in 2013 and 2021. Going through the full professional standards framework, I have identified (and directly quoted) the following sections to reference students' mental health or safeguarding needs (Figure 9).

Figure 9

Relevant Sections of the English Framework of Standards (direct quote, Department for Education, 2011, pp.10-14)

Part One: Teaching

(1.a): establish a safe and stimulating environment for pupils, rooted in mutual respect

(5.b): have a secure understanding of how a range of factors can inhibit pupils' ability to learn, and how to best overcome them

(5.c): demonstrate an awareness of the physical, social and intellectual development of children, and know how to adapt teaching to support pupils' education at different stages of development

(5.d): have a clear understanding of the needs of all pupils, including those with special education needs; those of high ability; those with English as an additional language; those with disabilities; and be able to use and evaluate distinctive teaching approaches to engage and support them.

(7.a): have clear rules and routines for behaviour in classrooms, and take responsibility for promoting good and courteous behaviour both in classrooms and around the school, in accordance with the school's behaviour policy

(7.b): have high expectations of behaviour, and establish a framework for discipline with a range of strategies, using praise, sanctions and rewards consistently and fairly

(7.c): manage classes effectively, using approaches which are appropriate to pupils' needs in order to involve and motivate them

(7.d): maintain good relationships with pupils, exercise appropriate authority, and act decisively when necessary.

(8.e): communicate effectively with parents with regard to pupils' achievements and well-being.

Part Two: Personal and Professional Conduct

(1): Teachers uphold public trust in the profession and maintain high standards of ethics and behaviour, within and outside school, by:

(1.a): treating pupils with dignity, building relationships rooted in mutual respect, and at all times observing proper boundaries appropriate to a teacher's professional position

(1.b): having regard for the need to safeguard pupils' well-being, in accordance with statutory provisions

(1.c): showing tolerance of and respect for the rights of others

(1.d): not undermining fundamental British values, including democracy, the rule of law, individual liberty and mutual respect, and tolerance of those with different faiths and beliefs (1.e.) ensuring that personal beliefs are not expressed in ways which exploit pupils' vulnerability or might lead them to break the law

Children's mental health, wellbeing, and safeguarding needs are well integrated throughout the professional standards document (in both the teaching and personal/professional conduct sections). These standards are easily adaptable to each school and therefore can look very different in terms of practice. However, it is very clear that teachers have been given significant responsibility for students' mental health and wellbeing beyond basic teaching standards.

5.4.3. Statutory Regulation

Teachers have long had the right and responsibility to act *in loco parentis* while students are on school grounds. According to Besley & Peters (2019) this means that teachers and educational institutions have been given responsibility for "specific functions as a parent to act in the best interests of students without infringing on the civil rights of students" (p.181). According to many pieces of legislation, this is the responsibility of not just particular teachers, but rather all school staff. For example, KSCIE (Department for Education, 2024) says that: "Safeguarding and promoting the welfare of children is **everyone's** responsibility. **Everyone** who comes into contact with children and their families has a role to play" (p.7; words highlighted in original text). In addition, according to this regulation, "all staff should be prepared to identify children who may benefit from early help [...] all staff should be aware of their local early help process and understand their role in it [... and] all staff should be made aware of the process for making referrals to children's social care and for statutory assessments under the Children Act 1989" (Department for Education, 2024, pp. 8-9).

Within KSCIE, teachers are asked to be aware of students requiring 'early help.' It reiterates comments made in the non-statutory 'Mental Health and Behaviour in Schools,' where the Department for Education (2018) writes that:

> Only appropriately trained professionals should attempt to make a diagnosis of a mental health problem. Staff however, are well placed to observe children day-to-day and identify those whose behaviour suggests that they may be experiencing a mental health problem or be at risk of developing one. (p.11)

As illustrated by the above quote, while government guidance emphasises that teachers should not replace mental health specialists, it also recognises teachers as well placed to observe children and monitor their mental health. Beyond traditional classroom observations, KSCIE (Department for Education, 2024) also recommends for use online systems and monitoring in Annex C, outlining four areas of risk: content (being exposed to harmful material), contact (being exposed to harmful interactions with others), conduct (poor personal behaviour, e.g., online bullying), and commerce (risks including online gambling and financial scams). However, while KSCIE argues that it is up to schools to "ensure appropriate filters and appropriate monitoring systems are in place" (Department for Education, 2024, p.38), it also says that "the appropriateness of any filters and monitoring systems are a matter for individual schools and colleges and will be informed in part, by the risk assessment required by the PREVENT Duty" (Department for Education, 2024, p.40), which focuses on counter-terrorism. Therefore, nowhere in this document does the DfE mandate, or even recommend, monitoring for suicide risk (this is instead an additional capacity of some EdTech systems).

While not intended for suicide prediction, as discussed in Chapter 1, under the PREVENT duty, the Home Office mandates that UK schools conduct online monitoring and recommends the implementation of comprehensive online risk assessment processes (Home Office, 2024b). The purpose of this is "to help prevent the risk of people becoming terrorists or supporting terrorism" (Department for Education, 2023, para. 1). While the primary aim is countering terrorism and radicalisation, as mentioned in Chapter 1, the requirement for such technology often means it is additionally used for suicide prediction. This is also reflected in Chapter 4, where most EdTech companies used for suicide prediction highlight their compliance with the PREVENT duty (Appendix 1).

Under PREVENT duty guidance, teachers are required to monitor students and refer them to the PREVENT program if the students display signs of risk or vulnerability to terrorism or radicalisation (Bryan, 2017). Referral processes place a strong emphasis on shared responsibility, involving schools to collaborate with other institutions such as the police and local authorities. Specifically, according to the Department for Education (2023),

> To comply with the Prevent duty, schools, colleges and further education independent training providers must show evidence of: productive co-operation, in particular with local

Prevent staff, the police and local authorities; co-ordination through existing multi-agency forum (para. 4)

Across the Home Office and DfE websites, there is a clear emphasis on the need for schools to build partnerships with local authorities, safeguarding partnerships, and policing teams (Department for Education, 2023; Home Office, 2024b). Although the Home Office does not explicitly state this, it follows that technology used for both anti-terrorism and suicide prediction would also be subject to the same collaborative requirements with police and other agencies.

Finally, it is important to note that, while teachers are currently bound by statutory guidance such as KCSIE, there is currently no standalone legal duty specifically requiring teachers to report safeguarding concerns such as suicide (Foster, 2025). Therefore, teachers' current duties are enforced through institutional policies and professional standards rather than via direct legal obligation. However, the UK government plans to introduce a formal statutory duty of care for professionals working with children in the near future, including teachers (Foster, 2025).

According to a 2023 House of Commons Library briefing, and reported by Foster (2025), the proposed legislative changes aim to place a legal obligation on individuals to take appropriate steps to prevent or report harm, particularly in cases of child sexual abuse. If introduced, this duty is expected not only to formalise existing safeguarding responsibilities but also to provide a framework for other areas of mandated reporting, such as suicide risk or signs of serious mental ill-health.

5.5. Analysis

The above professional standards, statutory advice, laws and regulation, offer valuable insights into how responsibility is conceptualized by policy and professional communities within the UK. The following section analyses how these documents enhance, refine, or complicate the original definitions of role responsibility and shared responsibility outlined in Chapter 3.

Specifically, the first section of the analysis explores what responsibility looks like according to the documents listed above, and whether (/how) definitions of shared responsibility diverge from both Hart's original framework (as well as how responsibility is codified into the technology itself). The second section identifies potential mediators that may explain *why* there are gaps between ideal notions of responsibility (as seen in ethics and law) and how responsibility is operationalized in practice (including in the design of these technologies). This includes divergent underlying theories of education and a mis-interpretation of responsibility as accountability.

5.5.1. Shared Responsibility

As schools become increasingly complex environments, schools have embraced a whole-school approach to mental health and wellbeing, which is supported by current legislation and outlined in the 'Working Together to Safeguard Children' statutory guidance (HM Government, 2023). A whole-school approach ensures that, rather than the responsibility for safeguarding falling on only one teacher/counsellor, it is the responsibility of all school staff to work together with other organisations (e.g., health services, police), to safeguard students - thus engaging in a multi-agency approach (Public Health England, 2021). Multi-agency approaches could be particularly important for students with multiple or severe mental health needs who require multidisciplinary or specialist services.

Teachers are not the only ones involved in the process of mental health monitoring (which suicide monitoring falls under). Instead, there are multiple spheres of influence, including industry, police, families, the NHS, and government. For example, in industry, there is a commercial ecosystem that already exerts considerable influence on schools and the process of monitoring suicide. This includes UK based companies such as Impero (Ativion), as well as US based companies Social Sentinel and Gaggle.

Regarding the police, there has been a substantial increase in the number of school officers put into schools, both in the USA[22] and the UK, in part in response to the PREVENT duty. Not only has there been an increase in the number of school police/safety officers, there also has been an increase in their role and remit. There are many written data-sharing agreements between law

[22] According to the National Association of School Resource Officers, school-based policing is the "fastest growing area of law enforcement" within the USA (Canady et al., 2012, p.1).

enforcement agencies and school districts, as well as with families and parents (Collins et al., 2021). As more groups become involved in monitoring mental health, models of responsibility should become equally more complex.

Therefore, while Chapter 4 outlined an initial map of shared responsibility with key actors limited to teachers, parents, students, EdTech companies, and occasionally the police, Chapter 5 reveals that the network of responsibility is broader than initially mapped. For instance, UK law and regulations suggest that additional parties such as specialist Children and Young People's Mental Health Services (CYPMHS), and local authorities, are also involved in developing children's care, and one way to do this is with different types of multi-agency approaches (explored in greater length in Chapter 7).

On paper and in regulation there are some clear divisions of responsibility and labour. For example, teachers are responsible for referral if they believe a student is at risk of poor mental health or suicide, however they are not legally responsible for assessment, diagnosis, or social and emotional care. Indeed, guidance is clear that only appropriately trained professionals should attempt to diagnose students' mental health conditions. However, these seemingly separate responsibilities in practice are often blurred (Fischman et al., 2006), e.g. as illustrated by the gap between responsibility as reported by tech companies (Chapter 4) and responsibility as per codified into UK regulation (Chapter 5).

The implications of this 'responsibility gap' are explored further in Chapter 7, which examines the concept of shared responsibility in greater depth. While that chapter offers a more detailed discussion of the gap between normative ideals of shared responsibility and the current reality of how responsibility is distributed, two preliminary hypotheses are introduced here.

The first hypothesis suggests that there may be two conflicting roles assigned to teachers: *teacher-as-academic-educator* and *teacher-as-one-who-surveilles*. This tension may fundamentally reshape how teachers perceive and enact responsibility in the context of suicide prediction. This theory is introduced below and will be further examined through empirical research in Chapter 6.

5.5.2. Hypothesis One: Divergent Roles / Theories of Education

An emergent theme that may shape the kinds of responsibilities teachers hold in relation to EdTech for suicide prediction and help explain why these responsibilities diverge from normative ideals, is the role conflict between *teacher-as-academic-educator* and *teacher-as-one-who-surveilles.* In schools, surveillance can take several forms. Teachers may monitor students' activity on school devices during lessons, a task that can conflict with their instructional role. By contrast, observing students' behaviour and presentation in the classroom is a long-established aspect of teaching, integral to maintaining safety, wellbeing, and a learning-conducive environment.

In this chapter, however, I focus specifically on the surveillance of students' online behaviour conducted by digital monitoring software, particularly systems that flag concerning keyword searches, as discussed in Chapter 5.

Although maintaining a safe classroom has long been part of teachers' professional responsibilities, under PREVENT duty guidance, teachers are required to explicitly monitor students and refer them if they express a risk or vulnerability of terrorism/radicalisation (Bryan, 2017). As such, it is up to the teacher to decide who is at risk. This duty, and the success of teachers in performing it, is monitored by the Office for Standards in Education, Children's Services, and Skills (Ofsted). Stanley et al. (2018) have argued that the PREVENT Duty has increased the 'securitization' of the previously non-juridical school, and Arthur (2015) shows that this means teachers have become increasingly responsible for security, thus creating conflict with their other, educational duties.

It is important to consider the impact of PREVENT within this thesis, as much of the EdTech brought about for mental health and suicide monitoring has its origin in PREVENT and supporting anti-terrorism initiatives in schools. Specifically, the PREVENT Duty has had unintended consequences regarding the monitoring of many behaviours beyond extremist activity. This is, according to Hope (2019), a form of 'concept creep,' where "categories such as bullying and violence have expanded their meanings so that they now include a much broader range of activities than before" (p.59).

Those against the use of these technologies often invoke Foucault's conceptualisation of school as a form of surveillance. Researchers including Taylor (2013) have argued that, while surveillance tools have been marketed as 'tech for safety' or 'tech for mental health,' instead we should see them through the lens of 'tech for surveillance' or 'tech for social control.' Rule (1974) therefore defines the role of the teacher not as academic-educator nor as one who safeguards children, but instead one who uses different tools to "discourage or forestall disobedience … [to] either punish such behaviour once it has occurred or prevent those with inclinations to disobedience from acting on those inclinations" (Rule, 1974, p.19).

A second example of 'concept creep' is where teachers may be gaining additional responsibilities from the medical field. Although above, I illustrated that statutory (and non-statutory) regulations are clear that teachers should not take up any official diagnostic or intervention role, in prior qualitative research teachers have argued that they are taking up increasing safeguarding and mental health responsibilities. For example, a teacher is quoted in Fischman et al. (2006) saying: "we are expected to be parent, psychologist, and then teacher, and that's very difficult …" (p.387). In this instance, teachers take on additional social duties beyond their legal or statutory responsibilities (and beyond the normative ideal) because their unique skills and knowledge allow them to recognise preventable risks that other professionals (e.g., a GP) or lay-people (e.g., parents) might not identify. However, as shown in this Chapter, there remains no consensus on the methods or thresholds for determining when this supplementary role is inadvisable, advisable, or mandatory (Levinson et al., 2020).

All these examples demonstrate how the role of a teacher sometimes shifts towards being responsible for medical care (including mental health care), surveillance, and/or 'crime punishment' rather than education, and that teachers' roles may be regularly changing and adapting based on function or concept creep. My hypothesis (that the divergence between normative and real-world views of shared responsibility stems from differing definitions of the teacher's role), will be explored in greater detail through the empirical interviews presented in the next chapter.

5.5.3. Hypothesis Two: Neoliberalism and the Move to Accountability

The previous section hypothesised that intended-actual divergences in teacher's responsibility (specifically: the difference between policy, which is clear that teachers are not required to provide any official diagnostic or intervention role, and practice, in which teachers are often at the forefront of both of those tasks, and also being held accountable if errors occur), stem from long-standing differences in how teachers' roles are perceived (for example, is their role to surveil or to educate). My second hypothesis, outlined in Section 5.5.3., suggests that these divergences are also rooted in the economic and political realities that continue to shape both educational technologies and the teaching profession today.

Besley and Peters (2019) have written that the shift from liberalism (autonomy/moral agency) to neoliberalism (market accountability) has triggered a further shift in responsibility to accountability. In their work, they argue that this shift has resulted in both conceptual and operational ambiguity within the responsibility field, ambiguity which has changed our understanding of responsibility away from the moral/legal, and towards regulation/accountability. Besley and Peters (2019) go on to describe four forms of accountability, including state-mandated agency (regulated from a state/federal level), professional accountability (through professional associations), consumer accountability (through the market), and democratic accountability (to the electorate).

While not the primary research question of this thesis, a question that may arise from this shift is: if teachers, students, clinicians, tech companies etc., are all invested with responsibility, who should be held accountable in the case of system failure (e.g., if a student dies by suicide, or if the algorithms deprive students of their privacy)? Not only who should be held accountable, but by what process (e.g., democratic agreement, through professional associations, through the market)?

5.6. Conclusion

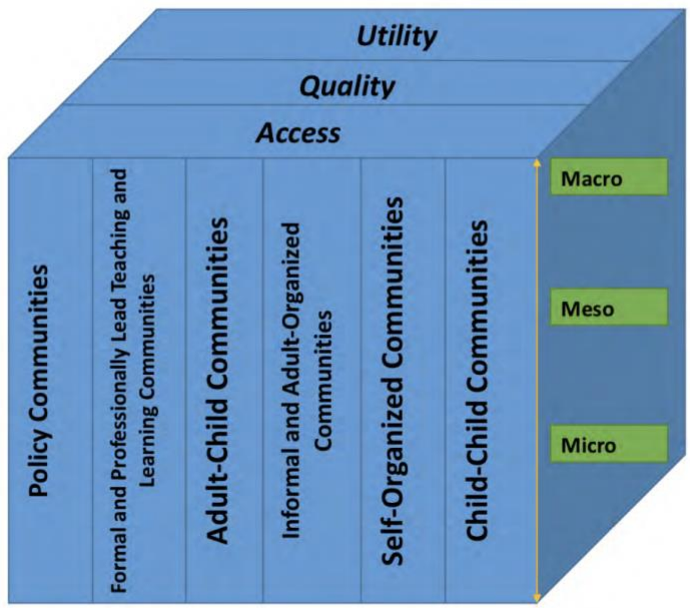
Overall, this chapter introduced Systems Theory and began to identify the macro-level policies that regulate how teacher and policy communities engage with suicide prediction technologies. The analysis then drew on normative theory to help explain differences between teachers' responsibility as defined by the technology companies themselves, versus law and policy. This chapter then suggested hypotheses as to why there were tensions between normative definitions of responsibility and how responsibility is embedded within the technology itself. This included two hypothesised mediators: (1) how a teacher's role is conceptualised from a philosophical standpoint, as well as (2) the political and economic environments in which school-based suicide prediction takes place.

This chapter only considered the macro-level of Bronfenbrenner (1979)'s systems theory, which is a significant limitation due to time constraints with the DPhil. For instance, it fails to consider the mesosystem, i.e. the interactions between different microsystems, such as the relationships among peers, parents, and teachers, which may influence or shape the school ethos (and the school's culture around mental health). Another limitation of the analysis above is that it relies on a relatively simplified model of systems thinking - primarily listing relational elements such as policies, laws, and actors. While this level of analysis was appropriate for the purposes of this thesis, future studies may benefit from employing more complex and nuanced forms of systems analysis. For example, another model that could be used is Schuelka and Engsig (2022) 'Complex Educational Systems Analysis' (CESA). According to the authors, this is a way to map interventions not only by their elements, but also by considering the evolving nature of resources, infrastructure, and contextual changes (also discussed in Berry et al. (2018). This may also include dynamic behaviour and feedback loops among these elements (Figure 10).

Figure 10

Complex Educational Systems Analysis (Schuelka & Engsig, 2022, p.6)

Figure 1
The Complex Educational Systems Analysis Cube (CESA³)



Within this model, and under the guidance set out in Schuelka and Engsig's (2022) paper, EdTech for suicide prediction could be analysed in the following way:

1. *The multilevel dimension (micro, meso, macro)*: This is influenced by Bronfenbrenner's Ecological Theory (1979) and encourages researchers to consider the embeddedness of an individual child into multiple systems and consider the relationships between the two. For example, how parents might be interacting with their child, and inputting into a EdTech programme (micro-level); how teachers and parents may interact with one another in order to read the results of the EdTech prediction or create a school culture around mental health (meso-system), or how an EdTech programme may relate to a school's official mental health policy, specific laws, and national policies such as KCSIE (macro-systems).
2. *The community dimension:* According to Schuelka and Engsig (2022), it is important to consider not only an individual classroom or school as a learning community, but also consider the existence of less formal communities. The authors argue that a successful Complex Educational Systems Analysis should be able to map out the following

dimensions of community: policy communities, formal and professional-led teaching and learning communities, adult-child communities, informal and adult-organised communities, self-organised communities, and child-child communities. With regard to the example of EdTech for suicide prediction, adult-to-child community may include interpersonal interactions between students and teachers during recess and how/if these are being monitored by technology. Child-to-child community could include how peers might be able to individually report each other within the application, or how the technology monitors interactions on social media.

3. *The educational attribution dimension (access, quality, utility):* While originally conceptualised within the field of Disability and Inclusive Education, this dimension may include enrolment and attendance rates, as well as an insight into which students the technology best supports (and which students may be underserved). In addition, it may bring into question whether or not the EdTech programme is effective.

Ultimately, Systems Analysis, as operationalised within Schuelka and Engsig's (2022) CESA approach, could help future researchers expand my initial analysis, and explore how Suicide Risk Prediction technology is being used *in situ* within UK schools. Furthermore, the CESA approach can also be used as a framework to explore responsibility. For example, by examining each of these levels (e.g. educational attribution dimension; community dimension; and the multi-level dimension), I may have been able to develop a deeper ethical understanding on how responsibility is held (and ought to be held) within UK schools. However, while I acknowledge the simplicity of my initial model and suggest that future studies could adopt a more complex approach (e.g. the one proposed by Schuelka and Engsig (2022) discussed above), I maintain that the original model was sufficient for the aims of this analysis. In fact, increasing complexity may not have necessarily strengthened the analysis in a directly or positively correlated way. In addition, within this thesis, I do not have access to the data needed to conduct a full analysis. This is primarily because of time limitations, as well because the educational attribution dimension (point 3) involves data to which I currently do not have access. As mentioned in the previous chapter, a key limitation of these software programmes is that they are proprietary, and therefore there is limited public data on their efficacy. As such, this aspect of the analysis is currently on hold.

Chapter 6. Interviews

In Chapter 3, I outlined the methodological rationale for empirical work, situating qualitative interviewing methodology within my larger empirical bioethics approach, in order to achieve a deeper understanding of teachers' views on suicide-prediction technologies. I explained how semi-structured interviews might inform and refine normative analysis by providing context on how teachers think and act in real-life situations. I also explored how interviews can be used to provide 'thicker' descriptions of normative topics; for example, interviews may help researchers understand teachers' values, and how these values interact with larger ethical concepts such as responsibility (for which I provided philosophical definitions developed in earlier sections of this thesis). Finally, I argued that these interviews may also help both generate new hypotheses, which could be tested and explored further through other methods (e.g. policy analysis, additional interviews, and/or normative work), as well as resolve hypotheses generated in earlier chapters (e.g. the mapping review (Chapter 4) and the analysis of law and policy documents (Chapter 5)).

Therefore, this chapter presents findings from an empirical project that investigated the values and preferences of teachers regarding the use of EdTech to monitor young people's risk of suicide online, answering the question: *what responsibilities do teachers and schools have when using technology for suicide prediction, and how do these compare to what teachers think their responsibilities should be?* The empirical project presented in the current chapter is key to answering the above research question, as it enables me to identify gaps between theory and practice, to develop more practical ethical guidelines, and to test the hypotheses generated in earlier chapters.

While the primary goal of this chapter is to contribute empirical findings toward answering the research question outlined above, there are also important political and social rationales for this project's empirical work (both for the inclusion of empirical work overall, and also for the specific methods chosen). As Stromquist (2006, p.14) notes, the methodologies we choose "change the distribution of power both in interpersonal relationships and in institutions throughout society." The question, then, that I grappled with within this thesis, was how to choose and enact methodologies which challenge existing power structures.

Using qualitative methodologies does not inherently shift power, but there are two ways in which I tried to do so in my research. First, while I did not interview students, I worked to include the voices of students in this project. The study design was informed by a co-production approach, particularly through consultation with the NeurOx Young People's Advisory Group (YPAG) in 2021 and 2022 (Pavarini et al., 2019). The NeurOx YPAG is a group of young people (ages 14-22), based at the Department of Psychiatry, University of Oxford, who worked with me to develop research questions and methodologies. They were involved throughout the early stages of the project, shaping the overall direction of the research, contributing to the development of the research question, and helping to design the interview guide. Their participation went beyond providing feedback on study materials. For example, they engaged in substantive discussions about how digital technologies operate within their own schools and shared key concerns and recommendations for how teachers should navigate these systems. Ultimately, even though students were not directly interviewed in this project, involving young people in the research design gave them a degree of influence over how the issues that affect their lives are studied and discussed.

The second approach involved incorporating a knowledge-building element into the interview process, and recruiting a range of teachers, including those who hold less formalised power within their schools. Specifically, I aimed to not only include teachers who actively have a say in whether or not these EdTech programmes are used in schools (e.g. head teachers), but also teachers who are less involved in decision making, in order to raise awareness about the existence of EdTech tools, and ethical implications of their use.

Both of the methods listed above were important to include within my research, and in my attempts to use this thesis to "change the distribution of power," as per Stromquist (2006, p.14). Foucault (1980) has argued that knowledge and power are intimately linked: those who control knowledge also shape what is considered legitimate authority. Therefore, building knowledge within both the YPAG and teacher populations was one of my key goals. Ultimately, within the UK education system, where top-down policy decisions can have significant real-world consequences, and where awareness of these tools is generally lower than in the US (as evidenced by fewer news reports), increasing understanding among teachers and students, and bringing their voices into dialogue with policy through empirical bioethics methodologies, enhances both the ethical integrity and practical relevance of the research. And while this is by no means a fully participatory research project, the elements of power-shifting it incorporates (as listed above) are valuable.

In what follows, I present the methodology and results of the analysis of the qualitative data, to present teacher's attitudes on responsibility within the context of new EdTech for suicide prediction. In Chapter 7 I conduct a philosophical and conceptual analysis of the data and draw normative conclusions about how ideas of responsibility underpin schools' use of new technologies for suicide prediction.

6.1. Methodology

6.1.1. Recruitment

Ethics approval was granted by the University of Oxford's CUREC Ethics Committee on February 2nd, 2022 (approval number R78840/RE001; both my formal CUREC approval and consent form template are included in the Appendix (Appendix 1 & 2)). Recruitment began shortly after February 2022, with the interviewing period lasting 18 months. Teacher participants were recruited through word of mouth and by directly emailing teachers on a school recruitment list developed by the Department of Psychiatry team for previous studies. As per my CUREC application, the inclusion and exclusion criteria for recruitment were as follows (Table 8).

Table 8

Inclusion and Exclusion Criteria

Inclusion Criteria	**Exclusion Criteria**
Hold a contract in a secondary school in the UK	Does not work in a secondary school in the UK
Holding any safeguarding role or capacity within their school	Is unable to complete the interview in English
Has access to a digital device and WiFi connection to complete the interview online	Does not have access to a digital device or WiFi connection to complete the interview online

It is important to note that recruitment took place during periods of COVID-19 restrictions and was also impacted by a period of personal leave, both of which significantly affected the pace of recruitment (and the former required interviews to be held online).

Recruitment ended based on four key factors: (1) achieving a clear mix of school types, (2) including both classroom teachers and senior leadership, (3) including a wide range of teaching experience, and (4) reaching a stage where interview responses consistently addressed all core research questions in full, and no new themes were emerging (Faulkner & Trotter, 2017; Francis et al., 2010; Morse, 2015). To achieve the first three components, I identified participants' roles in advance through preliminary informal emails in which I asked prospective participants to outline their current position (e.g., classroom teacher, head of department, assistant principal). This enabled me to monitor the emerging composition of the sample and ensure a balanced cohort across role types before concluding recruitment.

With these recruitment goals in mind, I included a total of ten participants in the study. An abbreviated table summarising the first three recruitment criteria is provided below (Table 9), followed by additional detail, with the full participant table available in the Appendix 5.

Table 9

Abbreviated Participant Summary

Participant ID	**School Type**	**Job Title**	**Years Teaching**
10	Community / local authority maintained	Assistant Head Teacher	7
11	Community / local authority maintained	English Teacher	5
12	Academy	Head of History	7
13	Private	Teacher	5
14	Private	Biology and PSHE tutor and SENCO	27
15	Grammar	Teacher of Religious Studies, EDI Student Coordinator	8
16	Private	Deputy Principal, Teaching + Learning	14
17	Private	Head of Religious Studies	9
18	Academy	Head of Sixth Form	8
19	Free School	Deputy Head of Sixth Form and Teacher of English	11

As per the first goal: participants were employed in various types of schools: community or local authority-maintained schools (n = 2), academies (n = 2), private schools (n = 4), a grammar school (n = 1), and a free school (n = 1). All participants worked at the secondary education level, with two participants involved in a mix of primary and secondary education.

As per the second recruitment goal: the job titles of the participants varied widely, including (1) an Assistant Headteacher, (2) an English Teacher, (3) a Head of History, (4) a general Teacher, (5) a Biology and PSHE Tutor and SENCo, (6) a Teacher of Religious Studies and EDI Student Coordinator, (7) a Deputy Principal for Teaching & Learning, (8) a Head of Religious Studies, (9) a Head of Sixth Form, and (10) a Deputy Head of Sixth Form and Teacher of English.

As per the third recruitment goal: teaching experience among the participants ranged from five to twenty-seven years, with two participants each having five, seven, or eight years of experience, and others having nine, eleven, fourteen, or twenty-seven years of experience. The majority held a PGCE certification (n = 8), with additional certifications including a School-Centred Initial Teacher Training (SCITT), Teach First PGCE, and a MSc (Learning and Teaching).

While the first three recruitment goals aimed to ensure a diverse range of perspectives in the qualitative interviews, achieving a fully representative sample was never the objective. Instead, the primary reason recruitment concluded was that I achieved thematic saturation. Importantly, this study focused on achieving thematic rather than data saturation, meaning that recruitment ceased once no new themes (rather than no new details) were emerging from the interviews.

While there is no universally agreed-upon method for determining when saturation is reached, several researchers have provided useful guidance. For instance, Glaser & Strauss (2017) define theoretical saturation as the point at which “the researcher becomes empirically confident that a category is saturated” because they have “seen similar instances over and over again” (p. 61). Building on this, Morse (2015, pp.587–588) writes that “we do not saturate particular details of individual events and random incidents. Rather, we saturate characteristics within categories.” Making this distinction is critical: while participants may express themselves differently, and therefore numerous codes may arise when conducting qualitative analysis, what matters for saturation are the recurring themes, not individual codes.

Furthermore, it is critical to note there are limitations to achieving true saturation. First, time-limitations, which are particularly relevant to a DPhil project. This is what Green & Thorogood (2010) and Guest et al. (2006) refer to as the limitation of coding during 'funded work.' Time is a particularly relevant limitation, as Green & Thorogood (2009) argue that the point of saturation may be "potentially limitless" (p.120).

With this limitation in mind, I followed the guidance of Francis et al. (2010) who recommended an initial sample of ten interviews (which I achieved). When interviewing these first ten participants, qualitative analysis was conducted iteratively after every two interviews, and recruitment was concluded once no new shared beliefs or themes emerged.

A full explanation of my analysis methodology is provided later in this chapter, but with regard to concluding recruitment, it is important to note that by the fourth interview, I had identified 10 preliminary themes, and by the sixth, this had increased to 13 - the final number of themes I ultimately identified. After the sixth interview, no new themes emerged, and this remained consistent throughout the remaining interviews. Therefore, while I cannot claim this is a fully representative sample of all UK teachers, I met the requirements outlined by Francis et al. (2010) and ended my study after ten interviews.

6.1.2. Additional Demographic Information

A full demographic breakdown is available in Appendix Five. However, particularly following up on the first three saturation goals, it is also worth noting that, in an effort to ensure diversity, I aimed to balance the sample by including schools with varying proportions of students eligible for Free School Meals (FSM), as well as those with and without a specific focus on Special Educational Needs (SEN). Concerning the provision of FSMs, three participants indicated that their schools provided FSMs, six participants reported that their schools did not, and one participant did not specify. Four participants were involved with students who had SEN, while six were not.

Of course, these efforts do not guarantee that I have captured *all* types of teachers, or *all* types of schools. Two teachers with the same job title may have very different responsibilities depending on their school and its policies. However, given the limitations of this study, the recruiting approach offered a reasonable attempt at diversity, and reaching saturation suggests that this was at least partially successful. Future research could build on these recruitment numbers by expanding recruitment strategies or increasing the sample size.

Finally, there were some limitations in recruitment that, while not critical for the purposes of this DPhil project, should be acknowledged and may warrant further consideration in future research. These include factors such as age, race, geographic location, and parental status. Specifically, the sample skewed younger, with most participants falling within the 25–34 age range (n = 6), which may reflect participants' greater ease with or exposure to digital technologies (in comparison to their older counterparts). The gender distribution included seven women and three men, and in terms of ethnicity, nine participants identified as White British or White English, and one as Chinese. Geographically, the sample was concentrated in the South East of England (n = 6), with the remainder from London (n = 2), the South West (n = 1), and the West Midlands (n = 1). These variables were not systematically accounted for but could meaningfully shape teachers' perspectives and experiences. Regarding parental status, only one participant reported having children, eight did not, and one was unknown. This final point is particularly noteworthy: parental status may shape how teachers perceive their responsibilities toward young people, potentially giving them a 'dual role' (of teacher/parent) that could influence their responses.

6.1.3. Interview Sessions

Interviews were conducted via Microsoft Teams, led by myself, and lasted approximately one hour. Each interview consisted of six core questions and a brief case study. Case studies are an established tool in empirical bioethics research and in qualitative study designs (Ulrich & Ratcliffe, 2007). A full version of the interview guide is available in the Appendix (Appendix Four).

Some questions were adapted from a prior focus group guide developed by Nadeem et al. (2011), but as this is an empirical bioethics project, the primary focus was on the concept of responsibility, drawing on both the normative and empirical insights from the previous chapters. Questions explored teachers' roles broadly, while paying particular attention to the two themes identified in the preceding chapter: (1) the expanding expectation for teachers to assume non-academic and pastoral responsibilities, and (2) the prioritisation of shared responsibility between schools and wider social and mental health services. In addition, although the interviews were not designed to evaluate teachers' technological expertise, I did briefly ask participants at the beginning about their knowledge of educational technologies used in their schools. Only two teachers were explicitly aware of the automated system that formed the focus of this thesis, while all others mentioned familiarity with alternative programmes such as CPOMS (Child Protection Online Monitoring System), but not with automated analytics or predictive tools. Including this question aimed to contextualise their perspectives on responsibility, without shifting the overall analytical focus of the DPhil away from its central conceptual lens.

Finally, in addition to asking teachers to reflect on the frameworks of responsibility outlined in earlier chapters of the thesis, I also investigated mediating factors that might explain the gap between normative ideals and practical realities. For example, Chapter 5 began to identify structural and contextual influences (e.g. such as role ambiguity and political-economic constraints) which may act as mediators. Ultimately, the interview format allowed for further exploration of these mediating factors. This helped in developing a richer, more nuanced conceptual framework around responsibility in school settings. With this in mind, the following section outlines the main interview themes, each accompanied by a sample question. The complete interview guide is included in Appendix Four.

Table 10

Condensed Interview Guide

Theme	Sample Question
Mental Health experience	What experience do you have with student mental health in your school?
Definition of Responsibility	In my readings, there is a lot of discussion about being 'responsible' for students' mental health. I've been trying to understand this in practice. What does "being responsible" for your students' mental health mean to you?
Source of Definition	How do you decide what you are responsible for, and where did you get this information from? Has this been covered in any of your training? For example, inset days, teacher training, university course(s).
Technology	Does your school monitor students' digital activities? If so, which part of the "pipeline" is this technology used for? Detection of student risk? Support during a crisis? Post-crisis support?
Data Handling	What kind of information is gathered by this software? (e.g. social media data, google searches)
Introduction of Case Study: *Your school has been using the software CommonX to filter online content (e.g. making sure students don't access inappropriate websites in school). You receive an email update from your IT manager, saying that CommonX has a new add-on: one which safeguards students and can provide suicide risk predictions. Its use will be covered by the general IT privacy agreement which parents signed earlier in the school year. Later in the week, on a Thursday evening at 8pm, you receive a message to say that one of your students is deemed high risk of suicide. When you click the link attached to the email, you discover that this student has googled "why does someone commit suicide" on their personal home computer.*	
Immediate Intuitions	What are your initial thoughts on this case study?
Risk and Risk Thresholds	The case study talks about a students' "risk of suicide." What does the phrase "risk of suicide" mean to you?
Protocol	How would you respond to a notification telling you that one of your students is at high risk?
Shared Responsibility	When you are dealing with student mental health, how often do you work with the NHS/police/other social services? Which group do you work with the most/least? (Rank: parents, NHS, the students themselves, tech companies, police/other social services)
Accountability	Who do you think should be held accountable if something goes wrong? (Includes examples of models of accountability to respond to)

It is important to clarify that although I drew upon the case vignette to highlight key issues for teachers, I am not analysing their individual responses to this vignette. Rather, the vignette served as a technique to ensure that teachers without direct experience of this technology could engage with a concrete example, and that all participants were considering responsibility in relation to the actual technological capacities of EdTech for suicide prediction.

6.1.4. Data Analysis

The interviews were transcribed by a professional transcription company. I then reviewed all transcriptions by listening to the recordings. This allowed me to both make necessary edits (e.g. if the original transcript was unclear or incorrect) and help familiarise myself with the data. After this step, the Word documents of the transcripts were then imported into the qualitative research software NVivo 11 for data analysis. Using Braun & Clarke (2022) approach to thematic analysis, and drawing upon support from previous NEUROSEC team members (e.g. Manzini, 2020), I employed both top-down and bottom-up coding strategies: the top-down approach was guided by the frameworks of responsibility (both individual and shared) developed in earlier chapters of this DPhil, while the bottom-up approach drew on principles from Grounded Theory methodology.

Grounded Theory methodology (Charmaz, 2017), includes a step-by-step coding process which I used to gather and explore teachers' insights on the ethical issues related to the implementation of monitoring systems for suicide risk. Specifically, as per the process outlined in Charmaz (2006) during the qualitative data analysis, I repeatedly reviewed my transcripts to become thoroughly familiar with the material. I also took memos to record reflections (both during the interview process, and after all interviews were complete; using operational approach by (Corbin & Strauss, 2008; Chametzky, 2023)), highlighted unclear arguments, identify emerging categories, and considered connections between categories, as well as between the empirical work and relevant literature on responsibility (including philosophical literature, and policy documents outlined in Chapter 5).

To further analyse the transcripts, I employed 'open coding,' and used brainstorming to develop codes that allowed for multiple interpretations of the data (Corbin & Strauss, 2008). After open coding, I then used 'conceptual coding,' which is the process of grouping the open codes into broader categories (Corbin & Strauss, 2008).

Due to the structured nature of my interview guide (Table 10), I started this process by dividing my open coding into two sections. In Section One, I created a coding scheme to describe the participants' roles in schools (both individual and shared), and in Section Two I explored their arguments for or against the level of responsibility which they hold for students' mental health within the school.

Then, during the conceptual coding phase, my coding scheme for Section One was split into two further sub-sections. The first included type of responsibility (e.g. prediction, prevention, and referral), and the second explored the framework of responsibility which was used by the teacher-interviewees (e.g. no framework, legal framework, beyond legal framework). Section One's focus was on what participants described as their roles in schools. In contrast, Part Two focused on why participants perceived responsibility as they did, and the coding for this involved more abstract names (taken from earlier chapters and top-down data from philosophy and education, particularly as it relates to barriers already identified within the thesis).

While engaging in both open and conceptual coding (particularly open coding), I used 'analytic induction,' to validate findings (Silverman, 2017). Specifically, analytic induction includes constant comparison, as well as deviant-case analysis, which allows me to compare data within and between transcripts to test my research questions and emerging hypotheses.

Constant Comparison: Within qualitative methods, constant comparison is used to analyse data continuously throughout the research process (Glaser & Strauss, 2017). It is a central feature of grounded theory but is also applied in other qualitative methodologies. The idea is to continuously compare pieces of data with one another to identify patterns, similarities, and differences. Pieces of data can include incidents, quotes, and/or coding concepts, either between or within interviews. It is important to note that this is not a linear method, but rather an iterative process, where when new data are collected, they are immediately compared with previous data to identify evolving patterns (or contrasting pieces of evidence) (Silverman, 2017; Glaser & Strauss, 2017).

Deviant-case analysis: This is sometimes referred to as negative-case analysis (Hanson, 2017). Deviant case analysis is a methodology which enables me to explore interviews (or specific data points) that deviate from the norm or contradict the initial patterns or hypotheses that are being developed in my earlier analysis (Hanson, 2017). The goal for this is to constantly challenge (and thus improve) my overall hypothesis.

In addition to using constant comparison and deviant-case analysis within my interviews, the final step of my methodology was in line with Dunn et al. (2012) who suggest that participant arguments (in this case, teacher arguments) must be subject to regular philosophical scrutiny. This simply meant checking whether participants' claims aligned with the core philosophical definitions developed earlier in the thesis; where a view clearly conflicted with these conceptual foundations, it was (a) recognized, but (b) not used to shape the normative argument. This ensured that the empirical data informed, but did not determine, the ethical conclusions.

Ultimately, the above methodological process allows for the integration between empirical and philosophical data, and ensures that, at the end of analysis, I have a coherent understanding of teachers' attitudes on how EdTech should be used for suicide prediction (an understanding that is in line with both universal bioethics standards and empirical data on the teachers' views). This, in turn, hopes to answer the thesis's primary research question: What responsibilities do teachers and schools have when using technology for suicide prediction, and how do these compare to what their responsibilities should be?

Figure 11

Coding Scheme, What are Teachers' Roles and Responsibilities?

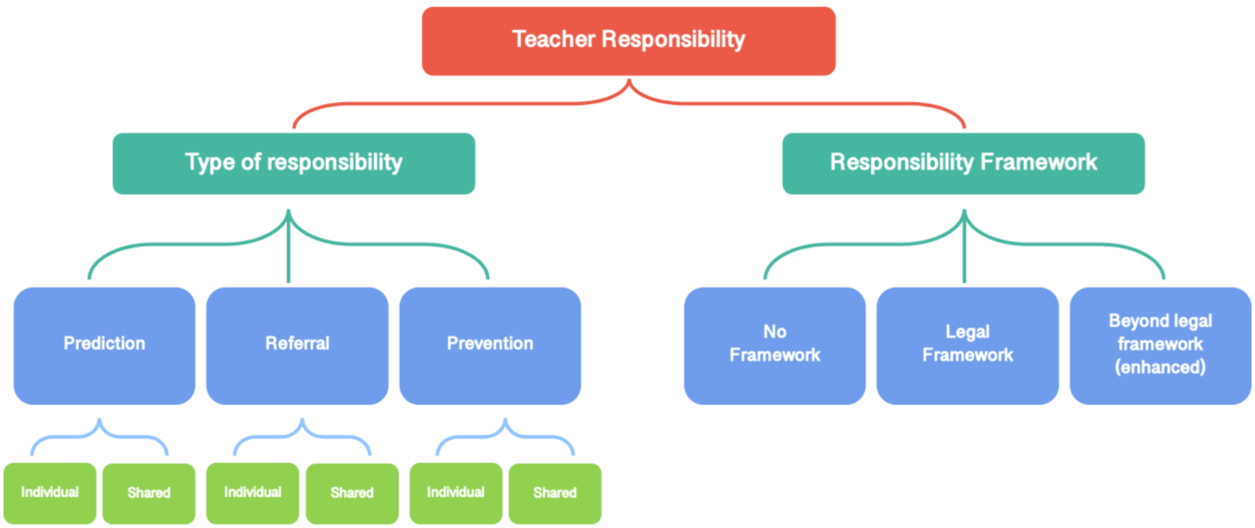
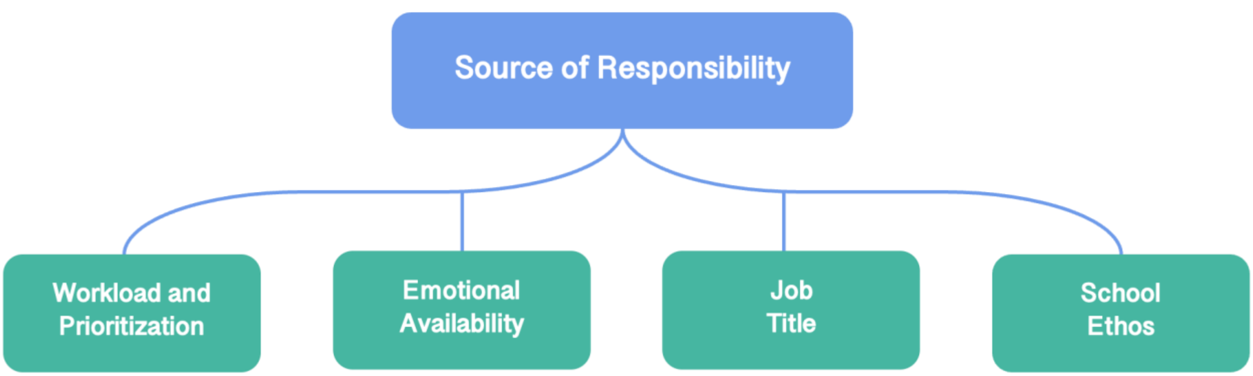


Figure 12

Coding Scheme, What are Teachers' Reasons for Responsibility?



6.2. Results

As outlined in Figures 11, to answer the question, "what are teachers' roles and responsibilities," the results are organized to show how two analytic dimensions relate to one another. The first dimension concerns the types of suicide-related responsibilities teachers take on in schools: prediction, referral, and prevention. The second-dimension concerns whether (and how) teachers draw on any clear framework of responsibility in carrying out these responsibilities. These frameworks fall into three categories: (1) the absence of a clear framework, (2) alignment with existing legal or ethical frameworks discussed earlier in the thesis, or (3) the adoption of an 'enhanced' framework that goes beyond formal requirements.

While summarizing the first section of my results (6.2.1), I map these two dimensions onto each other. In other words, for each of the three responsibility types (prediction, identification, and referral), I examine whether teachers describe acting simply 'because it's what they do,' or whether they understand their practice as tied (explicitly or implicitly) to one of the responsibility frameworks ('legal' or 'enhanced').

Third, the results section explores whether teachers identify any casual, or mediating factors in their definitions of responsibility. This includes contributing and mediating factors that influence the extent to which teachers feel responsible, such as workload, prioritisation (which links to the previous chapter's discussion of political and economic constraints), emotional availability, job title, and school ethos. Figure 13 and Table 11 below provide an overview of these themes, and the subsequent sections analyse each dimension in greater depth.

Figure 13

Flowchart Illustrating What Contextual Factors Shape Teachers' Responsibility Practices

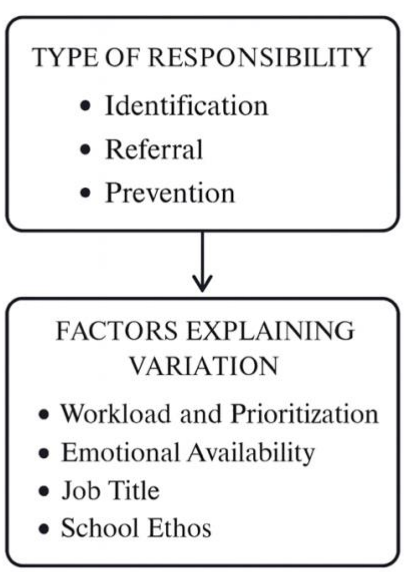


Table 11

Types of Responsibility and Factors Shaping Teachers' Role Perceptions, with Illustrative Quotes

Category	Theme	Description	Illustrative Quote
Type of Responsibility	**Identification**	Teachers universally see themselves as responsible for noticing early signs of poor mental health or emerging risk, consistent with DfE guidance.	"We're the first point where we might flag or notice signs." (P10)
Type of Responsibility	**Referral / Working Together**	Teachers emphasise that identification is only the first step: responsibility is shared through referral to other professionals or teams.	"We can't contain these issues on our own… making sure it's passed on to somebody else very quickly." (P14)
Type of Responsibility	**Prevention**	A smaller subset extend their role into active prevention, either through curriculum work or emotional support in high-risk moments.	"I step in at moments where things are too much… You are another trusted adult who can support them." (P17)
Factor Explaining Variation	**Workload and Prioritisation**	Teachers' willingness to take on enhanced responsibility is shaped by curriculum demands and time constraints, which limit their capacity to engage deeply with student wellbeing.	"We do have so much that we need to deal with… there does need to be a really clear line drawn between I am here to teach English and they are there to look after wellbeing." (P11)
Factor Explaining Variation	**Emotional Availability**	Teachers vary in their emotional capacity to manage pastoral issues; some find such work fulfilling, while others avoid it due to emotional load and risk of burnout.	"I steer clear of the pastoral stuff because I find it extremely emotionally draining and loaded." (P11)
Factor Explaining Variation	**Job Title**	Role and designation shape responsibility: pastoral or leadership staff see themselves as holding enhanced safeguarding roles, while classroom teachers tend to adopt a narrower remit.	"Any teacher has a responsibility for safeguarding, then as part of the leadership team… that responsibility is enhanced a little bit." (P10)
Factor Explaining Variation	**School Ethos**	School culture influences teachers' sense of responsibility: wellbeing-oriented schools encourage broader engagement with mental health, while academically focused schools reinforce narrower boundaries.	"It probably comes into the ethos of the school… you tend to soak up what the school is expecting of you." (P15)

6.2.1. Responsible for Identification, Referral, and/or Prevention

The first set of interview questions explored teachers' perceptions of their roles in monitoring student mental health. The findings are summarised in the table above, which outlines three primary types of responsibility: prediction, referral, and prevention of suicide. In this section, I further examine which aspects of responsibility teachers perceive as solely theirs, and which they view as shared.

In addition, as this is an empirical bioethics project, the qualitative data are also used to reflect back on earlier normative and policy findings. For example, in the document 'Mental health and behaviour in schools: Departmental advice for school staff', the Department for Education (2018) argues that teachers are key adults in adolescents' lives that may see the first signs of emergent mental health difficulties, including suicide and self-harm.

6.2.1.1. Identification

All teachers within my sample agreed with the DfE, in that they all argued that teachers play an integral role in *identifying* poor mental health, or risk factors for poor mental health, and then *referring* that student to another organisation/team (usually the school or council safeguarding team or an external agency). For instance:

> *We're the first point where we might flag or notice signs that would then be passed on to the safeguarding team.* (P10)

> *I think responsibility just means an awareness and being able to identify what is, if you like, typical neurotypical from the neurodiverse from the not quite adding up from the something is happening that's… Or spotting, even just maybe kind of an acute incident that's happened or… I think everybody does have an awareness, it's part of teaching, you wouldn't go into teaching unless you were prepared to take on that. And it's just a vigilance, I think, of just noticing when something is different. Not necessarily different bad, could be different good or just different. And just not necessarily raising the alert to everybody, but just being mindful of the fact of what the next step should be.* (P14)

Teachers' responses regarding their roles in identifying 'at-risk' pupils were closely aligned with existing legislation and policy, with participants frequently referencing official documents to support their views (including documents which came up in Chapter Five's literature review, e.g. KCSIE; Department for Education, 2020). For instance,

> *We're legally being held responsible as well. If something perhaps was revealed to me or if I saw something and then didn't go the proper route about reporting it, I would then be held legally responsible as well. (P15)*

6.2.1.2. Working Together (Through Referrals)

Teachers, by and large, do not see their ‘identification’ role as siloed. Instead, interviewed teachers regularly bring up the need to work with other groups on high-risk cases. This method of sharing responsibility is consistent with the philosophical and legal literature (e.g. statutory guidance, discussed in Chapter 5).

> *And also, not keeping it to yourself. I'm very big on making sure it's passed on to somebody else to share it very quickly. I don’t think we can keep… We can't contain these issues on our own, we're not… You know there are so many people looking after our students that one person alone really is unlikely to be able to solve their problems.* (P14)

Most teachers consistently maintained that there is a clear division of responsibility: teachers are expected to identify and refer students at risk, while other professionals are responsible for intervention. This perspective aligns with education policy documents, as discussed in Chapter 5. For instance, according to the 2018 DfE report titled ‘Mental Health and Behaviour in Schools,’ while teachers are required to make referrals if they suspect a student is at risk of poor mental health, they are not legally obligated to assess, diagnose, or provide ongoing social or emotional support.

> *So, my solution would be employing more people to do like school counsellors and people trained in this software whose role it is to deal with that and leave teachers to teach, whilst also having good relationships with the students. It's a delicate balance.* (P11)

However, despite this clear guidance from the DfE, some literature (e.g. O’Reilly et al., 2018), shows that, in practice, these seemingly separate responsibilities are often blurred, and teachers often go beyond making referrals, into prevention. This trend also emerged in the interviews, with some interviewed teachers appearing to blur the boundaries of responsibility, viewing their role as extending beyond only making referrals.

6.2.1.3. Working Together (Beyond Referrals, Into Prevention)

More specifically, while there was a consensus that, in practice and by law, teachers are responsible for the identification and referral of risk, some teachers who were interviewed believed that teachers are also responsible for additional tasks, including the *prevention* of suicide in cases of high risk students. This often took the form of actively supporting student mental health, such as developing PSHE curricula or leading school-wide awareness initiatives (as in the first quote), or working directly with student groups to manage emerging concerns (as in the second quote):

> *Being responsible would mean for me, before it ever becomes an issue, education for the students around their mental health and trigger possible issues with mental health. So, that might be increasing links with mobile phone use, positive friendships and relationships in schools*. (P10)
>
> *If some of these issues are arising out of friendship disputes and things like that, offering a different perspective which is almost therapeutic because it takes them out of their own experiences and you then give them a different inflexion.* (P17)

This also includes teachers who see their role as extending into prevention *only at certain moments*. As the participant below illustrates, some teachers straddle the line (e.g. primarily referring students, but stepping into prevention during what they perceive as 'big' or critical moments).

> *And then intervening at moments where I feel like maybe things are too much or if there are difficult things going on... So recently there was an incident at school which meant that a number of my tutees were very unhappy and upset about it. You are then also another trusted adult who can kind of look after them and look out for them and support them too. So, I do a bit of both.* (P17)

6.2.2. Why are there differences?

Overall, the previous section (6.2.1.) identified three main perspectives among teachers regarding their role for suicide prediction: identification (endorsed by all teachers), referral (endorsed by most), and prevention (endorsed by very few). However, not all teachers interviewed hold an "enhanced" view of responsibility that goes beyond the legal requirements set out by KCSIE. Therefore, this explores the reasons teachers gave for adopting (or rejecting) these different conceptions of responsibility.

First, to fully outline the differences in role perceptions emerging from these interviews, a specific tension arose between two categories of teachers, as illustrated on the next page:

A lot of people tend to think that teaching is the most important thing that you do, but actually it's the child, it's wellbeing and the safeguarding of the child (P15)	vs	*Obviously we all care very deeply about the children that we deal with, but I think there does need to be a really clear line drawn between I am here to teach English and they [safeguarding team] are there to look after your wellbeing* (P11)

More broadly, the first quote advocates for an integrated approach to education and care, where safeguarding may take precedence over a teachers' academic role. Contrarily, the second stresses the importance of delineating responsibilities between teaching and wellbeing, arguing that the teacher's responsibility is teaching, while safeguarding and wellbeing are designated to specialised staff[23].

Why are teachers divided in the ways illustrated above, and have a fundamental difference in what they think the role of a teacher should be? In addition, what problems may arise from these different views about enhanced monitoring roles, e.g. what may happen to students and teachers if a teacher takes on a minimalist vs enhanced view on responsibility?

In a previous study on stakeholder responsibility for adolescent mental health, O'Reilly et al. (2018) identified the following key themes to explain why many teachers rejected this 'enhanced' responsibility that went beyond identification: lack of capacity (including funding and time), teacher skills (such as limited training and knowledge), and secondary stress or burnout. In addition, the previous chapters of this DPhil, particularly the Systems Analysis in Chapter 5,

---

[23] It is important to note that teachers so far are only divided on whether they should be responsible for any 'enhanced' mental health monitoring roles beyond their statutory requirements (e.g. prevention or active monitoring). All teachers are clear they have the basic responsibility of reporting.

suggested that factors such as the UK's current economic and political landscape may play a significant role.

Similar themes emerge from my data. More specifically, four themes emerged from the data, and these themes provided the 'reasoning' behind how and why there were two sets of teachers: those who held this 'enhanced' idea of responsibility, and those who did not. These themes are workload and prioritisation, emotional availability, job title, and school ethos.

6.2.2.1. Workload and Prioritization

A teacher's responsibility for monitoring students' suicide risk is significantly impacted by their workload and prioritization challenges, as many educators struggle to balance various demands, including curriculum demands, external reporting, and standardized testing. With these competing priorities, teachers may find it difficult to allocate sufficient time and attention to monitoring students' mental health effectively. Consequently, the pressure to manage a heavy workload can limit their capacity to engage deeply with students' emotional needs. This is illustrated by the quote below:

> *without sounding dispassionate or unsympathetic, we do have just so much that we need to deal on a day-by-day basis and to have that taken out of our hands, just allows us to get on with our jobs. Obviously we all care very deeply about the children that we deal with, but I think there does need to be a really clear line drawn between I am here to teach English and they are there to look after your wellbeing.* (P11)

The quote reflects a teacher's perspective on the balance between their workload and the responsibilities of caring for students' well-being. The teacher acknowledges that, while they deeply care about the students, the demands of their daily tasks are already significant. By suggesting that certain responsibilities, such as student well-being, be handled by others (for instance a DSL), it highlights the need for boundaries to manage workload effectively while ensuring that students' well-being is appropriately addressed by the right professionals.

The current study is by no means the first to highlight the immense workload that teachers face. Across various studies, teachers consistently report that their ability to engage with adolescent mental health issues is largely determined by their workload, and that many are over-capacity (Dabrowski et al., 2025). Teachers argue that they lack resources including time or funding and are often constrained by broader school-conditions and country-wide economic conditions. For instance, schools have been significantly affected by the UK government's policies of austerity (Sims-Schouten, 2017) and struggle to provide adequate resources (e.g. enough teachers or specialised support staff). As a result, teachers find themselves ill-equipped to manage complex mental health issues, such as adolescent suicide, which require specialised attention and care.

6.2.2.2. Emotional Availability

Next, the interview data shows that a teacher's responsibility for monitoring students' suicide risk is closely tied to their capacity for an emotional load, as those who are more emotionally available are often better positioned to provide the support needed. However, engaging deeply with students' mental health can lead to emotional burnout, especially if the teacher is already managing a high workload or feels overwhelmed by the emotional demands. The interviews (as illustrated by the quote below), make it clear that this responsibility requires careful consideration to prevent burnout while ensuring that students receive the necessary care and support.

> *I tend to steer clear of the pastoral stuff because I find it extremely emotionally draining and loaded. So that's a personal choice that I've made. So, steer more towards the curriculum, academia side of it, but massive respect for everyone who does deal with the pastoral side of things. And thank god they're there.* (P11)

According to the academic literature, engaging with others' traumatic experiences can heighten stress and anxiety (Devilly et al., 2009). Unlike mental health practitioners, who often receive supervision or therapy to manage the risk of burnout, the teachers interviewed typically lack access to these forms of support. As a result, when students disclose serious mental health issues, such as self-harm or suicidal thoughts, teachers may experience a significant emotional burden, which many feel ill-equipped to manage effectively.

Ultimately, the data (particularly the quote above), underscores the emotional toll that pastoral care can have on teachers and reflects an understanding that different teachers may have varying levels of comfort (or capacity) when it comes to managing students' emotional well-being.

6.2.2.3. Job Title

A teacher's responsibility for monitoring students' suicide risk can vary based on their job title and designation, with those in leadership or pastoral roles typically holding greater accountability for safeguarding and mental health support. While all teachers share a general duty to ensure student well-being, according to the interviews, it is clear that those in senior positions or with specific pastoral duties are often more involved in the direct oversight and intervention regarding students' mental health risks. For instance:

> *any teacher in the school obviously has a responsibility for safeguarding, then it's part of the leadership team obviously. But that responsibility is enhanced a little bit. As part of my role, I line-manage the person who looks after one of the year groups in the school.* (P10)

The above quote shows that, for those in leadership roles, there may be a stronger inclination to engage with EdTech for suicide prediction because their roles already involve more responsibility for student welfare (e.g. being directly responsible for a specific year group's wellbeing). This quote also illustrates the speaker's acknowledgement of a shared responsibility among all teachers for safeguarding students, with an added layer of responsibility for those in leadership roles.

However, it is important to note that even these teachers might experience concern about the emotional and professional burden of constantly monitoring students' mental health through technology, as well as the risk of overstepping their professional boundaries, into the role of a parent.

> *Parents play the main role, because they stay with their children 24 hours, you can say. And a teacher, we may just stay with them maybe one hour or maximum two hour every week. Because we don't have lesson with them every day, maybe just twice a week. So we're with them maybe twice a week for one hour every time.. I feel the parents need to play the main role. Parents need to have some training about it, I think, definitely.* (P13).

Therefore, some may resist involving themselves too deeply in the use of such technology to preserve their focus on educational responsibilities and have a clearer role between teacher and parent.

6.2.2.4. School Ethos

The way a school uses EdTech to predict suicide risk is closely tied to its ethos, or the values and priorities that define its approach to education and student care. The interviews show that schools with a strong focus on mental health, pastoral care, or those striving to become 'wellbeing champions' (where student well-being and emotional health are central to the school's ethos) may be more likely to embrace technologies aimed at predicting and managing suicide risk. These schools see mental health as a core responsibility and would naturally integrate tools that support that mission.

> *it probably comes into the ethos of the school as well, because you do tend to soak up that ethos and soak up what the school are expecting of you as well.* (P15)

The quote above highlights how a school's ethos influences teachers' attitudes and practices. Teachers absorb and adapt to the expectations set by the school's culture, meaning that in a school where student mental health is a priority, staff may feel more inclined (or even obligated to) engage with EdTech programmes which are used for suicide prediction.

Conversely, in schools where the focus is more academic or where mental health is not as emphasised, there may be less support for using EdTech in this way. Teachers in these schools may feel that their primary role is educational, and the introduction of suicide prediction technology could feel like an additional burden that stretches their responsibilities beyond what they believe is appropriate.

The importance of a school's ethos in promoting positive behaviour and mental health has been highlighted by authors such as Spratt et al. (2006a; 2006b). For example, Spratt et al. (2006b) have seen clear differences between schools in which pastoral and mental health care were an integral feature of discipline, versus schools who see these as two separate, unrelated systems, within a school (e.g. within schools that see mental healthcare as a feature of discipline, interventions often "inflamed the situation and led to escalation" (Spratt et al., 2006b, p.17).

Yet, it is clear that a school's ethos does not play the only role in a teacher's actions, and decision on which intervention to engage. Spratt et al. (2006b) and Spratt et al. (2006a) found that, regardless of whole-school ethos, the decision to respond to pupils' behaviour, and decide which intervention to enact, is held by the individual classroom teachers in which the 'risky' behaviour is found.

6.3. Conclusion

In this chapter I investigated teachers' role within the school, exploring teachers' perspectives on their role in monitoring student mental health, specifically suicide. The first section examined how this responsibility is currently implemented in practice. This included teachers' roles in identifying signs of poor mental health and referring students to appropriate support teams. Within the interviews, it was clear that teachers consistently recognized their role in spotting potential issues and reporting them, and followed government guidelines like KCSIE. However, the data found a distinction between teachers' basic responsibilities of identifying and referring issues and a more extended view of a teacher's role, which included suicide prevention and broader mental health education. While there was general agreement on teachers' statutory duty to report, opinions diverged on whether teachers should undertake additional tasks related to mental health beyond these requirements.

The second section of my results explored why teachers' opinions diverged. The reasons behind varying perspectives on teachers' responsibilities for mental health monitoring can be categorised into four main themes:

1. *Workload and Prioritisation:* Teachers faced challenges balancing their heavy workloads with mental health responsibilities. Competing demands like curriculum, reporting, and standardised tests can limit their capacity to focus on students' emotional needs.
2. *Emotional Availability:* Teachers' ability to manage students' mental health risks were influenced by their emotional availability. Those who find pastoral care emotionally draining, in general, had a preference to focus on academic duties.
3. *Job Title:* Responsibilities for mental health monitoring often aligned with the teacher's title. For instance, teachers in leadership or pastoral roles typically had a greater focus on

safeguarding and mental health, whereas teachers in other roles, on average, felt less inclined to engage deeply with these issues.

4. *School Ethos:* A school's ethos significantly impacts how EdTech programmes were (or would be) adopted. Based on the interview results, schools that prioritise mental health and well-being would be more likely to use such technology, whereas schools focused more on academics may resist additional responsibilities that come with new monitoring technologies.

6.3.1. Limitations

A limitation of this project is that I was not able to complete full inter-rater reliability, primarily due to time constraints associated with the DPhil. Ideally, I would have conducted independent double-coding of a subset of transcripts followed by a more systematic comparison and reconciliation process. Nevertheless, I took several steps to mitigate this limitation. In the early stages of analysis, I arranged a meeting with three other bioethics DPhil colleagues, during which we all read the same transcript and then engaged in a facilitated group discussion. This session involved an in-depth review of the themes each person identified, including both the consistent patterns and the deviant or outlier cases. We specifically focused on developing hypotheses about why the identified participant embraced a large degree of individual responsibility for their student's mental health. Through discussion, we ultimately realised that their unusually positive stance was linked to the fact that they did not work full-time, which reduced factors such as emotional burnout. This reflective process aligns with the approach taken in other DPhil studies within the research group (e.g., Manzini, 2020).

In addition, samples of the interview data were presented and discussed at three academic meetings, including the Australasian Association of Bioethics and Health Law (2023) and the Neuroscience, Ethics, and Society meeting at the Department of Psychiatry (2024). These opportunities provided further external input on my developing interpretations, reduced the risk of individual bias, and strengthened confidence in several emerging hypotheses (e.g., reasons underlying discrepancies in teachers' sense of "enhanced" responsibility).

Overall, this chapter presented the results of my empirical research. While it has some limitations (e.g. sampling constraints and the absence of interrater reliability, listed above), it nevertheless offers valuable insights into teachers' views and values towards being responsible for suicide prediction with new technology.

In the next chapter (Chapter 7), I will analyse the interview data in greater depth through the lens of shared responsibility. This includes comparing and contrasting the qualitative interview findings with earlier chapters, such as Chapter 5 (my analysis of education systems in which these software programmes are deployed). By bringing these data sources (the mapping review, systems analysis, and interviewing work) into dialogue and applying philosophical scrutiny, I aim to strengthen the normative argument at the heart of this thesis (Dunn et al., 2012). Specifically, this asks, through a shared responsibility lens, *what responsibilities do teachers and schools have when using technology for suicide prediction, and how do these compare to what their responsibilities should be?*

Chapter 7. Shared Responsibility

The first six chapters of this thesis outlined the current ethical landscape around the use of EdTech for suicide prediction (Chapter Two), the technological capacities of EdTech programmes used within the UK (Chapter Four), the embedded use of these tools within the school and larger educational structures (Chapter Five), and insights from teacher interviews, specifically focused on what interviewees believe a teacher's responsibility is when it comes to student mental health and suicide prediction (Chapter Six).

Chapter Seven works to summarise and integrate the key ideas from the first six chapters, and uses the framework of shared responsibility to answer the core (primary and secondary) research question(s), specifically:

- *What responsibilities do teachers and schools have when using technology for suicide prediction, and how do these compare to what their responsibilities should be?*
- *Given that teachers do not work in isolation (and that children are embedded within multiple overlapping support systems), how might a model of shared responsibility (involving teachers, clinicians, parents, technology developers, and/or the students themselves) function? How should responsibility be shared (as indicated by legal, policy, and ethical frameworks), and is it being done so in practice?*

Using a framework of shared responsibility, Chapter Seven argues that while responsibility *should* be shared (as indicated by statutory, policy, and ethical frameworks) it is not being effectively implemented in practice.

This chapter draws on both theoretical and empirical evidence to explore the differences between the philosophical concept of shared responsibility among parents, social services, and students and its practical enactment within secondary schools. The chapter includes data from the teacher interviews, systems analysis, and mapping review. Core themes that emerge from each of these perspectives are presented across two subheadings. These two subheadings (and their associated sub-headings) are:

- 7.1 An Exploration of Shared Responsibility in Schools and the Use of EdTech
    - 7.1.1 Discussing the Multi-Agency Approach
    - 7.1.2 Examining Differences and Alignment Between Theoretical Shared Responsibility and Its Enactment in Practice
- 7.2 Ethical Challenges Stemming from These Differences

By using EdTech as a case study (both in its current capacity and by exploring potential future applications in school mental health), this chapter examines not only the limitations of shared responsibility, as enacted in its current form, but also how these limitations may affect students and society (Chapter 7.2). This includes effects related to broader issues introduced in earlier chapters, surrounding student privacy, beneficence, autonomy, the clinical utility of such tools, and their potential to discriminate against and/or to disproportionately harm certain groups of students (originally outlined in Chapter Two).

For example: if teachers are responsible for using monitoring tools while collaborating with parents, then the distribution of responsibility might look one way, and ethical concerns (e.g. consent or privacy) might take on certain characteristics. Conversely, if teachers are the sole users of the system, and interventions are confined within the school (e.g., excluding input from parents or other stakeholders), then the distribution of responsibility and connected ethical issues could look entirely different.

Throughout the chapter, I strongly advocate for a shared responsibility approach to suicide prediction in schools, though not in its current, enacted form. I conclude this chapter by exploring future scenarios of the use of EdTech for suicide prediction, specifically examining how the introduction of virtual therapists may impact shared responsibility. By integrating different scenarios of shared responsibility into the existing ethical literature on the impact of EdTech programs for mental health more broadly, I demonstrate the scholarly significance and broader implications of my research. Chapter 8 will build on this work by exploring the implications of my research for schools, policymakers, and new legislation, while also reflecting on research gaps, the latter of which informs future research directions.

7.1. Whole School and Multi-Agency Approaches

The definition and distribution of shared responsibility in safeguarding student mental health is growing in complexity as students and schools navigate increasingly intricate ecological environments and education policies (Schuelka & Engsig, 2022). To account for these increasingly complex ecological systems, schools are moving away from a model of individual responsibility that is typically placed on an individual teacher or counsellor and embracing a model of shared responsibility. Within school policy, this is usually classified in one of two different ways: a whole-school approach or a multi-agency approach.

The whole-school approach to a school's responsibility for mental health and wellbeing is emphasised in current government guidance (Chapter 5) and includes mental health strategies where the entire school community (teachers, staff, students, and families) collaborate to address the needs of all students (Cefai et al., 2021). Although there is currently no legal requirement for mandated reporting within schools, there is an expectation that school staff adhere to safeguarding and reporting guidance (and growing political momentum toward making this a legal obligation in the near future (Foster, 2025)[24])).

Meanwhile, a multi-agency approach, also outlined in 'Working Together to Safeguard Children' statutory guidance (HM Government, 2023), emphasizes that safeguarding is not only the responsibility of a single teacher or of all school staff, but also that of external agencies, and community organizations such as health services and police (Public Health England, 2021). This move, to include a broad range of external agencies, also becomes especially critical for students with severe or complex mental health needs, such as seen in cases of suicide, as they often require a multidisciplinary support system that reaches beyond the school walls.

[24] This movement towards mandated reporting being codified into UK law, and the impact of this research on the (potential) change in legislation, is discussed in more detail in Chapter 8.

Table 12

Whole School v Multi-Agency Approach to Responsibility

	Definition	Example
Whole School Approach to Responsibility	A form of responsibility where the entire school community (teachers, staff, students, and families) share responsibility for their students' mental health.	Everyone in a school has clear safeguarding responsibility and there are clear guidelines in place so each party knows their responsibility in: monitoring, intervention, data-sharing (and so forth).
Multi-Agency Approach to Responsibility	A collaborative approach where different professional organizations (e.g., education, health, social services) work together to support a child or family, and have clear guidelines on responsibility.	A child who is at risk of suicide is supported by a team consisting of teachers, a school counsellor, a social worker, and a healthcare provider, each contributing their expertise and working together to develop an individualized support plan. Sometimes called 'Team Around the Child' meeting (Liverpool Safeguarding Children Partnership, 2024). Within the multi-agency approach there are clear guidelines in place so each party knows their responsibility in: monitoring, intervention, data-sharing (and so forth).

As seen in the definitions above, a multi-agency approach not only considers the roles of individual agencies but also emphasises how these agencies interact with children and families to provide coordinated support. An example of the multi-agency approach 'in-action' are multi-agency planning meetings, for example a Team Around the Child (TAC) meeting, where professionals from different sectors (including education and social care) are brought together with the student and their family to create a coordinated care plan (outlined on council websites, including Liverpool Safeguarding Children Partnership, 2024; Sutton Council, 2025). Other approaches to multi-agency care could include integrated, online care platforms, allowing different stakeholders to contribute data or input within the same software for more seamless collaboration, or data-sharing agreements between organisations, such as the NHS, schools, and local authorities.

It is important to note, that typically multi-agency approaches focus on well-established community organizations, e.g. the NHS and social care teams, which may fail to account for broader, less-established communities and new private partners, such as EdTech companies[25], which have been seen as critical actors within recent cases of EdTech for suicide prediction. Therefore, the data collected in this thesis examined real-life examples of multi-agency approaches to EdTech for suicide prediction, e.g. through data sharing across and within specific technological platforms, highlighting these 'less traditional' actors, such as private technology companies.

While the first six chapters of this DPhil examine how a whole school/multi-agency approach functions in practice (particularly in the context of mental health monitoring apps), this chapter aims to collate these data points to answer two critical questions:

1. Do these real-life examples of multi-agency approaches to EdTech for suicide prediction reflect the normative ideal of shared responsibility, particularly from a philosophical perspective? (7.1.1.)
2. What are the limitations of the current multi-agency approach? (7.1.2.)

7.1.1. Differences / Alignment

Responsibility in schools is clearly being shared, to a certain degree, among multiple stakeholders. For example, one interview participant noted: *You know there are so many people looking after our students that one person alone really is unlikely to be able to solve their problems.* (Participant 14) But how is this responsibility being shared? And do 'real-life' examples align with the philosophical ideal of shared responsibility?

This section examines whether real-life examples of multi-agency approaches to EdTech for suicide prediction embody the normative ideal of shared responsibility. To address this, I first describe the concept of shared responsibility, both as it comes about in the data and based on bioethical theory, before analysing the extent to which these align.

---

[25] Chapter Five discussed this changing policy landscape, and further argued that the development of a commercial 'ecosystem' of EdTech for suicide prediction has made this model increasingly complex. For instance, while collecting more data may be helpful to predict risk more accurately, and sharing data between multiple parties (e.g. schools, parents, law enforcement, etc) can lead to more integrated interventions, it also results in more stringent (and complicated) scenarios.

As outlined in Chapter Three (3.2), shared responsibility refers to the collective and distributed roles of diverse stakeholders in achieving a common goal (as framed in the context of role responsibility; Hart, 1968). It implies that responsibility is equitably shared across all participants, each of whom has clearly defined roles and an active, collaborative role in decision-making processes (Kon et al., 2016)[26]. Crucially, this definition of shared responsibility requires active engagement by all stakeholders to prevent the dilution or abandonment of accountability.

By contrast, the multi-agency approaches examined in this thesis focus on coordination between multiple organizations or entities. While using a multi-agency framework can facilitate information sharing and resource allocation, it does not necessarily ensure that responsibilities are equitably distributed or that all stakeholders are meaningfully engaged (similar issues arise in other forms of collaborative governance structures, explored in Ansell & Gash (2008)). Thus, while these two concepts are often conflated (shared responsibility and multi-agency approaches to safeguarding), they are distinct both in theory and practice, which can be demonstrated by the data collected in this thesis. Specifically, the findings from the first six chapters of this thesis highlight several critical discrepancies between the normative ideal of shared responsibility and the implementation of multi-agency approaches in EdTech for suicide prediction, including: the exclusion of key stakeholders (7.1.1.1.); lack of shared commitment (7.1.1.2.), the lack of clear roles (7.1.1.3.), and the blurring of geographic boundaries (7.1.1.4.).

7.1.1.1. Exclusion of Key Stakeholders

According to the definition of shared responsibility laid out in Chapter Two, a fundamental component of shared responsibility is the active inclusion, or engagement, of all affected parties (Kon et al., 2016). However, the data from Chapter Four reveal that key stakeholders; most notably students and their families, are often excluded from these EdTech programmes. For instance, while some EdTech programs (e.g., Lightspeed, Smoothwall, Securly, and Impero) reference parental involvement, this feature is inconsistently implemented and often limited in scope (described

[26] A study by Kon and colleagues (2016), which discusses shared decision making in Intensive Care Unit, defines shared responsibility as: “a collaborative process that allows patients, or their surrogates, and clinicians to make health care decisions together, taking into account the best scientific evidence available, as well as the patient’s values, goals, and preferences” (p.2).

earlier in Chapter Four). Furthermore, none of the programs analysed allowed students to view or contribute to their own data, despite the growing emphasis in the literature on incorporating personal care networks (e.g., family, peers) into mental health interventions (Panaite et al., 2024). Therefore, while EdTech systems are advertised to support collaborative mental healthcare, Chapter Four findings indicate that these programmes often centralize control within specific actors, such as school administrators or even the EdTech companies themselves. The exclusion of students, in particular, marks a departure from the normative ideal, which prioritizes equitable and inclusive participation by all stakeholders.

7.1.1.2. Lack of Commitment towards Shared Responsibility

In Chapter Two, when outlining an ‘ideal’ framework of shared responsibility, it became clear that one of its core requirements was a shared commitment from each group to work collaboratively. However, the data presented in Chapters Four and Six reveal significant ambiguities in how different parties understand and value shared responsibility, and indicate that some stakeholders lack a consistent commitment to working with others. The data does not suggest an absolute absence of commitment from any particular group, but rather highlights specific practices that can undermine the possibility of *truly* shared responsibility and risk the dilution or abandonment of accountability, if they were to continue. Specific examples of this from the thesis include EdTech companies whose policies limit data sharing (Chapter 4), and some parents who, according to teachers, resist engaging as collaborative partners (empirical data, Chapter 6).

With regards to the first example, Chapter Four shows that although the platforms analysed (such as Impero and Lightspeed), collect substantial data on student behaviour, they rarely integrate this information with broader health or education datasets. This fragmentation hinders collaboration between stakeholders, making it difficult to establish who should act, when intervention is necessary, and how responsibilities ought to be coordinated. According to the Digital Futures Commission (2023), this limited data sharing is often driven by efforts to protect intellectual property or comply with data protection regulations.

In terms of my second example, as discussed in Chapter Five, teachers identify other parties (such as parents) as crucial actors, yet often encounter resistance from them. For example, one teacher described experiencing parental pushback in the following way:

> *I think part of the broad pushback we've had from the parents is encroachment of the school into the private family life. Even if it's well intentioned, even if our motivations for the policy and things like that is around supporting students with their social interactions, with their mental health, and things like that, some parents' pushback is around, you shouldn't be parenting my child. It's my child and I'll do that as I see fit.* (P10)

Both of these examples suggest that the issue may not be any specific group's opposition to the principle of shared responsibility, but rather to the practicalities of how shared responsibility is currently enacted, and barriers that may limit each party's commitment. These include the demands of corporate profit, and debates over data sharing, which become particularly relevant in the context of using EdTech software for suicide prediction.

7.1.1.3. Lack of Clear Roles

According to the Commission for Social Care Inspection (CSCI, 2005):

> All the evidence indicates that children are safeguarded best where there is clarity and understanding between different agencies about roles and responsibilities, underpinned by good working relationships at all levels. (p.33)

As seen in the description above, a key component of effective shared responsibility is that all parties clearly understand their own roles and responsibilities, as well as those of any other agencies, in safeguarding. This includes not only each group's individual duties and how those intersect, but also the shared responsibility of working collaboratively. However, the data collected in this thesis shows a lack of transparency and unclear role definitions regarding how teachers collaborate with other groups. Unclear role definitions are seen throughout the data collected in this thesis - in law, teacher interviews, and within the technology itself. For instance, in terms of individual responsibility, to some teachers, it is unclear that teachers are expected to identify students at risk for mental health concerns but are not legally obligated to conduct mental health assessments or provide social and emotional care (Foster, 2025). These uncertainties extend beyond teachers, as parents also play a crucial role in safeguarding but are not always clear about the responsibilities expected of them or how they should coordinate with schools. In addition to these ambiguities, there is also limited clarity about how data generated through digital monitoring should be interpreted and acted upon, including who is responsible for responding to flagged concerns and what processes should follow. This further complicates attempts to establish coherent and coordinated shared responsibility.

The interview section of my thesis showed that teachers think that the lack of clear roles may be exacerbated by the use of technology, specifically. For example, Participant 10, in the quote below, highlights the issue of unclear responsibility and the distribution of responsibility in safeguarding within schools, particularly in the context of digital technology for suicide prediction.

> *If you're an everyday classroom teacher, you notice a concern, you pass it on to someone else and then….you then take a step back away from it and you don't necessarily see the process that then follows. You've done your responsibility as a classroom teacher to pass it on….Would having this digital technology, if it's flagging up things more regularly, does that have an impact on a teacher's day-to-day responsibilities? Are they seeing more of these things? And if so, what is the impact on that individual teacher?* (P10)

According to this teacher, traditional safeguarding is a well-defined process: a teacher may identify a concern, report it to the appropriate safeguarding lead, and then step back. This would indicate a clear division of responsibility between teachers (who teach) and safeguarding leads (who are responsible for interventions). However, according to the teacher interviewed, the introduction of digital technology may disrupt this structure. For example, if teachers receive more alerts or are expected to engage more in the follow-up process, their role may become less clearly defined.

The potential for EdTech to further complicate how responsibility should be shared is evident not only in teacher interviews but also within the technology itself. Chapter Four, which analysed the technology, revealed that while companies claimed their systems enabled multi-party collaboration (e.g., between schools and external services), the specifics were often vague and varied between providers. For instance, only one company, Smoothwall, explicitly outlined its protocol for engaging emergency services: stating that emergency services would only be contacted if a teacher could not be reached. In contrast, other programs were less transparent about when and how they would involve external services such as the NHS or police (see: Appendix One).

Smoothwall's protocol is also a clear example of a new party emerging within the safeguarding process – the EdTech company itself. Because Smoothwall's system actively scans, interprets, and elevates risk (rather than passively storing data), and works with both teachers and emergency services, the company itself (and those within it, including human moderators (if available), engineers, and leadership personnel), inevitably becomes an actor within that process. Having an EdTech company positioned as an actor in safeguarding raises significant ethical and legal questions about the responsibilities Smoothwall (and other providers) hold, both as designers and operators of safeguarding technologies, within the context of the larger safeguarding network.

For example, ethically, Smoothwall must ensure that its risk-detection algorithms are accurate, that alerts are escalated appropriately, and that teachers are not placed in situations where system errors could cause harm. Legally, Smoothwall and other companies operate within a complex regulatory landscape shaped by data-protection law (e.g., GDPR), safeguarding statutory guidance (such as KCSIE), and product-safety expectations, and will increasingly need to learn how to work effectively with these regulatory bodies, and alongside other groups (such as teachers).

These examples show that, while theory shows there should be a clear distinction between teachers, EdTech companies, and those who are mental health trained or designated safeguarding leads, in practice, these boundaries often become blurred. This lack of clarity introduces further ambiguity regarding responsibility, making it difficult for schools to understand where accountability lies when an automated system becomes involved in a safeguarding process.

It is important to note, however, that blurred boundaries are not inherently unique to the context of EdTech for suicide prediction, nor are they necessarily always a negative. Instead, the academic literature acknowledges that role fluidity is common, particularly within multi-agency teams. According to Frost & Robinson (2007),

> as professionals move between communities in the workplace, professional identity is renegotiated, integrating forms of individuality and competence through participation in work activities (p.186).

According to Frost and Robinson (2007), as well as other theorists they drew on for their conclusions about multi-agency responsibility (e.g. Wenger, 1998), roles within multi-agency safeguarding teams are frequently re-negotiated over time. These changes are influenced by professionals' practice, participation (including shared experiences), and procedures for mental health identification (Frost & Robinson, 2007). These revised roles are often necessary for the full team to adapt to the child's changing conditions and provide adequate care.

Therefore, while the data in this thesis suggests that teachers face uncertainty about their roles and responsibilities in predicting suicide risk (potentially raising concerns and negatively impacting the care of students) it also highlights the regular re-negotiation and adjustment of roles over time, particularly among those teachers who hold dual positions (e.g., as both a maths teacher and a SENCO). Whether this reflects a routine process of role re-negotiation, necessary for providing adequate care, or a more troubling trend of increasing complexity in school mental health, especially with the introduction of technology, these shifts have ethical implications for both teachers and students.

Helping teachers adapt to changing roles around mental health and responsibility, particularly in relation to new technology, might be supported through improvements to teacher training. Although a detailed examination of teacher training lies beyond the scope of this thesis, it is important to note that several teachers raised concerns about the adequacy of the training they receive. While safeguarding training should be delivered on a recurring basis, often annually, teachers reported that such sessions do not always provide sufficiently clear guidance about role boundaries, shared responsibilities, or how emerging tools such as digital monitoring should be managed. Instead, teachers are often required to make independent judgements or learn through their own trial and error.

> *And so the training leaves a lot on the shoulders of individual teachers to make the judgements which also yes, you can say comes with experience and comes with time but*

> *then is also very difficult to inculcate new teachers particularly or indeed in teachers who are very fixed in their ways because they've been working in a certain environment or a certain ways for a long, long time.* (P17)

Teachers suggested that these gaps in training are frequently the result of financial and time limitations.

> *The financial requirements shall we say, get tighter and tighter and tighter and funding gets less and less and less and they're expected to do more with it. And actually, in order to be able to get effective training a lot of the time you need that money to do it.* (P15)

7.1.1.4. Blurring of Geographic Boundaries

The role of the teacher is often confined within the physical boundaries of the school, with responsibilities traditionally defined by that geographic context. However, the introduction of new mental health EdTech tools blurs these boundaries: between school and home, and thus also between teachers and parents. In many ways, the introduction of new technology shifts responsibility away from being strictly context-dependent (e.g., teachers at school, parents at home, clinicians in hospitals) and instead assigns teachers a more context-independent role in monitoring students' mental health.

According to a report by the Centre for Data and Technology (CDT, 2022, p.24): "nearly half of students and teachers in schools that use student activity monitoring report that this monitoring takes place outside of school hours. Only 45 percent of teachers report that student activity monitoring is limited to when school is in session." In addition, in the United States, Laird et al. (2022) showed that 65% of teachers reported monitoring student activity beyond the classroom, for example that the school continued to monitor student internet usage at home, and that the majority of reports were occurring outside of school hours.

Does constant monitoring equate to teachers being held constantly responsible? Within interviews, teachers expressed that they view their teaching responsibilities as being limited by time and space. Specifically, they believe they should not be held accountable for scenarios involving EdTech platforms like CommonX (the fictional example of an EdTech company discussed within interviews to avoid bias) when reports or incidents occur outside of work hours or outside the physical school environment. Their responses reflect a concern that tools like CommonX are increasingly blurring the lines between teachers' school duties and personal life, creating an encroachment on their time and responsibilities beyond the classroom, and they would find this difficult to navigate. For example, within the interviews, one teacher referred to navigating the use of home-based internet usage as "tricky," especially when it comes to understanding where their responsibilities lay in comparison to parents:

> *I guess there's a crossing a boundary between the school and the parent, and I think we would find tricky, particularly if it's on a personal computer* (P14)

7.2. Ethical Challenges

The instances outlined above (where the 'ideal' model of shared responsibility is not fully realised in the current landscape of teachers using EdTech for suicide prediction), raise a number of clinical concerns. For example, in the context of mental health monitoring, where immediate action is often required, role ambiguity, lack of shared commitment, and the exclusion of key stakeholders can result in delayed responses or, in the worst cases, serious consequences such as a student's death. Beyond concerns over the clinical utility of monitoring tools, these instances raise ethical issues introduced in Chapter Two and the literature review, such as student privacy, beneficence, autonomy, and their potential to disproportionately harm or discriminate against certain groups of students (described in more detail in Chapter Two, and by activists and scholars including (ACLU, 2023; Gomes de Andrade et al., 2018; Meredith et al., 2018; Reilly, 2017; Warren & Markey, 2022).

In the following section, I demonstrate how the instances highlighted above (where the multi-agency care model falls short of a 'gold standard' of shared responsibility) are directly linked to specific ethical concerns. This includes: (1) the impact of unclear roles on the medicalisation of the classroom, (2) conflicts of interest among stakeholders, and (3) the exclusion of key stakeholders, which raises fundamental concerns of justice.

7.2.1. Medicalisation of the Classroom

A key difference between the current case and the ideal definition of shared responsibility is the lack of clear boundaries between each party's roles (explained further in section 7.1.1.3). While both the law, and teachers themselves, emphasised the need for role clarity, teachers often found themselves uncertain about their responsibilities, with roles becoming blurred. As a result, the teachers interviewed frequently took on additional duties, such as that of a counsellor/therapist. The ethical implications of this are discussed extensively in education and sociology literature as part of 'the medicalisation of the classroom' / 'the medicalised classroom' (e.g. discussed by Cohen et al. (1983) and expanded upon by Petrina (2006)). These scholars critique how education is increasingly framed as a psychotherapeutic practice, shifting responsibility for mental health interventions from clinicians to educators.

In addition, beyond teachers, when monitoring extends into students' homes (e.g. by tracking search queries on personal devices), parents are inadvertently recruited into this medicalised framework, becoming de facto mental health gatekeepers. For example, the software Impero / Ativion (discussed in Chapter 4), which allows parents to track and monitor their children's devices. This added responsibility can place undue stress on families, particularly those with limited resources (time, economic or otherwise). According to Participant 14,

> *There are some parents that are incredibly invested and have lots of time. And there are other parents that either because they're incredibly busy or because they're very remote. I mean I don't mean just geographically remote, but perhaps their working hours are such that they can't get involved in the day-to-day life of their son or daughter. And that's why they've sent them to us, the responsibility lies on us.* (P14)

Technology companies also often end up playing a medical role in the example of EdTech for suicide prediction. By embedding predictive risk models within their platforms, they position themselves as stakeholders in student mental health, even while remaining classified as a 'non-medical device' (defined by Nehme et al. (2024, p.1) as "informational tools that do not handle patient data or provide recommendations for treatment or diagnosis") to most regulators (similar statements have also been made about Chatbots, e.g. by Nehme et al. (2024)).

Results from this DPhil show that in UK schools using EdTech for suicide prediction, teachers, parents, and technology companies play key roles in interpreting algorithmic risk assessments and making intervention decisions. However, they are not trained mental health professionals and may lack the expertise to assess or diagnose suicidality accurately (Ball & Anderson-Butcher, 2014). This reliance on non-clinical staff raises concerns about the validity of risk predictions, leading to the possibility of false positives (leading to unwarranted distress and unnecessary interventions), or false negatives, which could result in missed warning signs.[27]

7.2.2. Conflicts of Interest

Two factors may contribute to conflicts of interest in the use of EdTech for suicide prediction. First, the multi-agency approach (where multiple parties are involved in a child's care) introduces the potential of tensions between the roles of different groups. Second, the multi-functional nature of the technology (whereby one software is used to assess multiple risks and track metrics like attendance) adds a further level of complexity and conflict. Addressing both the multi-agency approach and the multi-functional nature of the technology is essential to understanding, and mitigating, ethical challenges related to conflict of interest.

For instance, Chapter Four highlights the role of private companies like Impero/Ativion in shaping school mental health practices. While these organisations play a central role in monitoring student behaviour, their prioritisation of business objectives (e.g. the need to profit from student data) raises concerns about the alignment of their goals with the wellbeing of students and clinical best practice**.** In practice, business objectives may involve not only the commercial need to sell their programmes but also the potential for companies to benefit from the data itself, for instance through its use in the development of new AI-driven tools, or for advertising. Although companies operating within the UK are formally bound by data protection and privacy regulations (e.g. GDPR), questions remain regarding the extent to which such frameworks meaningfully constrain

[27] The timing of alerts further complicates the effectiveness of interventions. In some cases, a school counsellor may receive a high-risk notification at 8 PM, raising questions about the feasibility of providing immediate support outside school hours. Without a coordinated crisis response plan, students flagged as high risk may remain vulnerable overnight, highlighting the need for clearer protocols and external mental health partnerships.

how student data can be repurposed, especially when platforms operate across medical (/mental health) and commercial boundaries.

The tension between data as a commodity and data as a common good is explored in multiple research papers, including Galetsi et al. (2023), who examines this ethical tension within medical apps, and Kaplan (2016) who considers ethical concerns about the commodification of wellbeing (and student mental health information which comes from 'big data' sources becoming a resource for technological development). These concerns also raise broader questions about transparency and accountability, particularly given that public money is used to procure such systems. In Chapter 8, I provide recommendations to address this issue, including the suggestion of greater regulatory oversight and clearer requirements for companies to disclose how data is being used, shared, or integrated into product development.

The risk of conflict of interest is further compounded by the dual use of EdTech for both mental health support and behavioural management, as well as for functions related to law enforcement. As explored in Chapters One and Two, schools and other agencies often use the same programs for suicide prevention while simultaneously deploying them for protest surveillance. For example, take the scenario below, where School Z relies on law enforcement as a key resource for intervention, illustrating how these overlapping roles can blur the boundaries between education, healthcare, and policing. *(Please note, this scenario is different to that shown to interview subjects in Chapter 6).*

<u>Current Scenario</u>

School Z has been using the software CommonX to filter online content and predict risk of suicide, behavioural issues, and bullying. On a Thursday evening at 8pm, the school counsellor receives a message to say that one of their students has been deemed high risk of suicide, and has been googling firearms, alongside the phrase "I want to die." The school has a data-sharing agreement with the local police force, and therefore this technology sends an alert to the police to go visit the student at their home. The police arrive within an hour.

This example shows how a system that flags multiple types of risk (for example self-harm or aggression) and is sometimes used for other purposes (such as protest surveillance or anti-terrorism) can unintentionally trigger responses from different agencies. This can increase the chance that a student who needs mental health support is directed toward a behavioural or even criminal intervention instead of appropriate clinical care (Collins et al., 2021).

In addition, this example shows how issues arising from conflicts of interest may also extend to broader ethical questions about justice. For instance, there is a potential risk of criminalising mental health issues for students of colour. Research, such as that by Auguste et al. (2023) highlights that psychiatric illness and violence are often conflated in ways that unfairly impact young people of colour, resulting in both stigmatisation and uneven rates of criminalisation.

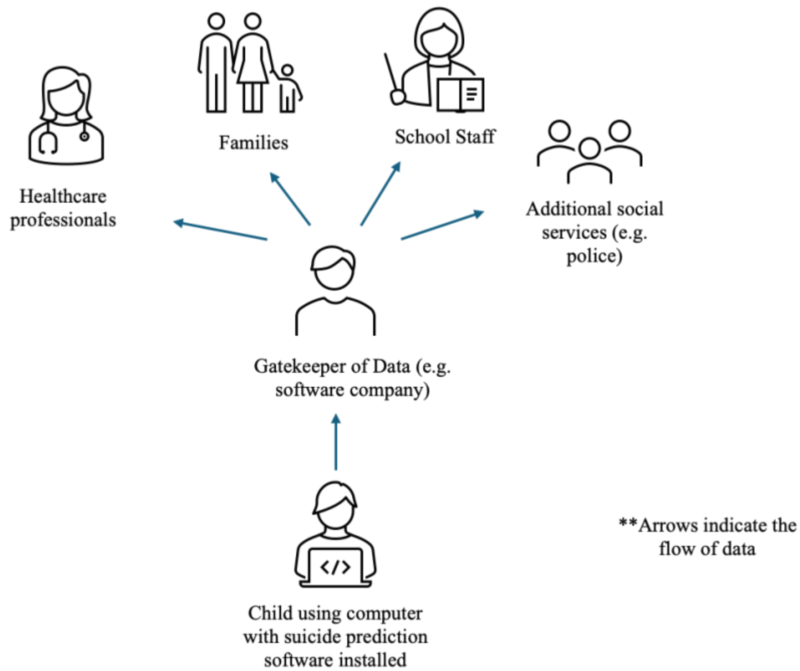
Ultimately, this section highlights two key tensions: (1) between the differing agencies within multi-agency safeguarding teams, and (2) between the dual purposes of the technology itself. It illustrates these tensions through two examples: the conflict between profit and care (i.e., commercial vs. safeguarding actors) and the conflict between criminalisation and care (i.e., teachers vs. law enforcement).

7.2.3. Single Data Controller

In some examples of EdTech for suicide prediction, responsibility may be technically shared, but unevenly distributed. This means multiple stakeholders are involved in a child's care, but control ultimately rests with a single data controller or decision-maker. For example, a suicide prediction tool might centralize responsibility by designating a single data controller, such as a school administrator, who determines when and with whom to share data that is collected on the EdTech programme, or the technology company itself, who collects data and then, using an automated system, decides (a) whether data should be shared, and (b) which parties it should be shared with (See Figure 14).

Figure 14

Software Companies as Gatekeepers of Student Data



The exclusion of parties in favour of a singular data controller is already occurring within the use of EdTech for suicide prediction, as discussed in Section 7.1.1.1. Livingstone et al. (2024) argue that, in instances where companies have this power of data control (or, what they refer to as a 'stranglehold' on data), ethical issues such as privacy and safety arise. Although it may seem intuitive that fewer groups accessing data could enhance privacy, the reality is more complex. When companies hold the majority power to control data, the risk to privacy can increase, especially if their data regulation policies are poorly written or weakly enforced. As Livingstone et al. (2024) note, data regulation policies are being increasingly weakened in the UK; "in post-Brexit UK, a revised data protection regime is proposed that weakens the provisions of the UK GDPR, purportedly to reduce the regulatory burden for business, arguably at the expense of children's rights" (p.5). Furthermore, excluding actors further from decision-making processes not only limits their agency but may also violate the principle of justice, as it denies young people the opportunity to contribute to interventions that directly impact on their lives (Livingstone et al., 2024).

By centralising control within a limited group of actors such as EdTech companies, current frameworks may also fail to respect the autonomy of students' families, diminishing their agency in addressing mental health concerns. Interview data from Chapter Six indicates that teachers observe a shift in parents seeking greater agency, expressing frustration when they feel excluded from decisions about their child's wellbeing. As one teacher recounted, *"God forbid the worst happens and the parents hadn't been informed, you can just imagine, can't you? Like how dare you? Why didn't you disclose this to us at the time?"* (P12). Another teacher relayed a similar sentiment from a parent who insisted, *"You shouldn't be parenting my child. It's my child, and I'll do that as I see fit"* (P10). Together, these accounts highlight the tension between collective decision-making and individual agency. This is what Nollkaemper (2018) refers to as the duality of shared responsibility.

Ultimately, as previously discussed in this thesis, the same software may be used very differently within different schools. These differences (in how the software is used) may depend on factors such as the school's leadership structure, existing mental health policies, and available resources (discussed further in Chapter Five). With thousands of schools across the UK using these programmes, differences in school structure can lead to thousands of unique cases, each with their own unique ethical dilemmas around justice, conflicts of interest, consent (and so forth). It is crucial to consider these factors now to develop an ethical framework for evaluating future iterations of this technology, especially as they become increasingly autonomous.

7.3. Increasingly Autonomous Systems

New software updates may further complicate ideas of shared responsibility, and thereby also this DPhil's ethical analysis. For example, schools may, in the future, integrate EdTech with virtual counselling services, as is the case in the United States with the software Gaggle (expanded to offer virtual therapy in September 2020) (Gaggle, 2025), or fully automated conversational agents (chatbots; e.g. Woebot, Joy, or Wysa; Kretzschmar et al., 2019)), or other non-human virtual therapists (e.g. AI).

I therefore conclude Chapter Seven by suggesting a future scenario, in which virtual, autonomous, therapy is integrated within UK schools, and suggest ways in which the introduction of this new actor (the virtual, autonomous, therapist) may impact both my model of shared responsibility and ethical analysis. The inclusion of this future scenario in my analysis is based on the premise that UK EdTech may evolve in a similar direction to that of technology within the United States, making it a relevant example for exploring potential future developments in digital safeguarding and suicide prediction. It is important to note that, although this scenario is similar to the one used in Chapter Four (/interview guide), the additional element involving automated referral to a virtual therapist was not included in the interview guide used with teachers, as it is not currently available in the UK. As a result, while teachers discussed aspects of automated monitoring and digital intervention more generally, I do not present data on their views regarding referral to a virtual therapist specifically.

Future Scenario - The Virtual, Autonomous Therapist

School Y has been using the software CommonX to filter online content and predict risk of suicide, behavioural issues, and bullying. School Y's counsellor receives an email update from the IT manager, saying that CommonX has a new add-on: virtual therapy. Its use will be covered by the general IT privacy agreement which parents signed earlier in the school year. Later in the week, on a Thursday evening at 8pm, the school counsellor receives a message to say that one of their students has been referred to the online counselling service. When the counsellor clicks the link attached to the email, they discover that this student has googled "why does someone commit suicide" on their personal home computer, and has already been scheduled to see a virtual, autonomous, therapist.

Unlike the hybrid model used currently in UK schools, where teachers and counsellors interpret algorithmic risk assessments, this model introduces a new layer: virtual mental health services. While the introduction of a virtual therapist might seem to alleviate some of the burden on school staff, it also raises questions about the division of responsibilities between educators, counsellors, and virtual mental health services. Therefore, the following analysis is divided into two sections: the first outlines how incorporating virtual, autonomous, mental health services (such as a Chatbot), may strengthen shared responsibility and mitigate some of the risks previously discussed; the second explores how introducing additional autonomous agents could widen the responsibility gap.

7.3.1. Benefits

In the previous section, I connected EdTech for suicide prediction to the concept of the 'medicalisation of the classroom,' where schools are increasingly positioned as spaces for psychotherapeutic intervention. I also explored how, in addition to schools, technology companies were also taking on roles traditionally held by mental health professionals. The integration of virtual therapy into School Y's safeguarding practices further extends the medicalisation of the classroom and EdTech companies, specifically shifting from a model that relied on traditional counsellors for care to one where therapy is embedded directly within the app itself.

It is important to consider the potential benefits of this model, particularly as it relates to the model of shared responsibility considered throughout this thesis. The Future Scenario described on the previous page could allow multiple points of entry to clinical care, particularly beneficial for individuals without access to traditional healthcare, due to geographic, financial, or systemic barriers (e.g. CAMHS waiting lists), as well as 24/7 access (Haque & Rubya, 2023; Khawaja & Bélisle-Pipon, 2023). Additionally, the introduction of a virtual therapist might also allow for greater autonomy of the individual person over their own medical care, particularly if users can access the app independently and choose (/'customise') the type of professional they engage with (Haque & Rubya, 2023; Khawaja & Bélisle-Pipon, 2023).

Furthermore, earlier sections of this thesis have highlighted teachers' concerns that EdTech could assign them medical responsibilities for which they are not qualified. Incorporating a virtual therapist within EdTech might help ensure that teachers are not held accountable for providing care beyond their professional scope. Instead, responsibility would rest with this virtual therapist, which could potentially be subject to regulation by healthcare professionals.

The final, and perhaps most important, (potential) benefit of the virtual therapist is that this type of intervention may prove equally effective (or even more so) than current in-person therapeutic programmes.

Specifically, with regard to the clinical utility of virtual therapists, the introduction of a virtual therapist introduces concerns about the effectiveness of the intervention. While this is a hypothetical scenario with no effectiveness data, it is important to note that the current evidence on the effectiveness of virtual therapists as compared to in-person care is inconclusive. Yet, some individual studies are supportive of virtual therapists. For example, Marcelle et al. (2019) report initial positive findings with platforms like BetterHelp, while Bulkes et al. (2022) suggest telehealth is broadly comparable to in-person care. Others suggest that virtual therapists would be less clinically effective than their traditional counterparts. For example, Prabhakar (2013) argues that a virtual therapist (e.g. one hosted on a EdTech platform) may be less effective compared to a traditional school-based approach as the virtual therapist lacks insight into the student's school environment, support network, and broader ecological system.

7.3.2. Growing Responsibility Gap

While this thesis focused on the roles and responsibilities of various parties in using EdTech for suicide prediction, it focused on traditional partners such as teachers, the NHS, social services, and the parents and families of students. However, the introduction of new technology for suicide prediction, particularly more autonomous systems, further complicates these dynamics.

Some academics argue that the engineers who develop EdTech should be considered a distinct party in the safeguarding process, collaborating with teachers, parents, and other stakeholders while using EdTech as a tool (e.g. similar to urban robots; Nagenborg et al. (2008). Others, however, argue that because these machines possess a degree of autonomy (e.g. they can modify their own operational rules), responsibility cannot be assigned to the software engineers (Santoro et al., 2008). Whether technology can be considered an autonomous actor (and thus if the technology makers themselves should bear responsibility), remains a central debate in AI and tech ethics. This raises important questions about where responsibility should sit within EdTech companies, including whether it is the coder, the engineer, senior leadership, or another clearly identifiable human agent who should be held accountable for the system's actions. If neither the technology nor any individuals are responsible, and/or if the roles of each party are not clearly defined, then a 'responsibility gap' arises (Matthias, 2004). While responsibility gaps are more commonly discussed in accounts that define responsibility as accountability, and/or to discuss "control misalignments" (Veluwenkamp, 2025, p.14), analogous responsibility gap may emerge not only in terms of who is to blame after a failure, but in terms of who holds which duties in the first place.

Matthias (2004), writes about the responsibility gap by saying:

> presently there are machines in development or already in use which are able to decide on a course of action and to act without human intervention. The rules by which they act are not fixed during the production process, but can be changed during the operation of the machine, by the machine itself. … Now it can be shown that there is an increasing class of machine actions, where the traditional ways of responsibility ascription are not compatible with our sense of justice and the moral framework of society because nobody has enough control over the machine's actions to be able to assume the responsibility for them. (p.17)

Scholars (e.g., Matthias, 2004) argue that the introduction of artificial intelligence has contributed to a growing responsibility gap, as roles are not clearly defined and it becomes difficult to determine each party's responsibilities or assign them to a specific individual. According to some, referred to by Tigard (2021) as techno-pessimists, with intelligent artificial programs, there will eventually be no way to deem a specific person responsible, and as such, the gap cannot be bridged.

However, the 'techno-optimists,' on the other hand, have a number of different ways they believe responsibility can still be allocated. Tigard (2021) and Santoro et al. (2008) say that there should be people (e.g., computer scientists) who evaluate the specific ML programs to determine their risks and benefits and create clear and explicit rules about both prospective and retrospective responsibilities. Rahwan (2018) proposes a larger group of people responsible for creating such a list of rules, beyond just computer scientists/technology specialists. Finally, Hellström (2013) says that "society may decide to collectively share responsibility" (p.105), through the democratic process (also, Archard, 2013).

The full impact of this will be explored in the next chapter. However, it is already evident that the uptake of virtual therapists further complicates existing models of care. Not only are decisions about risk now being made by virtual agents, but these agents are now also making decisions about treatment. This introduces a range of new actors such as computer scientists, technology specialists, and other stakeholders, into frameworks of shared responsibility, complicating the already complex systems of care around a student (illustrated in Chapter 5). Ultimately, this may lead to two growing risks. First, there is a risk that too many parties become involved in children's care, making it increasingly likely that no single party can ultimately be held responsible. Second, even when individuals are assigned role-responsibilities, they may not feel a strong sense of responsibility, especially if their role seems like only a minor part within a wider network of distributed duties.

7.3.3. Complicated Processes of Data Sharing

Including medical data, and medical decision-making, within a third-party virtual therapy service in this way complicates data-sharing. For instance, a virtual therapy service (such as CommonX, in the example above) will likely hold sensitive information about students, including information about their clinical care and suicide risk. This increases the risk of breaches or misuse (recent examples of data misuse, including data leaks, by BetterHelp were explored in work by Henson et al. (2019)).

In addition, the risk around data sharing may be exacerbated due to a lack of clear informed consent policies. Consider the programme NetSupport DNA, which collects student data within the student's home (Chapter Four). While a school may justify the use of this programme under a general IT privacy agreement signed by parents, it remains unclear whether students and families fully understand the extent of data collection and its implications.

Finally, questions arise as to whether a virtual therapist, held within a EdTech platform (which is also used for non-therapeutic purposes, such as school administration), constitutes medical or non-medical platform [28], which may influence the regulatory frameworks governing EdTech programmes and data sharing between parties. While the inclusion of a third-party therapy

[28] As a reminder, a 'non-medical device' is defined by Nehme et al. (2024, p.1) as "informational tools that do not handle patient data or provide recommendations for treatment or diagnosis."

provider is not inherently problematic, these challenges highlight the need for clear and robust data-sharing policies to mitigate risks and ensure effective support.

Overall, the scenario above highlights how the introduction of virtual therapists may enhance some of the ethical dilemmas presented in my earlier analysis. Data sharing between virtual therapists, schools, and EdTech companies, while potentially beneficial for streamlining interventions, may introduce new risks related to privacy and consent. In addition, the medicalisation of the classroom becomes even more pronounced, as schools and technology companies expand their roles in mental health care. Ultimately, these developments underscore the need for clear ethical guidelines and protocols to ensure that EdTech prioritises student wellbeing. This is relevant not only in its current forms but also in future applications. For instance, while this chapter focused on virtual therapists, additional challenges may arise with the introduction of fully automated conversational agents (chatbots; e.g. Woebot, Joy, or Wysa; Kretzschmar et al., 2019), or other non-human virtual therapists (e.g. AI).

7.4. Conclusion

Ultimately, this chapter has demonstrated two key points. First, it highlighted how current practices of shared responsibility diverge from an 'ideal' framework, characterised by unclear roles, limited shared commitment, the exclusion of key stakeholders, and the blurring of geographic boundaries. It also examined the ethical dilemmas that these discrepancies may give rise to, such as conflicts of interest, the medicalisation of the classroom, and broader issues of justice.

Second, this chapter explored a potential future scenario in which mental health support becomes increasingly virtual (e.g. such as through the introduction of virtual therapists) further complicating questions of role responsibility. This section discusses potential benefits of an increasingly virtual platform, before exploring areas of concern, which include a growing responsibility gap and complications related to data sharing.

Mapping how responsibility is shared in the context of EdTech for suicide prediction has been a central theme throughout this thesis, for two main reasons. First, the rapid pace of technological development means there are currently no clear guidelines outlining how teachers should engage with these tools or what an effective partnership model should entail. Second, there are, as yet, no legal requirements designating teachers as mandated reporters (both for suicide, but also for other safeguarding concerns, e.g. child abuse). As a result, although both the DfE and professional bodies provide statutory guidance on how teachers should act, decisions about whether (and how) teachers should take on responsibility for safeguarding are often left to individual schools. Chapter 8, which concludes the thesis, will therefore offer policy recommendations that address these challenges: both acknowledging both the absence of existing guidelines and the possibility of future legal mandates around reporting. It will also outline key directions for future research.

Chapter 8. Conclusion

In this final chapter, I begin by outlining the contribution my research makes to the broader field of EdTech for suicide prediction (8.1), showing how this thesis provides a framework for understanding teachers' responsibilities when using said technologies. Within this section I identify key barriers to effective shared responsibility and show how my methodology centres teachers' lived experiences within complex school systems.

This is followed by a clearly delineated section on recommendations (8.2). Within this section I strongly recommend that teachers' roles remain as primarily referral-based rather than intervening on suicide risk or poor mental health. I then outline additional recommendations, grounded in ethical theory, including the need to integrate digital monitoring tools within the wider ethos and statutory responsibilities of schools, coordinate with school mental-health teams, and greater obligations on EdTech companies regarding open data sharing, transparency about their operations, and accountability for the use of public funds.

In the second half of this chapter, I address the limitations of the study, (8.3), before concluding with potential directions for future research (8.4), including expanding the project through larger and more diverse samples, engaging parents and students, examining school ethos and prevention practices, conducting national surveys of EdTech use, and investigating new conceptual frameworks such as abolition and trust. The thesis ultimately concludes by questioning whether EdTech for suicide prediction should be used at all.

8.1. Contribution to Academic Field

I began this DPhil by reviewing the EdTech programmes themselves (Chapter 4), and relevant policy documents (Chapter 5), before moving on to report on interviews with teachers (Chapter 6). I found that while theories of teacher responsibility (particularly those drawing on Hart's 1968 framework of role responsibility), emphasised the significance of clearly defined roles, the actual role of the teacher in using EdTech for suicide prediction remains largely ambiguous. This ambiguity is partly due to the rapid development of new technologies (and the inability of regulatory bodies to keep pace) and reflects broader tensions regarding the role of a teacher: e.g.

as *one-who-surveilles* versus *one-who-educates*. Relatedly, policy documents also position teachers and schools as key actors in promoting good mental health and well-being, recognising that learning is undermined when students are stressed, bullied, anxious, or under excessive pressure. This further complicates teachers' responsibilities, as expectations around care and emotional support intersect with those emerging from new predictive technologies. These tensions, and related role-uncertainties, are embedded both within policy and teaching practice, as well as the technology itself.

Regarding the latter point, different software programmes differed in the degree to which they positioned teachers as active decision-makers (for determining intervention pathways) or passive facilitators (e.g. having view-only access to a student's account). Different software companies differed both in how they explicitly referred to teachers' roles, for example, in advertising material, and implicitly, as explored in Chapter 4's analysis of how 'high-risk' cases were triaged, and by whom.

In my analysis, I also identified a discrepancy between the normative ideals of shared responsibility and how shared responsibility was enacted in practice. While my original definition of shared responsibility emphasised equal contribution, mutual commitment, and clearly defined roles (Chapter 3, explored in definitions by Hart, 1968; Kon et al., 2017), these principles were rarely realised in practice. Chapter 7 showed how the gap between normative theory and practice manifested, including: the exclusion of key stakeholders, a lack of shared commitment, ambiguous role definitions, blurred geographic boundaries, and unresolved questions about whether responsibilities should be shared at all. Additionally, there was considerable uncertainty regarding how access to the software should be allocated, and to which other parties (for example, parents, social services, and to the students themselves).

8.2. Recommendations

Within this section, I set out four concrete recommendations for researchers, practitioners, and policymakers: (8.2.1) promoting a model of referral rather than direct intervention and emphasising shared (not individual) responsibility; (8.2.2) addressing the key barriers that currently hinder effective shared responsibility in school settings; (8.2.3) strengthening ethical and transparent practice in EdTech governance; and (8.2.4) integrating socio-political and historical contexts into the evaluation and development of mental-health technologies in education.

8.2.1. Recommendation 1: Referral, Not Intervention, and Shared, Not Individual

In response to my primary research question, *What responsibilities do teachers and schools have when using technology for suicide prediction, and how do these compare to what their responsibilities should be?*, this thesis demonstrates that teachers should be positioned primarily as responsible for referring safeguarding concerns, rather than intervening directly. More broadly, both schools as institutions (led by governing bodies and Senior Leadership Teams), and individual teachers (whose responsibilities may vary depending on their specific roles), were found to be tasked with the general promotion of students' mental health and well-being, as well as with monitoring and referring when there is perceived suicide risk. This conclusion is supported by interview data and by analysis of teacher guidance and policy materials (Chapters 4–7). Although a small number of teachers felt their role extended into treatment and intervention, these views were the exception. These wider expectations provide an important backdrop to understanding how responsibilities are distributed when suicide-prediction technologies are introduced.

In addition, given that teachers increasingly interact with a wide range of actors (including parents, mental health professionals, social services, and software providers) this thesis argues that responsibility must be understood as *shared* rather than individual. Throughout the empirical chapters, the thesis provides definitions, methodological techniques, and analytical frameworks (including systems theory) that help conceptualise and evaluate this shift toward shared responsibility. However, while these conceptual tools establish the foundations for understanding how responsibility is distributed, there remains a need for policymakers to develop specific, detailed, and well-thought-out frameworks that can translate these insights into practice. Such policy frameworks should clearly articulate roles, boundaries, and expectations across the network of actors involved, ensuring that shared responsibility is not only theorised but operationalised in ways that support teachers and students, and align with broader systems of care.

8.2.2. Recommendation 2: Overcoming Key Barriers to Shared Responsibility

This thesis recommends that policymakers, researchers, and practitioners work to overcome the two key barriers that currently prevent effective implementation of shared responsibility: (1) unclear frameworks, and (2) limited resources.

Clearer frameworks are needed to define how responsibility should be distributed across teachers, parents, EdTech companies, and other stakeholders. These frameworks must specify where responsibility begins and ends (e.g., school grounds versus the home), outline appropriate data-sharing protocols, and establish accountability mechanisms. They must also account for moderating factors that shape teachers' willingness or capacity to assume responsibility, including workload, available resources, and structural inequalities (see Chapters 6 and 7).

One pathway for clarifying shared responsibility, and developing these clear frameworks, is to ensure that all relevant stakeholders (e.g., teachers, students, parents, and EdTech developers) are meaningfully consulted when developing future policies and guidelines for suicide-prediction EdTech. This recommendation is supported by existing research: Hulme et al. (2009) highlight the value of collaborative consultation in establishing clear professional boundaries, and Salmon (2004) found that referrals to specialist services became more effective when school staff were given opportunities to consult directly with specialists.

A second pathway for clarifying shared pathway is to integrate digital tools within the wider statutory and ethical responsibilities of schools, rather than treat them as a standalone technical solution. Their use must align with safeguarding duties, data-protection requirements, and a school's pastoral ethos, as well as other, pre-established shared-responsibility pathways. Specifically, this could include close coordination with Mental Health Support Teams (MHSTs), which are National Health Service (NHS) teams based in schools and colleges that provide early, evidence-based mental health support, deliver interventions such as CBT, advise staff, and help embed a whole-school approach to wellbeing (Ellins et al., 2023). By working alongside MHST practitioners, and the MHST system of shared responsibility, schools can ensure that risk signals or alerts produced by digital tools are interpreted by trained professionals and integrated into established pathways of support, avoiding inappropriate escalation or punitive responses.

Next, in addition to the need for clearer frameworks, this thesis finds that a lack of resources (both individual and on a societal level) presents a significant underlying barrier to teachers enacting their 'ideal' model of shared responsibility. The availability and distribution of resources may be fundamental to how responsibility is understood, allocated, and sustained.

In terms of individual resources, teacher interviews revealed that, although teachers recognise pastoral and mental health care as part of their professional responsibility, they often feel unable to fulfil this duty due to limitations in economic resources, time, and emotional energy (Chapter 6). Time and emotional capacity are also closely linked to staffing shortages, which represent another resource-related challenge faced by schools. At the societal level, CAMHS budgets have been significantly reduced (Local Government Association, 2023; Ofsted, 2020), and these services face substantial backlogs. As a result, teachers may feel compelled to intervene rather than refer, or to engage with CAMHS differently, simply because the system is overstretched.

Policymakers should build on these findings by considering, in specific and practical terms, how resources should be allocated and managed to strengthen shared responsibility. This includes developing a detailed resource map or economic cost-benefit model (covering individual, community, and societal costs) to assess whether hiring additional staff, increasing pastoral capacity, or implementing other support mechanisms would improve shared responsibility models or shape how teachers understand their responsibility for student mental health. Resource planning should also address broader structural issues, such as CAMHS capacity and local authority funding, to ensure that referral pathways are viable in practice and that teachers are not compelled to intervene due to systemic shortfalls.

8.2.3. Recommendation 3: Enhancing Ethical and Transparent Practice in EdTech Governance

My third recommendation is emphasising the need for more robust ethical standards in the governance of digital technologies used in education. Central to this is a call for stronger obligations on EdTech companies to operate with transparency, particularly when their products influence safeguarding decisions or receive public investment.

To achieve this, EdTech providers should adopt open-data principles that allow independent researchers, school leaders, and policymakers to examine how their systems work in practice. This means offering accessible explanations of what data is collected, how algorithms operate, and how outputs inform educational or safeguarding decisions, reflecting broader calls for explainable and auditable AI in the public sector (Floridi & Cowls, 2019). According to Raji and colleagues (2020), independent auditing of algorithmic processes should be end-to-end, and standardised, enabling scrutiny of accuracy, bias, and unintended consequences (including false negatives and positives). Companies benefiting from public funds, in particular, must demonstrate the effectiveness of their tools through evidence rather than marketing claims. Ultimately, these measures would empower educators, families, and oversight bodies (including government) to determine whether digital monitoring technologies genuinely enhance student wellbeing.

These recommendations speak directly to the core ethical issues originally raised in Chapter 2 (and brought up throughout the thesis), including concerns about opacity, fairness, and the potential misuse of sensitive data. These recommendations on transparency also have implications for informed consent. Students and their families must be given clear, age-appropriate explanations of what data is being collected, how it will be used, and who will have access to it, rather than being asked to agree to opaque or generic 'internet use' policies. Ultimately, ethical practice should therefore include clear commitments to open data sharing, transparent reporting on system operations, and full accountability for the deployment of tools.

8.2.4. Recommendation 4: Integrating Socio-Political and Historical Contexts into EdTech Evaluation and Policy

The final recommendation of this thesis is directed toward researchers, practitioners, and policymakers, and concerns the need to incorporate political, economic, and historical contexts into the evaluation of key mental-health technologies.

First, I argue that applying a Systems Theory approach provides a valuable framework for understanding how EdTech is embedded within school environments, ensuring that any assessment of its impact accounts for contextual factors such as a school's values, ethos, organisational structures, and its connections to wider systems of community mental-health support. In this DPhil, the Systems Approach has been used to ensure such context-sensitivity, demonstrating that intervention pathways are often school-dependent, and that institutional characteristics (especially a school's ethos) shape how teachers understand their own responsibilities (Chapter 6) and how they conceptualise shared responsibility (Chapter 7). Recognising the school as an embedded institution with its own practices and constraints is therefore crucial for developing effective policy. A systems-based perspective, such as the one outlined in Chapter 5, can generate a more accurate picture of how schools are organised, thereby informing school-level decisions about technology use, suicide prediction, and suicide prevention more broadly (Institute of Medicine (US) Committee on Pathophysiology and Prevention of Adolescent and Adult Suicide, 2001). This approach also underscores the importance of building flexibility into future decision-making processes and into the design and implementation of technology programmes themselves.

Building on this, I also recommend that future ethical analyses and policy decisions explicitly integrate macro-level social, political, economic, and historical dimensions, complementing traditional bioethical principles (Beauchamp & Childress, 2013) and assessments of clinical or educational utility. This socio-political-economic-historical approach aligns with Green's (2021, p.1) "socio-technological" ethics, Birhane's (2021) work on relational ethics, and the broader scholarship reviewed in Chapter 5, collectively offering a richer and more context-sensitive foundation for evaluating responsibility, risk, and the appropriate use of mental-health technologies in schools.

By using a systems approach within my own thesis, I have been able to account for the broader context in which these technologies operate, and this has generated several additional, targeted recommendations. One such recommendation is the need for equity-focused and anti-racist policymaking to prevent harm and discrimination, particularly in light of the historical (mis)use of related technologies described in Chapters 1 and 2. A second recommendation is that the UK develop its own context-sensitive framework for evaluating and governing EdTech, as much of the existing research on socio-political dimensions of educational technologies is based on US schooling systems. Any UK policy should be informed by UK-specific factors, including the historical deployment of surveillance technologies within the PREVENT strategy and the chronic underfunding of CAMHS (as outlined in Chapter 1 and 2).

8.3. Limitations

Despite the richness of the data developed, doctoral research projects are limited by time, resources, and experience. As such, my model of responsibility within the context of EdTech for suicide prediction remains incomplete, and more work could be done to develop this project further. The following explores four fundamental limitations of this research project (the lack of evaluation, limitations in participation sampling, limitations of online interviewing, and the limited insight available into the mesosystem).

8.3.1. Limitation 1: The Need for Evaluation

Within this thesis I conducted no statistical or evaluative analysis on program efficacy (e.g., PPV) or clinical utility. Such an analysis would have enabled a better assessment of these tools; however, this would require more time and data than currently available.

This study is not unique in this limitation - there is a striking lack of transparency and robust evaluation surrounding EdTech programmes for suicide prediction within the current literature. This opacity surrounds the tools' clinical utility (as explained in Chapter 4). I found only one paper reporting on outcome measures: a PhD dissertation from the US; Shelton, 2022), as well as reporting on general algorithmic decision making. There is a lack of transparency regarding what data is used within these tools, how data is utilised, where data is stored, and how risk decisions are made (the document 'A Blueprint for Education Data' further describes this opacity and includes suggestions to improve transparency in the broader EdTech context; Digital Futures Commission, 2023).

Within this DPhil project, Chapter 4 highlighted the lack of transparency within EdTech companies' publicly available data, while Chapter 7 clearly discussed why this opacity limits teachers' ability to successfully implement ideal versions of shared responsibility. However, only by fully understanding how EdTech companies operate (including their intervention pathways and positive predictive value (PPV)) can regulators guide teachers on engaging with these companies and distributing roles and responsibilities.

8.3.2. Limitation 2: Participant Sampling

As discussed in the conclusion of Chapter 6, this project faced a number of limitations arising from both the COVID-19 context and the time constraints of a DPhil programme. First among these was the restricted number of participants who could be recruited. The small sample limited the range of professional experiences represented (for example, it prevented comparisons between safeguarding leads and non-safeguarding teachers), and meant that several important stakeholder groups, such as parents, EdTech developers, and young people themselves, could not be included. Their absence ultimately constrains the extent to which the findings can speak to the broader, multi-stakeholder ecosystem surrounding the use of EdTech for suicide prediction.

In addition, the reliance on existing professional networks for recruitment (often the most feasible option) introduced a further layer of bias. Teachers who were already known to the research team, already engaged in related initiatives, or geographically close to Oxford were disproportionately likely to participate. This convenience-based recruitment strategy, while practically necessary, means that the resulting sample is unlikely to be fully representative of teachers' experiences nationally or across different school types. Teachers with prior involvement in similar projects may also have been more comfortable with research processes or more predisposed to reflect on technology use.

8.3.3. Limitation 3: Online Interviewing

The need for online interviewing limited what types of teaching practice I was able to see/record. Because school-based engagement was effectively impossible during much of the research period (due to COVID-19), the project relied heavily on online interviews. While these interviews provided valuable insights, they also introduced limitations: they prevented observation of classroom practice, restricted opportunities for informal follow-up or contextual probing, and made it difficult to capture the organisational dynamics or material conditions of schools. Online-only interviews therefore provide a necessarily partial picture of teachers' work with digital tools.

8.3.4. Limitation 4: Limited Systems Analysis

A further critical limitation is that my systems approach addressed only a subset of Bronfenbrenner's ecological levels, and in particular failed to engage meaningfully with the mesosystem. As demonstrated throughout the thesis, school ethos plays a vital role in shaping practices and decisions around suicide prediction, and the absence of direct examination of this level is therefore a notable constraint. This limitation was largely driven by the fact that I was unable to enter schools during the research period, due to school closures, as well as by the time restrictions inherent in a DPhil project.

8.4. Future Research

Within this final section, I offer suggestions for strengthening the research and addressing the thesis's limitations, including extending this project through a larger and more diverse sample, interviewing parents and students, and examining the influence of school ethos and prevention practices. Additional possibilities include conducting a national survey of EdTech use in schools to better understand the landscape, and considering emerging developments such as AI or chatbot therapists. I also propose two more substantial directions for future research that arise from this project (trust and abolition) both of which offer important conceptual frameworks for further exploration in the context of school-based mental-health technologies.

8.4.1. Project Extensions

Throughout this thesis, several opportunities for further inquiry have been identified. This section brings them together in a coherent set of proposed extensions to the project.

8.4.1.1. Extension 1: More Participants

It is important to ensure that a wider range of teachers is represented in future research, including those from different subject areas, levels of seniority, and school contexts, as their responsibilities and experiences with EdTech can vary considerably. A larger and more diverse teacher sample would also make it possible to explicitly compare different types of teachers. This was something was not feasible in this thesis due to the small sample size, which limits meaningful within-group analysis.

In addition, the interviews conducted for this DPhil capture only teachers' perspectives, with parents' and students' views presented indirectly. Although teachers' accounts constitute a valuable and novel contribution, incorporating perspectives from other stakeholder groups would significantly enhance understanding of the complex phenomenon of EdTech for suicide prediction in schools. In particular, parents' roles, expectations, and responsibilities warrant closer examination. As many school-based mental-health interventions position parents as key partners in responding to risk, their insights are crucial for shaping the design and implementation of responsible and effective predictive systems. For example, Ice et al. (2014) demonstrate that involving parents (either collaboratively or as primary agents) can reduce dropout rates and improve intervention outcomes, underscoring the importance of integrating parental perspectives into future research.

8.4.1.2. Extension 2: Mapping National and School-Level Procedures

A further extension involves developing a deeper understanding of national and school-level procedures and organisational structures, particularly in relation to two key areas: (1) school management, including how procurement decisions are made, how budgets and priorities are set, and how schools select and justify the adoption of mental-health technologies; and (2) school ethos, examining how values and wellbeing commitments are interpreted and enacted in everyday practice.

Two methodological approaches could support this work. First, a national survey could investigate broader procurement trends, levels of adoption, perceived challenges, and regional or institutional variations. Such data would help situate individual schools' experiences within wider systemic patterns, strengthening the basis for policy and ethical guidance around the use of predictive technologies in education.

Second, conducting school-based observations would offer insight into how communication and decision-making actually unfold within the students' mesosystem. For example, how information flows between teachers, safeguarding teams, senior leadership, EdTech companies, and parents.

This would provide a richer, practice-based understanding of how EdTech tools are interpreted and used within the real organisational life of schools.

8.4.1.3. Extension 3: Accounting for Increasingly Autonomous Agents

A third avenue for extending this project, is to continue ‘future-proofing’ analyses of these tools by explicitly examining the emerging role of increasingly autonomous AI agents.

Building on the insights of Chapter 7, future work should investigate how ongoing software updates may further complicate shared-responsibility models, for example by modifying data-sharing pathways or enabling the processing of additional forms of student data such as social media activity. Schools may also begin integrating EdTech systems with virtual counselling or AI-supported therapeutic services, which would introduce an additional agent into the ecosystem. As these systems become more autonomous, the gaps in responsibility identified in this thesis are likely to widen, exacerbating concerns around diluted accountability or the ‘problem of many hands’ (van de Poel et al., 2012), as explored in Chapters 3 and 7.

Given the rapid pace of technological development, periods of regulatory and practical lag are inevitable, meaning there will be times when no clear guidance exists regarding how teachers should engage with these tools or how effective partnership and responsibility-sharing models should operate. For this reason, future research should prioritise conducting rapid reviews and regularly updated mapping exercises, both within academia and in government, to ensure ongoing oversight and to support the development of robust and adaptive frameworks that can respond to continual technological change.

8.4.2. Substantial Direction for Future Research 1: Trust

Trust is a key bioethical concept in relation to shared responsibility (Wahlstrom & Louis, 2008). For shared responsibility to function effectively, each party must not only have the necessary resources to fulfil their individual role but also trust that others will competently and conscientiously carry out their responsibilities. For instance, when a teacher refers a student to external support, they must trust that those receiving the referral (e.g. CAMHS) will act appropriately and effectively.

However, researchers including Kenny (2001) have shown that teachers often express a lack of trust in external organisations to adequately support their students. While Kenny's findings predominately pertain to child protection services and abuse, similar concerns may arise in the context of school mental health and EdTech companies. Future research should explore the role of trust in shared responsibility models; especially how trust may influence collaboration between educators, mental health professionals, and the technological systems or EdTech providers.

8.4.3. Substantial Direction for Future Research 2: Abolition

Much of the discussion in this thesis has operated under the assumption that as EdTech is already integrated in school technology systems, the role of this project (and subsequent academic and policy work), should be to determine how teachers should use the technology and to create guidelines on how responsibility should be shared between teachers and other groups (e.g. parents, social services). However, a growing number of voices argue that we should instead be working toward the removal of this technology from schools altogether (Hooper, 2015; Thuy Vo & Aldhous, 2019; Warren & Markey, 2022). As discussed in the early chapters, these calls for abolition emerge in response to serious ethical concerns: limitations on individual liberty and justice, the anti-Muslim and anti-Black application of surveillance tools, and the use of student data in collaboration with law enforcement. As a result, the use of EdTech within schools is often met with protests and backlash; not only from educators, students, and the general public, but also civil rights organisations such as the ACLU.

In future research projects, abolition should be considered and evaluated. However, even if EdTech programmes are ultimately removed, suicide prediction and monitoring tools will almost certainly continue to appear in schools in new forms. This is because, in many ways, the current tools represent a continuation of earlier instruments, such as the C-SSRS. Suicide prediction technologies have existed in the past and are likely to evolve and re-emerge. What will remain constant is that teachers will continue to bear safeguarding responsibilities for their students.

Therefore, many of the insights developed in this thesis regarding responsibility will remain relevant, even if these EdTech programmes change or are replaced. For instance, the need to clearly define a teacher's role as one of referral rather than direct intervention will persist, as clarifying this role not only protects teachers from overreach and burnout but also affirms their place within a broader ecosystem of shared responsibility that includes parents, clinicians, and social services, and of course, the young people themselves.

Bibliography

Access Now. (2023). *Humanitarian Tech Mapping Survey*. https://www.accessnow.org/Humanitarian-Tech-Mapping-Survey/.

ACLU. (2023). *Digital Dystopia The Danger in Buying What the EdTech Surveillance Industry is Selling*. https://assets.aclu.org/live/uploads/publications/digital_dystopia_report_aclu.pdf

Ahola-Launonen, J. (2015). The Evolving Idea of Social Responsibility in Bioethics. *Cambridge Quarterly of Healthcare Ethics*, *24*(2), 204–213. https://doi.org/10.1017/S0963180114000516

Alderson, P. (2007). Competent children? Minors' consent to health care treatment and research. *Social science & medicine (1982)*, *65*(11), 2272–2283. https://doi.org/10.1016/j.socscimed.2007.08.005

Ansell, C., & Gash, A. (2008). Collaborative Governance in Theory and Practice. *Journal of Public Administration Research and Theory*, *18*(4), 543–571. https://doi.org/10.1093/jopart/mum032

Archard, D. (2013). Dirty Hands and the Complicity of the Democratic Public. *Ethical Theory and Moral Practice*, *16*(4), 777–790. https://doi.org/10.1007/s10677-012-9387-y

Arksey, H., & O’Malley, L. (2005). Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology*, *8*(1), 19–32. https://doi.org/10.1080/1364557032000119616

Arthur, J. (2015). Extremism and Neo-Liberal Education Policy: A Contextual Critique of the Trojan Horse Affair in Birmingham Schools. *British Journal of Educational Studies*, *63*(3), 311–328. https://doi.org/10.1080/00071005.2015.1069258

Asan, O., & Choudhury, A. (2021). Research Trends in Artificial Intelligence Applications in Human Factors Health Care: Mapping Review. *JMIR Human Factors*, *8*(2), e28236. https://doi.org/10.2196/28236

Aston, J., Davies, E., Guijon, M., Lauderdale, K., & Popov, D. (2022). *The education technology market in England Research report*. Government Social Research. https://assets.publishing.service.gov.uk/media/636e7717e90e07186280f7cf/Edtech_market_in_England_Nov_2022.pdf

Auguste, E., Bowdring, M., Kasparek, S. W., McPhee, J., Tabachnick, A. R., Tung, I., & Galán, C. A. (2023). Psychology's Contributions to Anti-Blackness in the United States Within Psychological Research, Criminal Justice, and Mental Health. *Perspectives on Psychological Science*, *18*(6), 1282–1305. https://doi.org/10.1177/17456916221141374

Ayer, L., Nickerson, K., Grumet, J. G., & Hoover, S. (2022). *Effective Suicide Prevention and Intervention in Schools* (pp. 31–40). https://doi.org/10.1007/978-3-031-06127-1_4

Ayers, L., Boudeaux, B., Welburn Paige, J., Holmes, P., Blagg, T. L., & Mendon-Plasek, S. J. (2023). *Artificial Intelligence Based Student Activity Monitoring for Suicide Risk*. RAND Corporation.

Baca-García, E., Perez-Rodriguez, M. M., Basurte-Villamor, I., Saiz-Ruiz, J., Leiva-Murillo, J. M., de Prado-Cumplido, M., Santiago-Mozos, R., Artés-Rodríguez, A., & de Leon, J. (2006). Using Data Mining to Explore Complex Clinical Decisions. *The Journal of Clinical Psychiatry*, *67*(07), 1124–1132. https://doi.org/10.4088/JCP.v67n0716

Ball, A., & Anderson-Butcher, D. (2014). Understanding Teachers' Perceptions of Student Support Systems in Relation to Teachers' Stress. *Children & Schools*, *36*(4), 221–229. https://doi.org/10.1093/cs/cdu017

Barak-Corren, Y., Castro, V. M., Javitt, S., Hoffnagle, A. G., Dai, Y., Perlis, R. H., Nock, M. K., Smoller, J. W., & Reis, B. Y. (2017). Predicting Suicidal Behavior From Longitudinal Electronic Health Records. *American Journal of Psychiatry*, *174*(2), 154–162. https://doi.org/10.1176/appi.ajp.2016.16010077

Bason, B. (2021). *Bark for schools response to inquiry, letter to Elizabeth Warren, Edward J. Markey, and Richard Blumenthal*. https://www.warren.senate.gov/imo/media/doc/2022.07.20%20Letter%20to%20Student%20Activity%20Monitoring%20Companies%20re%20Reproductive%20Health%20Care%20Data.pdf

Beauchamp, T., & Childress, J. (2001). *Principles of biomedical ethics*. Oxford University Press.

Beauchamp, T., & Childress, J. (2013). *Principles of bioethics* (7th ed.). Oxford University Press.

Beck, A. T., Kovacs, M., & Weissman, A. (1979). Assessment of suicidal intention: The Scale for Suicide Ideation. *Journal of Consulting and Clinical Psychology*, *47*(2), 343–352. https://doi.org/10.1037/0022-006X.47.2.343

Berry, H. L., Waite, T. D., Dear, K. B. G., Capon, A. G., & Murray, V. (2018). The case for systems thinking about climate change and mental health. *Nature Climate Change*, *8*(4), 282–290. https://doi.org/10.1038/s41558-018-0102-4

Besley, T., & Peters, M. A. (2019). *Teaching, Responsibility, and the Corruption of Youth*. Brill | Sense. https://doi.org/10.1163/9789004380776

Bierhoff, H. (2002). Just World, Social Responsibility, and Helping Behavior. In *The Justice Motive in Everyday Life* (pp. 189–203). Cambridge University Press. https://doi.org/10.1017/CBO9780511499975.011

Birhane, A. (2021). Algorithmic injustice: a relational ethics approach. *Patterns*, *2*(2), 100205. https://doi.org/10.1016/j.patter.2021.100205

Boxer, P. (2010). Covariation of self- and other-directed aggression among inpatient youth: continuity in the transition to treatment and shared risk factors. *Aggressive Behavior*, *36*(3), 205–217. https://doi.org/10.1002/ab.20343

Bradshaw, C. P., Sawyer, A. L., & O'Brennan, L. M. (2009). A Social Disorganization Perspective on Bullying-Related Attitudes and Behaviors: The Influence of School Context. *American Journal of Community Psychology*, *43*(3–4), 204–220. https://doi.org/10.1007/s10464-009-9240-1

Braun, V., & Clarke, V. (2022). Conceptual and design thinking for thematic analysis. *Qualitative Psychology*, *9*(1), 3–26. https://doi.org/10.1037/qup0000196

Brent, D. A., & Mann, J. J. (2005). *Family genetic studies, suicide, and suicidal behavior.* American Journal of Medical Genetics Part C: Seminars in Medical Genetics, 133C(1), 13–24.

Broer, T. (2022). The Googlization of Health: Invasiveness and corporate responsibility in media discourses on Facebook's algorithmic programme for suicide prevention. *Social Science & Medicine*, *306*, 115131. https://doi.org/10.1016/j.socscimed.2022.115131

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Harvard University Press.

Bryan, C., & Lurye, S. (2025, March 12). Student privacy vs. safety: The AI surveillance dilemma in WA schools. *The Seattle Times*. https://www.seattletimes.com/education-lab/student-privacy-vs-safety-the-ai-surveillance-dilemma-in-wa-schools/

Bryan, H. (2017). Developing the political citizen: How teachers are navigating the statutory demands of the Counter-Terrorism and Security Act 205 and the Prevent Duty. *Education, Citizenship and Social Justice*, *12*(3), 213–226. https://doi.org/10.1177/1746197917717841

Bulkes, N. Z., Davis, K., Kay, B., & Riemann, B. C. (2022). Comparing efficacy of telehealth to in-person mental health care in intensive-treatment-seeking adults. *Journal of Psychiatric Research*, *145*, 347–352. https://doi.org/10.1016/j.jpsychires.2021.11.003

Calman, S. K. C., & Royston, G. (1997). Personal paper: Risk language and dialects. *BMJ*, *315*(7113), 939–942. https://doi.org/10.1136/bmj.315.7113.939

Campbell, F., Tricco, A. C., Munn, Z., Pollock, D., Saran, A., Sutton, A., White, H., & Khalil, H. (2023). Mapping reviews, scoping reviews, and evidence and gap maps (EGMs): the same but different— the "Big Picture" review family. *Systematic Reviews*, *12*(1). https://doi.org/10.1186/s13643-023-02178-5

Canady, M., James, B., & Nease, J. (2012). *To Protect & Educate: The School Resource Officer and the Prevention of Violence in Schools*. www.nasro.org

Cane, P. (2016). Role Responsibility. *The Journal of Ethics*, *20*(1–3), 279–298. https://doi.org/10.1007/s10892-016-9235-8

Castel, R. (1991). From Dangerousness to Risk. In G. Burchell, C. Gordon, & P. Miller (Eds.), *The Foucault Effect: Studies in Governmentality* (pp. 281–298). The University of Chicago.

CDT. (2022). *Survey Research on Parent, Student, and Teacher Experiences*.

Cefai, C., Simones, C., & Caravita, S. (2021). *A systemic, whole-school approach to mental health and well-being in schools in the EU*. https://doi.org/10.2766/50546

Chametzky, B. (2023). Writing Memos: A Vital Classic Grounded Theory Task. *European Journal of Humanities and Social Sciences*, *3*(1), 39–43. https://doi.org/10.24018/ejsocial.2023.3.1.377

Chan, C. K. Y. (2025). AI as the Therapist: Student Insights on the Challenges of Using Generative AI for School Mental Health Frameworks. *Behavioral Sciences*, *15*(3), 287. https://doi.org/10.3390/bs15030287

Chan, M. K. Y., Bhatti, H., Meader, N., Stockton, S., Evans, J., O'Connor, R. C., Kapur, N., & Kendall, T. (2016). Predicting suicide following self-harm: systematic review of risk factors and risk scales. *British Journal of Psychiatry*, *209*(4), 277–283. https://doi.org/10.1192/bjp.bp.115.170050

Charmaz, K. (2017). Constructivist grounded theory. *The Journal of Positive Psychology*, *12*(3), 299–300. https://doi.org/10.1080/17439760.2016.1262612

Charmaz, Kathy. (2006). *Constructing grounded theory*. Sage Publications.

Chin, M. (2022, August 22). The Trevor Project cuts ties with student surveillance software after online backlash. *The Verge*. https://www.theverge.com/2022/10/3/23385024/trevor-project-gaggle-lgbtq-youth-student-surveillance-software-privacy

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A Global Measure of Perceived Stress. *Journal of Health and Social Behavior*, *24*(4), 385. https://doi.org/10.2307/2136404

Collins, S., Reddy, A., Sharifi, Y., & Vance, A. (2021). *The Privacy and Equity Implications of Using Self-Harm Monitoring Technologies: Recommendations for Schools*. Student Privacy Compass. https://studentprivacycompass.org/resource/self-harm-monitoring/

Cookson, R. (2015). Justice and the NICE approach. *Journal of Medical Ethics*, *41*(1), 99–102. https://doi.org/10.1136/medethics-2014-102386

Coote, L., Kelly, L., Graham, C., Curtis-Gretton, L., Green, M., Salhi, L., de Ossorno Garcia, S., Sefi, A., & Holmes, H. (2024). An early economic evaluation of Kooth, a web-based mental health platform for children and young people with emerging mental health needs. *Internet Interventions*, *36*, 100748. https://doi.org/10.1016/j.invent.2024.100748

Corbin, J., & Strauss, A. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed). Sage Publications, Inc. https://doi.org/10.4135/9781452230153

Crown Prosecution Service. (2021). *The Code for Crown Prosecutors*. Crown Prosecution Service.

CSCI. (2005). *Safeguarding children: the second joint Chief Inspector's Report on arrangements to safeguard children*. Commission for Social Care Inspection.

Dabrowski, A., Hsien, M., Van Der Zant, T., & Ahmed, S. K. (2025). "We are left to fend for ourselves": understanding why teachers struggle to support students' mental health. *Frontiers in Education*, *9*. https://doi.org/10.3389/feduc.2024.1505077

Dai, C.-P., & Ke, F. (2022). Educational applications of artificial intelligence in simulation-based learning: A systematic mapping review. *Computers and Education: Artificial Intelligence*, *3*, 100087. https://doi.org/10.1016/j.caeai.2022.100087

Department for Business & Trade, & Department for Science, I. & T. (2024). *Directory of UK Safety Tech Providers (updated March 2024)*. https://www.gov.uk/government/publications/directory-of-uk-safety-tech-providers/directory-of-uk-safety-tech-providers

Department for Education. (2011). *Teachers' Standards Guidance for school leaders, school staff and governing bodies*. Department for Education. https://assets.publishing.service.gov.uk/media/61b73d6c8fa8f50384489c9a/Teachers__Standards_Dec_2021.pdf

Department for Education. (2016). *Behaviour and Discipline in Schools: Advice for head teachers and school staff.* Department for Education. https://dera.ioe.ac.uk/id/eprint/25117/

Department for Education. (2018). *Mental health and behaviour in schools*. Department for Education. https://assets.publishing.service.gov.uk/media/625ee6148fa8f54a8bb65ba9/Mental_health_and_behaviour_in_schools.pdf

Department for Education. (2020). *Keeping Children Safe in Education* (Issue September). Department for Education. https://consult.education.gov.uk/safeguarding-in-schools-team/keeping-children-safe-in-education-2020/supporting_documents/KCSIE%202020%20%20draft%20guidance.pdf

Department for Education. (2022). *EdTech demonstrator schools and colleges: about the programme*. https://www.gov.uk/Government/Publications/Edtech-Demonstrator-Schools-and-Colleges-Successful-Applicants/about-the-Programme.

Department for Education. (2023). *The PREVENT duty: an introduction for those with safeguarding responsibilities.* Department for Education. https://www.gov.uk/government/publications/the-prevent-duty-safeguarding-learners-vulnerable-to-radicalisation/the-prevent-duty-an-introduction-for-those-with-safeguarding-responsibilities

Department for Education. (2024). *Keeping children safe in education 2024 Statutory guidance for schools and colleges*. Department for Education. https://assets.publishing.service.gov.uk/media/66d7301b9084b18b95709f75/Keeping_children_safe_in_education_2024.pdf

Department for Education. (2025, July 15). *Publication of the revised Relationships Education, Relationships and Sex Education (RSE) and Health Education statutory guidance and the Government's response to the public consultation*. UK Parliament Committees. https://committees.parliament.uk/publications/48969/documents/257319/default/

Department of Health & Social Care, Royal College of Psychiatrists, Royal College of General Practitioners, Royal College of Nursing, The Royal College of Midwives, Institute of Health Visiting, Directors of Adult Social Services (ADASS), The British Association of Social Workers, The British Psychological Society, & Mental Health Network NHS Confederation. (2021). *Information sharing and suicide prevention: consensus statement*. https://www.gov.uk/government/publications/consensus-statement-for-information-sharing-and-suicide-prevention/information-sharing-and-suicide-prevention-consensus-statement#consensus-statement

Department of Health and Social Care. (2023). *Suicide prevention in England: 5-year cross-sector strategy*. Department of Health and Social Care. https://www.gov.uk/government/publications/suicide-prevention-strategy-for-england-2023-to-2028/suicide-prevention-in-england-5-year-cross-sector-strategy

Desai-Hunt, K. (2021, March 14). Gaggle: MPS's new student surveillance software brings possible protection and danger. *The Southener*.

Devilly, G. J., Wright, R., & Varker, T. (2009). Vicarious Trauma, Secondary Traumatic Stress or Simply Burnout? Effect of Trauma Therapy on Mental Health Professionals. *Australian & New Zealand Journal of Psychiatry*, *43*(4), 373–385. https://doi.org/10.1080/00048670902721079

D'Hotman, D., & Loh, E. (2020). AI enabled suicide prediction tools: a qualitative narrative review. *BMJ Health & Care Informatics*, *27*(3), e100175. https://doi.org/10.1136/bmjhci-2020-100175

D'Hotman, D., Loh, E., & Savulescu, J. (2020). AI-enabled suicide prediction tools: ethical considerations for medical leaders. *BMJ Leader*, leader-2020-000275. https://doi.org/10.1136/leader-2020-000275

Digby, J. (1989). *Operations Research and Systems Analysis at RAND, 1948-1967*. RAND Corporation. https://www.rand.org/content/dam/rand/pubs/notes/2007/N2936.pdf

Digital Futures Commission. (2023). *A Blueprint for Education Data*. 5Rights Foundation. https://eprints.lse.ac.uk/119737/1/A_Blueprint_for_Education_Data_FINAL_Online.pdf

Donkersley, M. (2016, October 11). *Examination of witnesses Mark Donkersley, Managing Director, e-Safe Systems Limited, and Professor Derek McAuley, Professor of Digital Economy, University of Nottingham.* https://committees.parliament.uk/oralevidence/6076/html/

Dunn, M., Sheehan, M., Hope, T., & Parker, M. (2012). Toward Methodological Innovation in Empirical Ethics Research. *Cambridge Quarterly of Healthcare Ethics*, *21*(4), 466–480. https://doi.org/10.1017/S0963180112000242

EduGeek. (2019). *Cost of Impero Education Pro* [Online forum thread]. EduGeek. Retrieved November 27, 2025, from https://www.edugeek.net/forums/topic/183740-cost-of-impero-education-pro/

Efthimiou, O., Seo, M., Chalkou, K., Debray, T., Egger, M., & Salanti, G. (2024). Developing clinical prediction models: a step-by-step guide. *BMJ*, e078276. https://doi.org/10.1136/bmj-2023-078276

Ellins, J., Hocking, L., Al-Haboubi, M., Newbould, J., Fenton, S. J., Daniel, K., … Mays, N. (2023). Implementing mental health support teams in schools and colleges: the perspectives of programme implementers and service providers. *Journal of Mental Health*, *33*(6), 714–720. https://doi.org/10.1080/09638237.2023.2278101

Faulkner, S. L., & Trotter, S. P. (2017). Data Saturation. In *The International Encyclopaedia of Communication Research Methods* (pp. 1–2). Wiley. https://doi.org/10.1002/9781118901731.iecrm0060

Fischman, W., DiBara, J. A., & Gardner, H. (2006). Creating good education against the odds. *Cambridge Journal of Education*, *36*(3), 383–398. https://doi.org/10.1080/03057640600866007

Floridi, L., & Cowls, J. (2019). A Unified Framework of Five Principles for AI in Society. *Harvard Data Science Review*, *1*(1). https://doi.org/10.1162/99608f92.8cd550d1

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and Machines*, *28*(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Foster, D. (2025). *Duties to report child abuse in England*. UK Parliament. https://commonslibrary.parliament.uk/research-briefings/sn06793/

Foucault, M. (1980). *Power/Knowledge*. Harvester.

Francis, J. J., Johnston, M., Robertson, C., Glidewell, L., Entwistle, V., Eccles, M. P., & Grimshaw, J. M. (2010). What is an adequate sample size? Operationalising data saturation for theory-based interview studies. *Psychology & Health*, *25*(10), 1229–1245. https://doi.org/10.1080/08870440903194015

Frost, N., & Robinson, M. (2007). Joining Up Children's Services: Safeguarding Children in Multi-disciplinary Teams "Effective strategies for making multi-disciplinary teams work will combine inter-agency issues with internal team-specific aspects." *Child Abuse Review*, *16*, 184–199. https://doi.org/10.1002/car

Gaggle. (2025). *Gaggle Therapy*. gaggle.net. https://www.gaggle.net/therapy

Galetsi, P., Katsaliaki, K., & Kumar, S. (2023). Exploring benefits and ethical challenges in the rise of mHealth (mobile healthcare) technology for the common good: An analysis of mobile applications for health specialists. *Technovation*, *121*. https://doi.org/10.1016/j.technovation.2022.102598

Geertz, C. (1973). Thick Description: Toward an Interpretive Theory of Culture. In *The Interpretation of Cultures*. Basic Books.

Glaser, B. G., & Strauss, A. L. (2017). *The discovery of grounded theory : strategies for qualitative research*. Routledge.

Gomes de Andrade, N. N., Pawson, D., Muriello, D., Donahue, L., & Guadagno, J. (2018). Ethics and Artificial Intelligence: Suicide Prevention on Facebook. *Philosophy and Technology*, *31*(4), 669–684. https://doi.org/10.1007/s13347-018-0336-0

Gómez-Vírseda, C., de Maeseneer, Y. & Gastmans, C. Relational autonomy: what does it mean and how is it used in end-of-life care? A systematic review of argument-based ethics literature. *BMC Med Ethics* **20**, 76 (2019). https://doi.org/10.1186/s12910-019-0417-3

Gorwa, R., Binns, R., & Katzenbach, C. (2020). Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society*, *7*(1), 205395171989794. https://doi.org/10.1177/2053951719897945

Gough, D., Thomas, J., & Oliver, S. (2012). Clarifying differences between review designs and methods. *Systematic Reviews*, *1*(1), 28. https://doi.org/10.1186/2046-4053-1-28

Graham, S., Depp, C., Lee, E. E., Nebeker, C., Tu, X., Kim, H.-C., & Jeste, D. V. (2019). Artificial Intelligence for Mental Health and Mental Illnesses: an Overview. *Current Psychiatry Reports*, *21*(11), 116. https://doi.org/10.1007/s11920-019-1094-0

Grandclerc, S., De Labrouhe, D., Spodenkiewicz, M., Lachal, J., & Moro, M.-R. (2016). Relations between Nonsuicidal Self-Injury and Suicidal Behavior in Adolescence: A Systematic Review. *PLOS ONE*, *11*(4), e0153760. https://doi.org/10.1371/journal.pone.0153760

Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. In *Health Information and Libraries Journal* (Vol. 26, Issue 2, pp. 91–108). https://doi.org/10.1111/j.1471-1842.2009.00848.x

Green, B. (2021). The Contestation of Tech Ethics: A Sociotechnical Approach to Technology Ethics in Practice. *Journal of Social Computing*, *2*(3), 209–225. https://doi.org/10.23919/JSC.2021.0018

Green, J., & Thorogood, N. (2010). *Qualitative methods for health research*. SAGE.

Guest, G., Bunce, A., & Johnson, L. (2006). How Many Interviews Are Enough? *Field Methods*, *18*(1), 59–82. https://doi.org/10.1177/1525822X05279903

Gunawardena, H., Leontini, R., Nair, S., Cross, S., & Hickie, I. (2024). Teachers as first responders: classroom experiences and mental health training needs of Australian schoolteachers. *BMC Public Health*, *24*(1), 268. https://doi.org/10.1186/s12889-023-17599-z

Haas, A. P., Eliason, M., Mays, V. M., Mathy, R. M., Cochran, S. D., D'Augelli, A. R., Silverman, M. M., Fisher, P. W., Hughes, T., Rosario, M., Russell, S. T., Malley, E., Reed, J., Litts, D. A., Haller, E., Sell, R. L., Remafedi, G., Bradford, J., Beautrais, A. L., … Clayton, P. J. (2010). Suicide and Suicide Risk in Lesbian, Gay, Bisexual, and Transgender Populations: Review and Recommendations. *Journal of Homosexuality*, *58*(1), 10–51. https://doi.org/10.1080/00918369.2011.534038

Halsband, A., & Heinrichs, B. (2022). AI, Suicide Prevention and the Limits of Beneficence. *Philosophy & Technology*, *35*(4), 103. https://doi.org/10.1007/s13347-022-00599-z

Hanson, A. (2017). Negative Case Analysis. In *The International Encyclopedia of Communication Research Methods* (pp. 1–2). Wiley. https://doi.org/10.1002/9781118901731.iecrm0165

Haque, M. D. R., & Rubya, S. (2023). An Overview of Chatbot-Based Mobile Mental Health Apps: Insights From App Description and User Reviews. *JMIR MHealth and UHealth*, *11*, e44838. https://doi.org/10.2196/44838

Hardimon, M. O. (1994). Role Obligations. *The Journal of Philosophy*, *91*(7), 333. https://doi.org/10.2307/2940934

Harmer, B., Lee, S., Rizvi, A., & Saadabadi, A. (2024). Suicidal Ideation. In *StatPearls*. StatPearls Publishing.

Hart, H. L. A. (1968). *Punishment and Responsibility*. Oxford University Press.

Hawton, K., Lascelles, K., Pitman, A., Gilbert, S., & Silverman, M. (2022). Assessment of suicide risk in mental health practice: shifting from prediction to therapeutic assessment, formulation, and risk management. *The Lancet Psychiatry*, *9*(11), 922–928. https://doi.org/10.1016/S2215-0366(22)00232-2

Hellström, T. (2013). On the moral responsibility of military robots. *Ethics and Information Technology*, *15*(2), 99–107. https://doi.org/10.1007/s10676-012-9301-2

Henson, P., Wisniewski, H., Hollis, C., Keshavan, M., & Torous, J. (2019). Digital mental health apps and the therapeutic alliance: initial review. *BJPsych Open*, *5*(1), e15. https://doi.org/10.1192/bjo.2018.86

HM Government. (2023). *Working Together to Safeguard Children: A guide to multi-agency working to help, protect and promote the welfare of children*. https://assets.publishing.service.gov.uk/media/6849a7b67cba25f610c7db3f/Working_together_to_safeguard_children_2023_-_statutory_guidance.pdf

Home Office. (2024a). *Individuals referred to and supported through the Prevent Programme, April 2023 to March 2024*. https://www.Gov.uk/Government/Statistics/Individuals-Referred-to-Prevent-to-March-2024/Individuals-Referred-to-and-Supported-through-the-Prevent-Programme-April-2023-to-March-2024.

Home Office. (2024b). *Revised Prevent duty guidance: for England and Wales (2015)*. https://www.gov.uk/Government/Publications/Prevent-Duty-Guidance-England-Scotland-and-Wales-2015/Revised-Prevent-Duty-Guidance-for-England-and-Wales-2015.

Honig, M. (2006). Complexity and policy implementation: Challenges and opportunities for the field. In M. Honig (Ed.), *New directions in education policy implementation: Confronting complexity* . The State University of New York.

Hooper, S. (2015, October 4). UK: Keyword warning software in schools raises red flag. *AlJazeera*. https://www.aljazeera.com/features/2015/10/4/uk-keyword-warning-software-in-schools-raises-red-flag

Hope, A. (2019). Creep: The Growing Surveillance of Students' Online Activities. *Education and Society*, *36*(1), 55–72. https://doi.org/10.7459/es/36.1.05

Horowitz, L. M., Ballard, E. D., & Pao, M. (2009). Suicide screening in schools, primary care and emergency departments. *Current Opinion in Pediatrics*, *21*(5), 620–627. https://doi.org/10.1097/MOP.0b013e3283307a89

Hosseinabadi-Farahani, M., Fallahi-Khoshknab, M., Arsalani, N., Hosseini, M., & Mohammadi, E. (2021). Justice and unintentional discrimination in health care. *Journal of Education and Health Promotion*, *10*(1), 51. https://doi.org/10.4103/jehp.jehp_885_20

Hughes, J. L., Horowitz, L. M., Ackerman, J. P., Adrian, M. C., Campo, J. V, & Bridge, J. A. (2023). Suicide in young people: screening, risk assessment, and intervention. *BMJ*, e070630. https://doi.org/10.1136/bmj-2022-070630

Hulme, R., Cracknell, D., & Owens, A. (2009). Learning in third spaces: developing trans-professional understanding through practitioner enquiry. *Educational Action Research*, *17*(4), 537–550. https://doi.org/10.1080/09650790903309391

Hunt, M. R., & Carnevale, F. A. (2011). Moral experience: a framework for bioethics research. *Journal of Medical Ethics*, *37*(11), 658–662. https://doi.org/10.1136/jme.2010.039008

IBS. (2025). *Smoothwall Monitor*. https://ibsschools.Com/Products-Services/Safeguarding/Smoothwall.

Ice, C. L., Neal, W. A., & Cottrell, L. (2014). Parental Efficacy and Role Responsibility for Assisting in Child's Healthful Behaviors. *Education and Urban Society*, *46*(6), 699–715. https://doi.org/10.1177/0013124512468004

IFF Research. (2023). *2022-23 Technology in Schools Survey*. https://assets.publishing.service.gov.uk/media/655f8b823d7741000d420114/Technology_in_schools_survey__2022_to_2023.pdf

Ilgen, M. A., Downing, K., Zivin, K., Hoggatt, K. J., Kim, H. M., Ganoczy, D., Austin, K. L., McCarthy, J. F., Patel, J. M., & Valenstein, M. (2009). Exploratory Data Mining Analysis Identifying Subgroups of Patients With Depression Who Are at High Risk for Suicide. *The Journal of Clinical Psychiatry*, *70*(11), 1495–1500. https://doi.org/10.4088/JCP.08m04795

Impero. (2017a). *Impero Edlink: mobile device management for education made simple*. https://www.pugh.co.uk/wp-content/uploads/2017/10/Impero-EdLink.pdf

Impero. (2017b). *Impero Education Pro*. https://schoolict.org/Wp-Content/Uploads/2016/10/Impero-Education-Pro-Brochure-2017.Pdf.

Impero. (2018a). *Impero Edaware: Supporting effective safeguarding and child protection practices*. https://www.pugh.co.uk/wp-content/uploads/2019/04/Impero_EdAware_Temp_01-4pp-UK.pdf

Impero. (2018b). *Impero Software Announces a New Version of Education Pro and Updates to iSafeguard at the 2018 Bett Show.* Impero Website. https://www.imperosoftware.com/us/resources/press-releases/impero-software-announces-new-version-education-pro-updates-isafeguard-2018-bett-show/

Impero. (2024). *Impero Software Solutions*. https://www.imperosoftware.com/

Insel, T. R. (2017). Digital phenotyping: Technology for a new science of behavior. *JAMA - Journal of the American Medical Association*, *318*(13), 1215–1216. https://doi.org/10.1001/jama.2017.11295

Institute of Medicine (US) Committee on Pathophysiology and Prevention of Adolescent and Adult Suicide. (2001). *Suicide prevention and intervention: Summary of a workshop*. National Academies Press. https://www.ncbi.nlm.nih.gov/books/NBK223843/

Ives, J. (2008). 'Encounters with Experience': Empirical Bioethics and the Future. *Health Care Analysis*, *16*(1), 1–6. https://doi.org/10.1007/s10728-007-0077-1

Ives, J. (2014). A method of Reflexive Balancing in a Pragmatic, Interdisciplinary and Reflexive Bioethics. *Bioethics*, *28*(6), 302–312. https://doi.org/10.1111/bioe.12018

Ives, J., Dunn, M., & Cribb, A. (2017). *Empirical Bioethics: Theoretical and Practical Perspectives.* (J. C. S. Ives, M. Dunn, & A. Cribb, Eds.). Cambridge University Press.

IWF. (2014, February 17). *Securus Software joins the Internet Watch Foundation to bring cutting-edge safeguarding data*. https://www.iwf.org.uk/news-media/news/securus-software-joins-the-internet-watch-foundation-to-bring-cutting-edge-safeguarding-data/

Jackman, P. C., Sanderson, R., Allen-Collinson, J., & Jacobs, L. (2022). 'There's only so much an individual can do': an ecological systems perspective on mental health and wellbeing in the early stages of doctoral research. *Journal of Further and Higher Education*, *46*(7), 931–946. https://doi.org/10.1080/0309877X.2021.2023732

Jacobson, N. C., Weingarden, H., & Wilhelm, S. (2019). Using Digital Phenotyping to Accurately Detect Depression Severity. *Journal of Nervous & Mental Disease*, *207*(10), 893–896. https://doi.org/10.1097/NMD.0000000000001042

Jennings, S. (2021). *The development, transmission and enactment of policy messages in complex adaptive educational systems: Exploring implementation and pupil experiences of mental health and wellbeing interventions in Welsh primary schools*. University of Bristol.

Kaplan, B. (2016). How Should Health Data Be Used? *Cambridge Quarterly of Healthcare Ethics*, *25*(2), 312–329. https://doi.org/10.1017/S0963180115000614

Kapur, N., Cooper, J., O'Connor, R. C., & Hawton, K. (2013). Non-suicidal self-injury *v.* attempted suicide: new diagnosis or false dichotomy? *British Journal of Psychiatry*, *202*(5), 326–328. https://doi.org/10.1192/bjp.bp.112.116111

Kass, N. E. (2001). An Ethics Framework for Public Health. *American Journal of Public Health*, *91*(11), 1776–1782. https://doi.org/10.2105/AJPH.91.11.1776

Kaste, M. (2018, November 17). Facebook Increasingly Reliant on A.I. To Predict Suicide Risk. *NPR*. https://www.npr.org/2018/11/17/668408122/facebook-increasingly-reliant-on-a-i-to-predict-suicide-risk

Keierleber, M. (2022, June 29). Minneapolis schools to halt controversial student surveillance initiative. *Minnesota Reformer*. https://minnesotareformer.com/2022/06/29/minneapolis-schools-to-halt-controversial-student-surveillance-initiative/

Kenny, M. C. (2001). Child abuse reporting: teachers' perceived deterrents. *Child Abuse & Neglect*, *25*(1), 81–92. https://doi.org/10.1016/S0145-2134(00)00218-0

Kessler, R. C., Bossarte, R. M., Luedtke, A., Zaslavsky, A. M., & Zubizarreta, J. R. (2020). Suicide prediction models: a critical review of recent research with recommendations for the way forward. *Molecular Psychiatry*, *25*(1), 168–179. https://doi.org/10.1038/s41380-019-0531-0

Khawaja, Z., & Bélisle-Pipon, J.-C. (2023). Your robot therapist is not your therapist: understanding the role of AI-powered mental health chatbots. *Frontiers in Digital Health*, *5*. https://doi.org/10.3389/fdgth.2023.1278186

Khurana, D., Koli, A., Khatter, K., & Singh, S. (2023). Natural language processing: state of the art, current trends and challenges. *Multimedia Tools and Applications*, *82*(3), 3713–3744. https://doi.org/10.1007/s11042-022-13428-4

Kingori, P. (2013). Experiencing everyday ethics in context: Frontline data collectors perspectives and practices of bioethics. *Social Science & Medicine*, *98*, 361–370. https://doi.org/10.1016/j.socscimed.2013.10.013

Kon, A. A., Davidson, J. E., Morrison, W., Danis, M., & White, D. B. (2016). Shared decision making in ICUs: An American college of critical care medicine and American thoracic society policy statement. *Critical Care Medicine*, *44*(1), 188–201. https://doi.org/10.1097/CCM.0000000000001396

Koops, B.-J. (2021). The concept of function creep. *Law, Innovation and Technology*, *13*(1), 29–56. https://doi.org/10.1080/17579961.2021.1898299

Kretzschmar, K., Tyroll, H., Pavarini, G., Manzini, A., & Singh, I. (2019). Can Your Phone Be Your Therapist? Young People's Ethical Perspectives on the Use of Fully Automated Conversational Agents (Chatbots) in Mental Health Support. *Biomedical Informatics Insights*, *11*. https://doi.org/10.1177/1178222619829083

Laird, E., Grant-Chapman, H., Venzke, C., & Quay-De La Vallee, H. (2022). *Hidden Harms: The Misleading Promise of Monitoring Students Online*. https://cdt.org/wp-content/uploads/2022/08/Hidden-Harms-The-Misleading-Promise-of-Monitoring-Students-Online-Research-Report-Final-Accessible.pdf

Large, M., Kaneson, M., Myles, N., Myles, H., Gunaratne, P., & Ryan, C. (2016). Meta-Analysis of Longitudinal Cohort Studies of Suicide Risk Assessment among Psychiatric Patients: Heterogeneity in Results and Lack of Improvement over Time. *PLOS ONE*, *11*(6), e0156322. https://doi.org/10.1371/journal.pone.0156322

Large, M. M. (2018). The role of prediction in suicide prevention. *Dialogues in Clinical Neuroscience*, *20*(3), 197–205. https://www.ncbi.nlm.nih.gov/pubmed/30581289

Lee, V. E., & Loeb, S. (2000). School Size in Chicago Elementary Schools: Effects on Teachers' Attitudes and Students' Achievement. *American Educational Research Journal*, *37*(1), 3–31. https://doi.org/10.3102/00028312037001003

Levinson, A. H., Crepeau-Hobson, M. F., Coors, M. E., Glover, J. J., Goldberg, D. S., & Wynia, M. K. (2020). Duties When an Anonymous Student Health Survey Finds a Hot Spot of Suicidality. *The American Journal of Bioethics*, *20*(10), 50–60. https://doi.org/10.1080/15265161.2020.1806374

LightSpeed Systems. (2025a). *Compliance*. https://www.lightspeedsystems.com/en_uk/challenges/compliance/

LightSpeed Systems. (2025b). *Safety and Student Services*. https://www.lightspeedsystems.com/persona/safety-and-student-services/

Lim, W. M. (2024). What Is Qualitative Research? An Overview and Guidelines. *Australasian Marketing Journal*. https://doi.org/10.1177/14413582241264619

Liverpool Safeguarding Children Partnership. (2024). *Liverpool Safeguarding Children Partnership Procedures*. Https://Liverpoolscp.Trixonline.Co.Uk/Chapter/Multi-Agency-Working#principles-of-Multi-Agency-Working.

Livingstone, S., Pothong, K., Atabey, A., Hooper, L., & Day, E. (2024). The Googlization of the classroom: Is the UK effective in protecting children's data and rights? *Computers and Education Open*, *7*, 100195. https://doi.org/10.1016/j.caeo.2024.100195

Local Government Association. (2023). *Children and young people's emotional wellbeing and mental health – facts and figures*. https://www.local.gov.uk/about/campaigns/bright-futures/bright-futures-camhs/child-and-adolescent-mental-health-and

Lorenz, K., Freddolino, P. P., Comas-Herrera, A., Knapp, M., & Damant, J. (2019). Technology-based tools and services for people with dementia and carers: Mapping technology onto the dementia care pathway. *Dementia*, *18*(2), 725–741. https://doi.org/10.1177/1471301217691617

Lunenberg, M., Dengerink, J., & Korthagen, F. (2014). *The Professional Teacher Educator*. Sense Publishers. https://doi.org/10.1007/978-94-6209-518-2

Maass, D., Barnett, D., & Kelley, J. (2023). *GoGuardian: A Red Flag Machine By Design*. www.redflagmachine.org

Madhusudan, B. (2021). *Securly response to inquiry, letter to Elizabeth Warren, Edward J Markey, and Richard Blumenthal. https://www.warren.senate.gov/imo/media/doc/Securly%20Senate%20Response%20Final.pdf*

Mann, J. J., Michel, C. A., & Auerbach, R. P. (2021). Improving Suicide Prevention Through Evidence-Based Strategies: A Systematic Review. *American Journal of Psychiatry*, *178*(7), 611–624. https://doi.org/10.1176/appi.ajp.2020.20060864

Manzini, A. (2020). *Citizenship, Genomics, and Mental Health: An Empirical Bioethics Study of Young People's Attitudes towards Advances in Autism Genomics*. Retrieved August 9, 2025, from https://ora.ox.ac.uk/objects/uuid:fa01f84d-23be-4569-a0f6-09ef949c1a89

Marcelle, E. T., Nolting, L., Hinshaw, S. P., & Aguilera, A. (2019). Effectiveness of a Multimodal Digital Psychotherapy Platform for Adult Depression: A Naturalistic Feasibility Study. *JMIR MHealth and UHealth*, *7*(1), e10948. https://doi.org/10.2196/10948

Marks, M. (2019). Artificial Intelligence Based Suicide Prediction. *Yale Journal of Health Policy, Law and Ethics*, 98–121.

Marlow, C. (2025, May 8). Is your school spying on your child online? *The Guardian*. https://www.theguardian.com/commentisfree/2025/may/08/surveillance-schools-students-edtech

Maslaha. (2023). *How PREVENT Impacts Your Students*. Maslaha.Org. https://www.maslaha.org/project/how-to-make-a-safe-classroom-for-all-children

Matthias, A. (2004). The responsibility gap: Ascribing responsibility for the actions of learning automata. *Ethics and Information Technology*, *6*(3), 175–183. https://doi.org/10.1007/s10676-004-3422-1

McCosker, A., Farmer, J., & Kamstra, P. (2023, October). Mental Health and the Digital Care Assemblage: User and Moderator Practices. *AoIR Selected Papers of Internet Research.* https://doi.org/10.5210/spir.v2023i0.13461

McKernan, L. C., Clayton, E. W., & Walsh, C. G. (2018). Protecting Life While Preserving Liberty: Ethical Recommendations for Suicide Prevention With Artificial Intelligence. *Frontiers in Psychiatry*, *9*, 650. https://doi.org/10.3389/fpsyt.2018.00650

McMillan, J., & Hope, T. (2008). The possibility of empirical psychiatric ethics. In *Empirical ethics in psychiatry*. Oxford University Press.

Meadows, D. H., & Wright, Diana. (2015). *Thinking in systems : a primer*. Chelsea Green Publishing.

Meier, E., Rigter, T., Schijven, M. P., van den Hoven, M., & Bak, M. A. R. (2025a). Correction: The impact of digital health technologies on moral responsibility: a scoping review. *Medicine, Health Care and Philosophy*, *28*(2), 367–368. https://doi.org/10.1007/s11019-024-10248-1

Meier, E., Rigter, T., Schijven, M. P., van den Hoven, M., & Bak, M. A. R. (2025b). The impact of digital health technologies on moral responsibility: a scoping review. *Medicine, Health Care and Philosophy*, *28*(1), 17–31. https://doi.org/10.1007/s11019-024-10238-3

Meredith, J., McCarthy, S., & Hemsley, B. (2018). Legal and ethical issues surrounding the use of older children's electronic personal health records. *Journal of Law and Medicine*, *25*, 1042–1055.

Milne, D. N., McCabe, K. L., & Calvo, R. A. (2019). Improving Moderator Responsiveness in Online Peer Support Through Automated Triage. *Journal of Medical Internet Research*, *21*(4), e11410. https://doi.org/10.2196/11410

Milton, C. L. (2019). Privacy: Potential Violations of Human Dignity. *Nursing Science Quarterly*, *32*(2), 106–107. https://doi.org/10.1177/0894318419826215

Miser, H. J., & Quade, E. S. (1985). *Handbook of Systems Analysis, Part I. Overview of Uses, Procedures, Applications, and Practice.* . Elsevier Science Publishing.

Moore, G., Michie, S., Anderson, J., Belesova, K., Crane, M., Deloly, C., Dimitroulopoulou, S., Gitau, H., Hale, J., Lloyd, S. J., Mberu, B., Muindi, K., Niu, Y., Pineo, H., Pluchinotta, I., Prasad, A., Roue-Le Gall, A., Shrubsole, C., Turcu, C., … Osrin, D. (2021). Developing a programme theory for a transdisciplinary research collaboration: Complex Urban Systems for Sustainability and Health. *Wellcome Open Research*, *6*, 35. https://doi.org/10.12688/wellcomeopenres.16542.2

Morley, J., & Floridi, L. (2024). *The Ethics of AI in Health Care: An Updated Mapping Review*. https://doi.org/10.2139/ssrn.4987317

Morse, J. M. (2015). "Data Were Saturated . . . ." *Qualitative Health Research*, *25*(5), 587–588. https://doi.org/10.1177/1049732315576699

Moura, I., Teles, A., Silva, F., Viana, D., Coutinho, L., Barros, F., & Endler, M. (2020). Mental health ubiquitous monitoring supported by social situation awareness: A systematic review. *Journal of Biomedical Informatics*, *107*, 103454. https://doi.org/10.1016/j.jbi.2020.103454

Nadeem, E., Kataoka, S. H., Chang, V. Y., Vona, P., Wong, M., & Stein, B. D. (2011). The Role of Teachers in School-Based Suicide Prevention: A Qualitative Study of School Staff Perspectives. *School Mental Health*, *3*(4), 209–221. https://doi.org/10.1007/s12310-011-9056-7

Nagenborg, M., Capurro, R., Weber, J., & Pingel, C. (2008). Ethical regulations on robotics in Europe. *AI & Society*, *22*(3), 349–366. https://doi.org/10.1007/s00146-007-0153-y

Nass, S. J., Levit, L. A., & Gostin, L. O. (2009). *Beyond the HIPAA privacy rule : enhancing privacy, improving health through research*. National Academies Press.

Nehme, M., Schneider, F., Amruthalingam, E., Schnarrenberger, E., Tremeaud, R., & Guessous, I. (2024). Chatbots in medicine: certification process and applied use case. *Swiss Medical Weekly*, *154*, 3954. https://doi.org/10.57187/s.3954

Nelson, H. D., Denneson, L. M., Low, A. R., Bauer, B. W., O'Neil, M., Kansagara, D., & Teo, A. R. (2017). Suicide Risk Assessment and Prevention: A Systematic Review Focusing on Veterans. *Psychiatric Services*, *68*(10), 1003–1015. https://doi.org/10.1176/appi.ps.201600384

Netsupport. (2017). *Supporting the Prevent Duty Guidance*. https://web.archive.org/Web/20220422104513/Https://Www.Netsupportsoftware.Com/Prevent-Duty-Guidance/.

NetSupportDNA. (2025). *NetSupportDNA Education*. https://www.netsupportdna.com/education/

NetSweeper. (2021, May 27). *Netsweeper Helping Trading with Schools Manage Internet Access for 51,000 Staff and Students*. https://www.netsweeper.com/education-web-filtering/netsweeper-helping-trading-with-schools-manage-internet-access-for-51000-staff-and-students?

Netsweeper. (2023, January 13). *Ways to Improve School Safety Using Student Safety Alert Systems*. https://www.netsweeper.com/education-web-filtering/school-safety-using-student-safety-alert-systems

Netsweeper. (2025a). *OnGuard: Identify Students at Risk in Real Time AI-Detected, Human-Verified Digital Safety Monitoring*. https://go.netsweeper.com/onguardUK

Netsweeper. (2025b). *Reduce Administrative Burden for School IT Staff*. https://www.netsweeper.com/uk-web-filter-digital-safety/uk-solutions/policy-compliance

NetSweeper. (2025). *Support Student Mental Health with Digital Safety Monitoring*. Https://Www.Netsweeper.Com/Solutions/Education-Student-Mental-Health-Safety#:~:Text=Strengthen%20Student%20Safety%20and%20Wellbeing&text=Using%20AI%20technology%2C%20it%20detects,Initiate%20appropriate%20intervention%2C%20when%20necessary.

NHS. (2024, July). *NHS Screening*. https://www.nhs.uk/Conditions/Nhs-Screening/.

NICE. (2022). *Self-harm: assessment, management and preventing recurrence NICE guideline*. www.nice.org.uk/guidance/ng225

NIHR. (2021, June 16). *New wave of AI technologies in £36 million funding boost*. Nihr.Ac.Uk. https://www.nihr.ac.uk/news/new-wave-ai-technologies-ps36-million-funding-boost

NIHR. (2023, March 3). *Ground-breaking AI research aims to improve tests and treatments for thousands of patients*. Nihr.Ac.Uk. https://www.nihr.ac.uk/news/ground-breaking-ai-research-aims-improve-tests-and-treatments-thousands-patients

Noble, H., & Mitchell, G. (2016). What is grounded theory? *Evidence Based Nursing*, *19*(2), 34–35. https://doi.org/10.1136/eb-2016-102306

Nollkaemper, A. (2018). The duality of shared responsibility. *Contemporary Politics*, *24*(5), 524–544. https://doi.org/10.1080/13569775.2018.1452107

Nordin, N., Zainol, Z., Mohd Noor, M. H., & Chan, L. F. (2022). Suicidal behaviour prediction models using machine learning techniques: A systematic review. *Artificial Intelligence in Medicine*, *132*, 102395. https://doi.org/10.1016/j.artmed.2022.102395

Ofsted. (2020). *Making the cut: how schools respond when they are under financial pressure*. www.gov.uk/government/publications/amanda-spielman-letter-to-the-public-accounts-committee.

ONS. (2024). *Suicides in England and Wales 2023 registrations*.

Opfer, V. D., & Pedder, D. (2011). Conceptualizing Teacher Professional Learning. *Review of Educational Research*, *81*(3), 376–407. https://doi.org/10.3102/0034654311413609

O'Reilly, M., Adams, S., Whiteman, N., Hughes, J., Reilly, P., & Dogra, N. (2018). Whose Responsibility is Adolescent's Mental Health in the UK? Perspectives of Key Stakeholders. *School Mental Health*, *10*(4), 450–461. https://doi.org/10.1007/s12310-018-9263-6

Orlando, C. M., Broman-Fulks, J. J., Whitlock, J. L., Curtin, L., & Michael, K. D. (2015). Nonsuicidal Self-Injury and Suicidal Self-Injury: A Taxometric Investigation. *Behavior Therapy*, *46*(6), 824–833. https://doi.org/10.1016/j.beth.2015.01.002

Osman, A., Bagge, C. L., Gutierrez, P. M., Konick, L. C., Kopper, B. A., & Barrios, F. X. (2001). The Suicidal Behaviors Questionnaire-Revised (SBQ-R):Validation with Clinical and Nonclinical Samples. *Assessment*, *8*(4), 443–454. https://doi.org/10.1177/107319110100800409

Panaite, A., Desroches, O., Warren, É., Rouly, G., Castonguay, G., & Boivin, A. (2024). Engaging with peers to integrate community care: Knowledge synthesis and conceptual map. *Health Expectations*, *27*(2). https://doi.org/10.1111/hex.14034

Parsons, J. A., Johal, H. K., Parker, J., & Romanis, E. C. (2024). Translational or translationable? A call for ethno-immersion in (empirical) bioethics research. *Bioethics*, *38*(3), 252–261. https://doi.org/10.1111/bioe.13184

Pathirathna, M. L., Nandasena, H. M. R. K., Atapattu, A. M. M. P., & Weerasekara, I. (2022). Impact of the COVID-19 pandemic on suicidal attempts and death rates: a systematic review. *BMC Psychiatry*, *22*(1), 506. https://doi.org/10.1186/s12888-022-04158-w

Patterson, J. (2021). *Gaggle response to inquiry, letter to Elizabeth Warren, Edward J. Markey, and Richard Blumenthal*. https://perma.cc/GN6H-WDN9

Pavarini, G., Lorimer, J., Manzini, A., Goundrey-Smith, E., & Singh, I. (2019). Co-producing Research with Youth: The NeurOx Young People's Advisory Group Model. *Health Expectations*, *22*(4), 743–751. https://doi.org/10.1111/hex.12911

Pellegrino, E. D., & Thomasma, D. C. (1981). A philosophical basis of medical practice. New York: Oxford University Press.

Persson, J. (2022, October). *Policing thoughts, proactive technology, and the Online Safety Bill*. https://jenpersson.com/Caredata/Privacy/.

Petersen, K., Vakkalanka, S., & Kuzniarz, L. (2015). Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and Software Technology*, *64*, 1–18. https://doi.org/10.1016/j.infsof.2015.03.007

Peterson, D. (2016). Edtech and Student Privacy: California Law as a Model. *Berkeley Technology Law Journal*, *31*(2), 961–996.

Petrina, S. (2006). The Medicalization of Education: A Historiographic Synthesis. *History of Education Quarterly*, *46*(4), 503–531. https://doi.org/10.1111/j.1748-5959.2006.00030.x

Pirani, S., Kulhanek, C., Wainwright, K., & Osman, A. (2021). The Reasons for Living Inventory for Young Adults (RFL-YA-II). *Assessment*, *28*(3), 942–954. https://doi.org/10.1177/1073191119900242

Plutchik, R., van Praag, H. M., Conte, H. R., & Picard, S. (1989). Correlates of suicide and violence risk 1: The suicide risk measure. *Comprehensive Psychiatry*, *30*(4), 296–302. https://doi.org/10.1016/0010-440X(89)90053-9

Posner, K., Brent, D., Lucas, C., Gould, M., Stanley, B., Brown, G., & Mann, J. (2008). *Columbia-Suicide Severity Rating Scale (C-SSRS).* https://cssrs.columbia.edu/wp-content/uploads/C-SSRS_Pediatric-SLC_11.14.16.pdf

Poulin, C., Shiner, B., Thompson, P., Vepstas, L., Young-Xu, Y., Goertzel, B., Watts, B., Flashman, L., & McAllister, T. (2014). Predicting the Risk of Suicide by Analyzing the Text of Clinical Notes. *PLoS ONE*, *9*(1), e85733. https://doi.org/10.1371/journal.pone.0085733

Prabhakar, E. (2013). E-Therapy: Ethical Considerations of a Changing Healthcare Communication Environment. *Pastoral Psychology*, *62*(2), 211–218. https://doi.org/10.1007/s11089-012-0434-3

PSE. (2019, October 19). *Cheshire Councils adopt new eSafety solution for schools*. https://www.publicsectorexecutive.com/News/cheshire-councils-adopt-new-esafety-solution-for-schools

Public Health England. (2021). *Promoting children and young people's mental health and wellbeing A whole school or college approach Public Health England working with the Department for Education*.

Radden, J. (2007). Virtue Ethics as Professional Ethics: The Case of Psychiatry. In Working Virtue (pp. 113–134). Oxford University PressOxford. https://doi.org/10.1093/oso/9780199271658.003.0005

Rahwan, I. (2018). Society-in-the-loop: programming the algorithmic social contract. *Ethics and Information Technology*, *20*(1), 5–14. https://doi.org/10.1007/s10676-017-9430-8

Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). Closing the AI accountability gap. In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (pp. 33–44). FAT* ’20: Conference on Fairness, Accountability, and Transparency. ACM. https://doi.org/10.1145/3351095.3372873

Range, L. M., & Lewis, L. S. (1992). Life Orientation Inventory: A method of assessing suicide risk. *Journal of Psychoeducational Assessment*, *10*(3), 296–297. https://doi.org/10.1177/073428299201000310

RCPCH. (2021). *Adolescent Mortality*. State of Child Health. https://stateofchildhealth.rcpch.ac.uk/evidence/mortality/adolescent-mortality/

Regan, P. M., & Jesse, J. (2019). Ethical challenges of edtech, big data and personalized learning: twenty-first century student sorting and tracking. *Ethics and Information Technology*, *21*(3), 167–179. https://doi.org/10.1007/s10676-018-9492-2

Reilly, M. (2017, May 1). Is Facebook targeting advertising at depressed teens? *Technology Review*. https://www.technologyreview.com/604307/is-facebook-targeting-ads-at-sad-teens/

Rose, N. (2010). ‘Screen and intervene’: governing risky brains. *History of the Human Sciences*, *23*(1), 79–105. https://doi.org/10.1177/0952695109352415

Roy, A., Nikolitch, K., McGinn, R., Jinah, S., Klement, W., & Kaminsky, Z. A. (2020). A machine learning approach predicts future risk to suicidal ideation from social media data. *Npj Digital Medicine*, *3*(1), 78. https://doi.org/10.1038/s41746-020-0287-6

Rule, J. B. (1974). *Private Lives and Public Surveillance: Social Control in the Computer Age*. Schocken books.

Runeson, B., Odeberg, J., Pettersson, A., Edbom, T., Jildevik Adamsson, I., & Waern, M. (2017). Instruments for the assessment of suicide risk: A systematic review evaluating the certainty of the evidence. *PLOS ONE*, *12*(7), e0180292. https://doi.org/10.1371/journal.pone.0180292

Ryan, P., Porter, Z., Al-Qaddoumi, J., McDermid, J., & Habli, I. (2023). *What’s my role? Modelling responsibility for AI-based safety-critical systems*. http://arxiv.org/abs/2401.09459

Salloch, S., Vollmann, J., & Schildmann, J. (2014). Ethics by opinion poll? The functions of attitudes research for normative deliberations in medical ethics: Table 1. *Journal of Medical Ethics*, *40*(9), 597–602. https://doi.org/10.1136/medethics-2012-101253

Salmon, G. (2004). Multi-Agency Collaboration: The Challenges for CAMHS. *Child and Adolescent Mental Health*, *9*(4), 156–161. https://doi.org/10.1111/j.1475-3588.2004.00099.x

Santoro, M., Marino, D., & Tamburrini, G. (2008). Learning robots interacting with humans: from epistemic risk to responsibility. *AI & Society*, *22*(3), 301–314. https://doi.org/10.1007/s00146-007-0155-9

Schafer, A. (2015). Quiet sabotage of the queer child: Why the law must be reframed to appreciate the dangers of outing gay youth. *Howard Law Journal*, *58*(2), 597–636.

Schalock, H. Del. (1998). Student progress in learning: Teacher responsibility, accountability, and reality. *Journal of Personnel Evaluation in Education*, *12*(3), 237–246. https://doi.org/10.1023/A:1008063126448

Schicktanz, S., & Schweda, M. (2012). The Diversity of Responsibility: The Value of Explication and Pluralization. *Medicine Studies*, *3*(3), 131–145. https://doi.org/10.1007/s12376-011-0070-8

Schools Broadband. (2025a). *Homepage*. https://www.schoolsbroadband.co.uk/

Schools Broadband. (2025b). *Safeguarding, Filtering, Monitoring and Security Ensuring Online Safety in Schools with Advanced Filtering and Monitoring Solutions*. https://www.schoolsbroadband.co.uk/our-services/safeguarding-filtering-and-security/

Schools Broadband. (2025c). *Safeguarding Management*. https://www.schoolsbroadband.co.uk/our-services/safeguarding-filtering-and-security/safeguarding-management/

Schuelka, M. J., & Engsig, T. T. (2022). On the question of educational purpose: complex educational systems analysis for inclusion. *International Journal of Inclusive Education*, *26*(5), 448–465. https://doi.org/10.1080/13603116.2019.1698062

Securly. (2019). *Smart DNS to Anything How Securly Works?* https://www.securly.com/assets/images/Securly_b2b_SmartDNSwhitepaper_final.pdf

Securly. (2024). *Privacy Policy*. https://www.securly.com/privacy

Securus. (2025). *Managed full monitoring service for schools & all other education settings*. https://www.securus-software.com/securus-fms/

Sensocloud. (2025a). *Homepage*. https://senso.cloud/en-GB

Sensocloud. (2025b). *Microsoft Teams Monitoring*. https://senso.cloud/en-GB/safeguarding-microsoft-teams

Seymour, K., McNicoll, J., & Koenig-Robert, R. (2024). *Big brother: The effects of surveillance on fundamental aspects of social vision*. *Neuroscience of Consciousness, 2024*(1), Article niae039. https://doi.org/10.1093/nc/niae039

Shafti, M., Taylor, P., Forrester, A., Handerer, F., & Pratt, D. (2023). A systematic review of the co-occurrence of self-harm and aggression: Is dual harm a unique behavioural construct? In *Frontiers in Psychiatry* (Vol. 14). Frontiers Media S.A. https://doi.org/10.3389/fpsyt.2023.1083271

Shelton, L. C. (2022). *Identifying Suicide and Self-harm Behaviors in Students with Gaggle Safety Management Program* (Publication Number 29165789) [Doctoral dissertation, University of Colorado at Denver]. ProQuest Dissertations & Theses. https://www.proquest.com/openview/73eb00812a44d10ae35cfd2b26df7e50/1?pq-origsite=gscholar&cbl=18750&diss=y

Shinde, A. (2021). *GoGuardian response to inquiry, letter to Elizabeth Warren, Edward J. Markey, and Richard Blumenthal*. https://www.warren.senate.gov/imo/media/doc/GoGuardian%20Response%20_%20Re_Edtech%20Letter.pdf

Silverman, David. (2017). *Doing qualitative research*. SAGE Publications.

Sims-Schouten, W. (2017). 'Mental Health First Aid Training' in schools is a sticking-plaster solution. . *The Conversation*. https://theconversation.com/mental-health-first-aid-training-in-schools-is-a-sticking-plaster-solution-80166

Smoothwall. (2023a, August 21). *1 in 10 DSLs Admit to Not Being Able to Spot a Child with Mental Health Issues, Can You?* https://smoothwall.com/resources/can-you-spot-a-child-with-mental-health-issues

Smoothwall. (2023b, September 1). *The 2023 KCSIE Filtering and Monitoring Updates Every School Needs to Know*. Smoothwall. https://smoothwall.com/resources/kcsie-2023-filtering-and-monitoring-updates-smoothwall

Smoothwall. (2025a). *Detecting Unsafe Images in Cloud Storage: The Challenge and Solution*. https://smoothwall.com/Resources/Detecting-Unsafe-Images-in-Cloud-Storage-Challenge-and-Solution.

Smoothwall. (2025b). *Qustodio Parental App*. https://smoothwall.com/solutions/qustodio

Smoothwall. (2025c). *Smoothwall Pulse*. https://smoothwall.com/solutions/pulse

Smoothwall. (2025d). *Welcome to Smoothwall*. https://www.smoothwall.com/education/

Solomon, M. Z. (2005). Realizing bioethics' goals in practice: ten ways "is" can help "ought". *The Hastings Center Report*, *35*(4), 40–47.

Spratt, J., Shucksmith, J., Philip, K., & Watson, C. (2006a). Interprofessional support of mental well-being in schools: A Bourdieuan perspective. *Journal of Interprofessional Care*, *20*(4), 391–402. https://doi.org/10.1080/13561820600845643

Spratt, J., Shucksmith, J., Philip, K., & Watson, C. (2006b). 'Part of who we are as a school should include responsibility for well-being': Links between the school environment, mental health and behaviour. *Pastoral Care in Education*, *24*(3), 14–21. https://doi.org/10.1111/j.1468-0122.2006.00374.x

Stanley, T., Guru, S., & Gupta, A. (2018). Working with PREVENT: Social Work Options for Cases of 'Radicalisation Risk.' *Practice*, *30*(2), 131–146. https://doi.org/10.1080/09503153.2017.1414176

Steinhoff, A., Bechtiger, L., Ribeaud, D., Eisner, M., & Shanahan, L. (2023). Self-, other-, and dual-harm during adolescence: a prospective-longitudinal study of childhood risk factors and early adult correlates. *Psychological Medicine*, *53*(9), 3995–4003. https://doi.org/10.1017/S0033291722000666

Stromquist, N. P. (2006). The theoretical and practical bases for empowerment. In D. Bhaskara & D. Pushpalatha (Eds.), *Women, education, and empowerment*. Discovery Publishing House.

Sueki, H. (2015). The association of suicide-related Twitter use with suicidal behaviour: A cross-sectional study of young internet users in Japan. *Journal of Affective Disorders*, *170*, 155–160. https://doi.org/10.1016/j.jad.2014.08.047

Sutton Council. (2025). *Team Around the Child (formerly known as Impact)*. Https://Www.Sutton.Gov.Uk/w/Early-Support-Service-Team-around-the-Child-Formerly-Known-as-Impact.

Taylor, E. (2013). *Surveillance Schools*. Palgrave Macmillan UK. https://doi.org/10.1057/9781137308863

Thuy Vo, L., & Aldhous, P. (2019). Your Dumb Tweets Are Getting Flagged To People Trying To Stop School Shootings. *Buzzfeed*. https://www.buzzfeednews.com/article/lamvo/social-sentinel-school-officials-shootings-flag-social-media

Tigard, D. W. (2021). There Is No Techno-Responsibility Gap. *Philosophy & Technology*, *34*(3), 589–607. https://doi.org/10.1007/s13347-020-00414-7

Turoldo, F. (2009). Responsibility as an Ethical Framework for Public Health Interventions. *American Journal of Public Health*, *99*(7), 1197–1202. https://doi.org/10.2105/AJPH.2007.127514

Ulrich, C. M., & Ratcliffe, S. J. (2007). *Hypothetical Vignettes in Empirical Bioethics Research* (pp. 161–181). https://doi.org/10.1016/S1479-3709(07)11008-6

UNESCO. (2021). *One year into COVID-19 education disruption: Where do we stand?* UNESECO Website. https://www.unesco.org/en/articles/one-year-covid-19-education-disruption-where-do-we-stand

van de Poel, I., Nihlén Fahlquist, J., Doorn, N., Zwart, S., & Royakkers, L. (2012). The Problem of Many Hands: Climate Change as an Example. *Science and Engineering Ethics*, *18*(1), 49–67. https://doi.org/10.1007/s11948-011-9276-0

Vanderstraeten, R. (2023). Systems Theory Approaches to Researching Educational Organizations. In *Oxford Research Encyclopaedia of Education*. Oxford University Press. https://doi.org/10.1093/acrefore/9780190264093.013.1885

Varkey, B. (2021). Principles of Clinical Ethics and Their Application to Practice. *Medical Principles and Practice*, *30*(1), 17–28. https://doi.org/10.1159/000509119

Velupillai Sumithra, Epstein Sophie, Bittar André, Stephenson Thomas, Dutta Rina, & Downs Johnny. (2019). *Identifying Suicidal Adolescents from Mental Health Records Using Natural Language Processing*. https://doi.org/10.3233/SHTI190254

Vembye, M., & Jensen, H. S. (2018, March 23). How the Laws of Education Lie. *Philosophy of Education Society of Great Britain Conference*.

Von Bertalanffy, L. (1968). *General System Theory: Foundations, Development.* George Braziller.

Wahlstrom, K. L., & Louis, K. S. (2008). How Teachers Experience Principal Leadership: The Roles of Professional Community, Trust, Efficacy, and Shared Responsibility. *Educational Administration Quarterly*, *44*(4), 458–495. https://doi.org/10.1177/0013161X08321502

Warren, E., & Markey, E. (2022). *Prepared by Senators Constant Surveillance: Implications of Around-the-Clock Online Student Activity Monitoring Constant Surveillance*. https://www.warren.senate.gov/download/356670-student-surveillance

Wenger, E. (1998). *Communities of Practice*. Cambridge University Press. https://doi.org/10.1017/CBO9780511803932

West, M. (2023). *An ed-tech tragedy? Educational technologies and school closures in the time of COVID-19*. UNESCO. https://doi.org/10.54675/LYGF2153

Williamson, B. (2021). Meta-edtech. *Learning, Media and Technology*, *46*(1), 1–5. https://doi.org/10.1080/17439884.2021.1876089

Williamson, B., Eynon, R., & Potter, J. (2020). Pandemic politics, pedagogies and practices: digital technologies and distance education during the coronavirus emergency. *Learning, Media and Technology*, *45*(2), 107–114. https://doi.org/10.1080/17439884.2020.1761641

World Health Organization (2021). *Live life: An implementation guide for suicide prevention in countries*. Geneva: World Health Organization.

Zempi, I., & Tripli, A. (2023). Listening to Muslim Students' Voices on the Prevent Duty in British Universities: A Qualitative Study. *Education, Citizenship and Social Justice*, *18*(2), 230–245. https://doi.org/10.1177/17461979221077990

Appendix

Appendix One: Available EdTech Software to Monitor for Suicide Risk

Name	Use	Materials Screened	Risk of…	Scanning Technique	Moderator	Interventions Suggested	Compliance	Parental Role	Teacher Role	Student Role	Approved by..?
Ativion (Prior to 2024, Ativion was referred to as Impero Ed-aware and Ed-Protect) Sources: Impero 2017; 2018; 2024; 2025	According to their website, in 2024, Impero was used in 90+ countries, and 1,400 UK high schools (Impero, 2025)	According to their website (Impero, 2025) this software screens materials on the network (e.g. anything typed on HTML, web browsers, emails, applications). This software takes screenshot captures of flagged material and 30 second recordings. This software can be downloaded on school computers and personal devices.	According to their website, in 2024 this software searched risk for (non exclusively): attendance, bereavement, bullying, anxiety, discrimination, friendship, gender/sexuality, home life, home communication, hunger, inappropriate technology, mental health, personal hygiene, preparation/attitude to learning, school refusal, sexualize behavior	The software includes a library used for identifying key words, phrases, abbreviations or acronyms	There is no mention of a human moderator within any of the promotional materials.	According to their website, in 2024 this software allowed schools to "store and access over 100 safeguarding-related documents, such as key government legislation and memos, with links to templates and national guidance located in one place. Add and log your school's documents and policies" (Impero, 2024, p.1)	According to their website, in 2024 Impero followed FERPA and COPPA guidelines. (Impero, 2024) According to their website, schools did not have to get consent from parents or students to use this software, however schools did need to give students clear notification that the school was using this system. (Impero, 2025).	According to their website, in 2024, parents can track and monitor devices, including browsing history, location, and any contact accessed. Parents can also apply screen-time limits and block/control access. (Impero, 2017).	According to their website, teachers/school staff can "add a concern…by recording, analyzing, and storing multiple types of student safety information, including depressive episodes, risky online behavior. This enables informed counselling sessions, behavior management conversations and safeguarding interventions." (Impero, 2018).	Live chat available for students	Mental Health America (MHA), the Internet Watch Foundation (IWF), Victvs, Beat, ABA and SafeBAE
Securly	According to their	According to their online materials	According to their online materials,	According to their website,	According to Securly's	There is no mention of any suggested	According to Securly's website	According to Securly's website (Securly,	Securly allows a school to create a "delegated	According to their website,	There is no mention of any approval by

Sources: Securly, 2019; 2024	website (Securly, 2019), Securly reaches 20,000+ schools	(Securly, 2019), Securly screens students' activity feed, downloads, emails, email, docs, and drive materials.	Securly monitors for students' risk of nudity, bullying, "dark content," self-harm, and violence (Securly, 2019)	software uses smart DNS and NLP to scan students' computers based on context and key words. Includes "risk confidence score." (Securly, 2019)	website (Securly, 2019), they have specialists available 24/7 to analyze flagged activity. These specialists decide what needs immediate action, and then contact the school if needed.	interventions in this company's online materials.	(Securly, 2024), identifiable information can be shared with: service providers, in business transfers, and in response to legal processes. According to their website, students' aggregate/de-identifiable information is allowed to be used for third-party online behavioral advertising and research (Securly, 2024).	2024), the school is able to choose what they share with parents. This includes whether they display all activity, only home activity, or educational sites visited. Parents do, however, have access to a parent portal. This portal includes sites visited and searches. Also includes parental controls for school-issued devices.	reporting group," where the school determines which staff is closest to the student. (Securly, 2019)	Securly includes "TipLine", which is an anonymous reporting for students, parents and teachers. (Securly, 2024)	professional bodies in Securly's online materials.
NetSupport DNA Sources: NetSupportDNA; 2017; 2025	The exact number of schools that use NetSupport is not reported. However, overall they report having "18 million users" (NetSupportDNA, 2025)	Netsupport screens activity by a specific user of a school computer. Any websites, applications, services. Netsupport can also save screenshots, and the students' webcam can be activated (NetSupportDNA, 2025)	According to their online materials, Netsupport screens for (non-exclusively): racism, homophobia, self-harm, eating disorders, bullying, grooming, FGM, drugs addiction (NetSupportDNA, 2025)	Netsupport monitors for specific keywords and phrases. The context and history of a child's activity is also automatically calculated (NetSupport DNA, 2025).	There is no mention of a human moderator within any of the promotional materials.	The website says that "Online safeguarding resources are available" (NetSupportDNA, 2025). however I am unable to access anything specifically.	According to their online materials, they are compliant with PREVENT Duty and KCSIE (Netsupport, 2017).	Parents are able to download NetSupport on devices brought into schools, and can programme these with different options. E.g., if parents want restrictions turned off during the evening (NetSupportDNA, 2025).	Teachers and school safeguarding staff can "add a concern" manually if they are verbally notified of a student's concern (NetSupportDNA, 2025).	Students can report their own concerns (NetSupport DNA, 2025).	IWF, eCadets, the digital institute, go bubble, BESA, digital citizenship institute, GDPR in schools, American Association of Suicidology, Edugeek (NetSupportDNA, 2025).

LightSpeed Systems Sources: LightSpeed Alert, 2025a; 2025b	According to their website (LightSpeed Systems, 2025b) LightSpeed Systems is used in 31,000 schools worldwide, including 90 UK schools under the "Cognita" school-group	According to their website (LightSpeed Systems, 2025), LightSpeed Systems screens anything on school owned devices. This includes search history, online documents, sites visited, and social media comments.	LightSpeed monitors for students' risk of issues such as cyberbullying, self-harm, school violence, and suicide. According to their website (LightSpeed Alert, 2025): "67% of student suicides had clear warning signs that could have been identified by Lightspeed Alert and Safety Specialists".	This software uses an "AI tool" which is integrated with Google Gemini. Teachers can watch student activity live. Materials sent to teachers include "screenshots, browser history, and case notes." (LightSpeed Alert, 2025)	According to their website, lightspeed includes in-house, human specialists, who work alongside the AI alert system to moderate/evaluate all alerts and identify high risk cases. (LightSpeed Alert, 2025)	The website includes policies specifically for different age groups (e.g. KS3 vs KS5) LightSpeed Alert, 2025)	LightSpeed includes a full webpage to discuss their compliance with legislation including GDPR and KCSIE (LightSpeed Systems, 2025a)	According to Lightspeed's website, parents get weekly reports on how their children used the internet, including visits to blocked sites.	According to their promotional materials, teachers can directly check-in with students on this software. LightSpeed Alert, 2025) The school is also able to delegate risk reports to safety resource officers, administrators, counselors or teachers.	Students can report their own concerns about others using the "StopIt" programme, and concerns about themselves using "HelpMe." Students can upload screenshots, photos, or multimedia. LightSpeed Alert, 2025)	There is no mention of any approval by professional bodies in LightSpeed Systems' online materials.
Netsweeper. Sources: NetSweeper, 2021; 2023; 2025a; 2025b	There is not enough data to conclusively say how many UK schools use Netsweeper, however, Netsweeper writes that the software is being used in at	According to Netsweeper (2021), the software scans internet and desktop content as well as user-submitted data in real time.	Netsweeper monitors for students risk of issues such as (but not limited to): cyberbullying, self-harm, school violence	According to their website, Netsweeper uses a "AI-driven, dynamic categorisation engine"	According to their website, administrators are available to verify and evaluate risk and context of each report (Netsweeper, 2025a)	There is no mention of any suggested interventions in this company's online materials.	Netsweeper includes a full webpage to discuss their compliance, including with KCSIE (Netsweeper, 2025b)	According to their website, parental access is controlled / varied by the specific school, through a parent portal. This allows parents to set parental controls (e.g. on content) or receive insight into any "OnGuard" results (e.g. if a student is deemed high risk).	According to their website, teachers receive real-time alerts about potentially harmful activities and are responsible for following up. Web materials specifically say that teachers should work with law enforcement agents when using their software (although the specific details of this partnership are not discussed). (Netsweeper, 2023)	Students do not have direct access to their own data, however they are able to submit reports.	There is no mention of any approval by professional bodies in Netsweeper's online materials.

	least 130 sites across Bristol City Council (NetSweeper, 2021).										
Senso.cloud Sources: Sensocloud, 2025a; 2025b	There is no data on how many schools in the UK use senso.cloud, but there are 10,000 schools using this software globally (and it is headquartered in Nottingham) (Sensocloud, 2025a)	This software is in the Microsoft Azure cloud and/or Google Classroom, and monitors text-based content, visual content, online activities, and communication. (Sensocloud, 2025b)	According to their website, senso.cloud detects, bullying, suicidal ideation, self-harm, potential threat to others. (Sensocloud, 2025a)	Includes “AI-based visual threat detection” and keyword monitoring using customizable libraries, including those developed from the IWF. (Sensocloud, 2025b)	Senso.cloud offers an “Assisted Monitoring Service” where trained professionals work alongside teachers to evaluate results. (Sensocloud, 2025b)	Does not offer an intervention. Specific interventions are determined by the school’s safeguarding policies, not senso.cloud.	The website says Senso.cloud is compliant with GDPR, KSCIE, and OFsted. (Sensocloud, 2025a)	The platform does not explicitly involve parents in monitoring.	Teachers monitor student activities, and receive alerts about potential risks. (Sensocloud, 2025b)	Students do not have direct access to their own data, or ways to report any risk.	Accredited with the IWF (Sensocloud, 2025a)
Securus Software Sources: Securus 2025; IWS, 2024; PSE, 2019	Different sources report different numbers of schools using Securus Software. For example, IWF (2014)	Monitors students’ online and offline activities, typed and untyped content. (Securus, 2025)	According to their website, Securus detects: cyberbullying, mental health issues including suicide and self-harm, radicalization, substance abuse.	Screenshots are used to capture incidents that are triggered by keywords. Securus’ Keyword Library guidance comes from IWF and	Safeguarding experts evaluate incidents from Severity 1 (low risk) to Severity 5 (high risk). If there is a high-risk incident, the school	Does not offer an intervention. Specific interventions are determined by the school’s safeguarding policies, not Securus.	According to their website, the use of Securus is compliant with GDPR, KCSIE, Ofsted, and ISI. (Securus, 2025)	The platform does not explicitly involve parents in monitoring.	According to Securus (2025), teachers and DSL’s receive alerts for high-severity incidents. In addition, Securus is integrated with other school-based programmes, e.g. CPOMs	Students do not have direct access to their own data, or ways to report any risk.	Works with IWF and UK Council for Child Internet Safety.

	reports that Securus Software is used in over 3,200 schools; while PSE (2019) reports 3,200.		(Securus, 2025).	UK Council for Child Internet Safety.	is alerted. (Securus, 2025)						
Smoothwall Sources: Smoothwall, 2023a; 2023, b; 2025a; 2025b; 2025c; 2025d, and Department for Business & Trade & Department for Science, 2024	According to the UK government's Directory of UK SafetyTech Providers (Department for Business & Trade & Department for Science, 2024). Smoothwall is used by approximately 40% of UK schools.	Monitors students' web use and digital communications (Smoothwall, 2025c)	According to their website, Smoothwall monitors for risks including: cyberbullying, Grooming, Radicalization, Violence, Child Pornography, Self-Harm/Suicide. "Every 37 minutes, a seriously vulnerable child was identified. This means that activity was detected including references or indicators of self-harm, suicide ideation and other health issues."	There is real-time, content-aware web filtering, and human-moderated monitoring. (Smoothwall, 2025a)	Depending on how much the school wants to pay, Smoothwall can either be a managed system (by Smoothwall itself), or self-managed by the DSL (Smoothwall, 2023b, 2025b, 2025a)	Does not offer an intervention. Specific interventions are determined by the school's safeguarding policies, not Smoothwall. However, Smoothwall does work with "The Key," who includes model policies and safeguarding training. (Smoothwall, 2025b)	According to their website, Smoothwall is compliant with GDPR and KCSIE. (Smoothwall, 2023b)	The name of the parental app associated with Smoothwall is called Qustodio (Smoothwall, 2025b) Promotional materials say parents can "check [their] child's digital activity and adjust your settings. Easily view their activity timeline, browsing history, YouTube views, screen time and more." (described within the website in section 2, title: "key benefits")	Teachers can monitor student activity, respond to alerts, and participate in training. In addition, Smoothwall's website explicitly says that if there is a case deemed as urgent and the software cannot get ahold of teachers, then they will call emergency services. (IBS, 2025; Smoothwall, 2023b, 2025a)	Includes "the pulse," which "Gives students a channel to speak up and schools a way to spot those in need - in just 60 seconds a week" (Smoothwall, 2025c)	Smoothwall works with "Iloveyouguys", IWF, National Online Safety, and is Also is an accredited filter provider by the UK Safer Internet Centre. (Smoothwall, 2025d)

			(Smoothwall, 2023a)								
Schools Broadband. Sources: Schools Broadband, 2025a; 2025b; 2025c; n.d.	The exact number is unknown, however, the website homepage says they have over 3,000 customers (Schools Broadband, 2025a)	According to their website, SchoolsBroadband monitors keyword searches, and " visited sites of interest" (Schools Broadband, 2025c, p.2).	Schools broadband materials refer to monitoring for bullying and safeguarding risk and safeguarding more generally (Schools Broadband, 2025b, 2025c)	Schools Broadband determines risk by using keywords taken from lists put together by: National Crime Agency, Ofcom, IWF. In addition, there is a Schools Broadband specific keyword list.	There is no mention of a human moderator within any of the promotional materials.	Does not offer an intervention, however, according to online materials, Schools Broadband does provide "schools with critical support to identify online incidents, and integrate the management of digital pupil behaviour with offline pastoral care" (Schools Broadband, 2025c, para. 1).	According to their website, Schools Broadband is compliant with KCSIE and PREVENT Duty, as well as with GDPR, the Internet Watch Foundation (IWF) Block List and Image Hash List Home Office Terrorism Block List UK Safer Internet Centre (Schools Broadband, 2025b)	The platform does not explicitly involve parents in monitoring.	The system sends immediate alerts to Safeguarding leads to allow close monitoring as well as immediate intervention (if required.) (Schools Broadband, 2025c, 2025b)	Students do not have direct access to their own data, or ways to report any risk.	Schools broadband reports working closely with the Interational Crime Agency, Ofcom, IWF (Schools Broadband, 2025c, 2025b)

Appendix Two: CUREC Approval

MEDICAL SCIENCES INTERDIVISIONAL RESEARCH ETHICS COMMITTEE
Research Services, Boundary Brook House, Churchill Drive, Headington, Oxford, OX3 7GB
Tel: +44(0)1865 616575
ethics@medsci.ox.ac.uk

UNIVERSITY OF OXFORD

CONFIDENTIAL

Professor Ilina Singh & Jessica Lorimer
Department of Psychiatry
University of Oxford
Warneford Hospital
Oxford

2 February 2022

Dear Professor Singh and Jessica,

Research Ethics Approval - CUREC 1

Ethics Approval Reference: R78840/RE001

Study title: Ed-Tech and Ethics: Monitoring Suicide Risk in UK Schools

Short title: Teacher Perspectives on Monitoring Students' Mental Health in UK Schools

The above application has been considered on behalf of the Medical Sciences Interdivisional Research Ethics Committee (MS IDREC) in accordance with the University's procedures for ethical approval of all research involving human participants.

I am pleased to inform you that, on the basis of the information provided to the IDREC, the proposed research has been judged as meeting appropriate ethical standards, and approval has been granted for a period of **18 months**, commencing on **2nd February 2022.**

Amendments
Should there be any subsequent changes to the study, you should submit details to the MS IDREC for consideration and approval. Details of changes must be listed on an amendment form.

Yours Sincerely

DocuSigned by:
Leah Butts
9F14889D2BC549A...

Mrs Leah Butts
Research Ethics Administrator

for
Dr Helen Barnby-Porritt
Research Ethics Manager

Appendix Three: CUREC Approved Consent Form

CONSENT FORM FOR TEACHERS

Central University Research Ethics Committee (CUREC) Approval Reference: R78840/RE001

Teacher Perspectives on Monitoring Students' Mental Health in UK Schools

Please initial each box

1	I confirm that I have read and understand the information sheet for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	
2	I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, and without any adverse consequences or penalty.	

3	I understand that research data collected during the study may be looked at by authorised people outside the research team. I give permission for these individuals to access my data.	
4	I understand that this project has been reviewed by, and received ethics clearance through, the University of Oxford Central University Research Ethics Committee.	
5	I understand who will have access to personal data provided, how the data will be stored and what will happen to the data at the end of the project.	
6	I understand how this research may be written up and published.	

7 I understand how to raise a concern or make a complaint.

8 I understand that researchers will observe lessons and other aspects of my teaching, as detailed on the information sheet, and discussed and agreed with the researchers.

9 [I consent to being audio recorded

12 I understand how audio recordings will be used in research outputs

13c I give permission to be quoted directly in research outputs but only fully anonymously

14 I agree to take part in the study

dd / mm / yyyy

Name of Participant Date

Signature

dd / mm / yyyy

Name of person taking consent Date Signature

Appendix Four: Interview Guide

Introduction: "Thank you for signing up for this interview, which is part of my PhD in the Neuroscience, Ethics and Society group at the University of Oxford. We will be talking about mental health and technology in schools, as I'm interested to hear teachers' opinions on what they think their roles and responsibilities are for monitoring students' mental health. Some of what I'm asking about is a bit sensitive, because it relates to student mental health; and it's important to know that I'm really just asking about your views and experiences – no right or wrong answers! I am recording the interview, but it will be fully anonymous, and I won't report anything you say individually to the school. We can pause or stop the interview at any point. The whole process should take no longer than one hour. Do you have any questions about this, or any of the forms I sent via email earlier?"

"Great, let's begin the interview. In the first section we're going to speak briefly about student mental health at your school."

Theme	**Questions**
Mental Health General	• Based on the questionnaire you filled out before this interview, I understand that your role in school is _____. Can you describe your role and how it relates to student mental health? • As part of my PhD, I have been reading about ways that (*YOUR ROLE, e.g. teachers*) can be involved in their students' mental health; for example: Detection of student risk, Support during crisis, and/or Post-crisis support. Are you involved in all of these areas? • It seems that lots of different people work with student mental health – teachers, administrators, parents, healthcare professionals. What is unique about your role as a ____? How is your role different to that of a parent or healthcare professional?

<u>Definition of Responsibility</u>	• In my readings, there is a lot of discussion about being 'responsible' for students' mental health. I've been trying to understand this in practice. What does "being responsible" for your students' mental health mean to you?
	• Can you tell me a time where you have taken responsibility for one of your student's mental health?
	• In the example you just gave, (*if no example was given, then "in general"*) do you think there was a difference between legal requirements, what you "ought" to do, and what you "actually" did?
	• In the example you just gave, (*if no example was given, then "in general"*), did you work with other groups of people? (e.g. parents or healthcare professionals)?
	• If you did work with other people, how were your responsibilities different?
	• Follow-up Prompts:
	    • Why did you say that teachers were responsible for ______ but not for ____?
	    • Why do you think ____ is the role of the parent instead of the teacher?
	• From what I heard, you are describing responsibility in ____way. Is that correct? How did you decide that you were responsible in ____ way and where did you get this information from? For example, did your school provide training or support?

"Now we're going to move on to the second section of this interview, where we talk about technology and the software available at your school."

Technology	Questions about Type of Software
	● In the news recently there has been a lot of discussion about different types of new technology brought into schools to do things like monitor student's internet use. Does your school monitor students' digital activities?
	● If yes:
	  ○ Do you know the name of the software programmes being used?
	  ○ In your own words, how does this technology work and what is it used for?
	  ○ Has this technology been around for a while, or did your school invest in new technology due to COVID-19?
	● From what I understand, these programmes are being used for a lot of things – for instance, making sure students don't connect to illegal websites or to monitor for bullying. Has this technology (or other types) impacted how you handle student mental health specifically?
	● If yes:
	  ○ Which part of the "pipeline" is this technology used for? Detection of student risk? Support during a crisis? Post-crisis support?
	  ○ Do you know how this technology works? For example:
	    ▪ What kind of information is gathered by this software? (e.g. social media data, google searches)
	    ▪ Is this data anonymous? What non-anonymized data does this gather about individual pupils?
	    ▪ Who has access to this information? (both anonymous and non-anonymized)
	● If this technology is used in your school, how do you specifically use the technology? Can you tell me of a time you used this technology.
	● Do you think you use this technology more or less than other staff members? Why/can you expand on this?

"I'm going to now ask you to read a short case study." (Show on screen).

Your school has been using the software CommonX to filter online content (e.g. making sure students don't access inappropriate websites in school). You receive an email update from your IT manager, saying that CommonX has a new add-on: one which safeguards students and can provide suicide risk predictions. Its use will be covered by the general IT privacy agreement which parents signed earlier in the school year. Later in the week, on a Thursday evening at 8pm, you receive a message to say that one of your students is deemed high risk of suicide. When you click the link attached to the email, you discover that this student has googled "why does someone commit suicide" on their personal home computer.

Case study Questions	• What are your initial thoughts on this case study? • What are the benefits of this program? • Are there any elements of this story that concern you? • Does something like this happen at your school? If so, how similar or different is this case study to what is happening in your school right now?

Risk and Risk Thresholds	• The case study talks about a students' "risk of suicide." What does the phrase "risk of suicide" mean to you?
	• In your opinion, what are the warning signs of suicide in a student?
	• In your opinion, what causes suicide?
	• Does your school monitor for students' risks of suicide? This can be with a program like CommonX or in some completely different way.
	• *If yes:*
	  ○ Could you expand on how you monitor students for risk of suicide?
	  ○ When a student at your school is considered at risk for suicide or mental health challenges, what happens next?
	• Do you think a programme like CommonX can help your school predict which students are at risk of suicide? Why or why not?

Protocol	● If your school used a program like CommonX, you may occasionally receive a notification that one of your students is at "high risk of suicide." How would you respond to a notification like this? For example: ○ Does your school have a procedure on student mental health/safeguarding that would apply in this situation? If so, what kinds of protocols/procedures are there? Are there written guidelines? ○ Is there a process of documentation? Should you inform other people (support services), or do you work alone? ● If your school already monitors students for risk of suicide and other mental health conditions, how regularly do you have to follow this protocol/report a student as "high-risk?"
Shared Responsibility	● When you are dealing with student mental health, how often do you work with the NHS/police/other social services? Which group do you work with the most/least? (*If the participant needs prompting, they can rank*: parents, NHS, the students themselves, tech companies, police/other social services) ○ When have you involved ____ in student mental health? Can you give an example? ○ How does your partnership with ____ differ from _____? ● When thinking about technology like CommonX in the case study, do you think it's best to work with other groups? If so, who and how? ● E.g. Parents, NHS, the students themselves, tech companies, police/other social services? ● If you listed multiple groups: ● How should responsibility be shared or distributed? ● Does one group have more "oversight"?

<u>Accountability and Governance</u>	• In my research I am reading about a lot of software programmes like CommonX. Some of these software programs work better than others, and there have been some cases where students have been flagged as "high risk" incorrectly. There have also been some cases where the software programs have missed students.
	• Are you worried about programmes like CommonX going wrong, for example flagging too many students or missing some?
	• If something does go wrong can anyone be "blamed"/who do you think should be held accountable?
	• You said that ____ should be held accountable if something went wrong. What do you think should happen to this person/group? Should there be legal repercussions?
	• The government is looking to regulate this technology and are coming up with ways to evaluate programmes like CommonX. This is called the "Online Safety Data Initiative." Some of the ways they are looking to evaluate these products are listed on the slide on the screen. These include checking for:
	    o Performance (e.g. how well the models do in predicting who is at risk accurately)
	    o Robustness (e.g. whether or not the models can still hold up after rigorous testing)
	    o Fairness (e.g. whether these models are impartial and do not discriminate)
	    o Scalability (e.g. can this technology be used on a large scale and still be accurate)
	    o Explain-ability (e.g. do these software companies provide rationale why certain people are flagged and not flagged)
	    o Privacy (e.g. do these programmes maintain young people's privacy?)
	• What do you think about these criteria?
	• Which of the factors discussed are you more or less worried about?
	• Are there any factors that you think might be missing and the government should be using to evaluate technology like CommonX?

"Do you have any other thoughts or comments about technology like CommonX that you would like to share?"

"Thank you very much. This is now the end of the interview."

Appendix Five: Demographic Questionnaire

Participant	Age	Gender	Ethnicity	Parent?	Region	Type of school	Age range	FSM	SEN	Job Title	Teaching Experience	Certification	training for MH	How?
10	25-34	M	White British	N	London	Community school/local authority maintained school	secondary only	Y	Y	Assistant Headteacher	7	Teach First PGCE	Y	Inset days
11	25-34	F	White English	N	South East	Community school/local authority maintained school	secondary only	N	Y	English Teacher	5	PGDE	N	-
12	25-34	F	White English	N	London	Academy	secondary only	Y	Y	Head of History	7	PGCE	Y	Teacher Training
13	25-34	F	Chinese	N	South East	Private School	secondary only	unknown	N	Teacher	5	PGCE	Y	Self study, teacher training
14	45-54	F	White English	?	South East	Private School	secondary only	N	N	Biology and PSHE tutor and SENCo	27	PGCE	Y	Inset days
15	35-44	F	White English	N	South West	Grammar School	secondary plus sixth form	N	(average)	Teacher of Religious Studies, EDI Student Coordinator	8	PGCE	Y	Inset days
16	45-54	M	White English	Y	West Midlands	Private School	secondary, but GCSE onwards	N	N	Deputy Principal, Teaching & Learning	14	PGCE	Y	self study, inset days
17	25-34	F	White English	N	South East	Private School	mix primary and secondary	N	N	Head of Religious Studies	9	PGCE	Y	inset days, teacher training
18	35-44	F	White English	N	South East	Academy	secondary only	N	Y	Head of Sixth Form	8	SCITT ITT and PGCE	Y	inset days, teacher training, mental health first aider
19	25-34	M	White English	N	South East	Free School	secondary only	Y	N	Deputy Head of Sixth Form and Teacher of English	11	PGCE (teach first) MSc (Learning and Teaching)	Y	Inset days, mental health first aid