Safe Spaces and Epistemic Character: Corruption, Resilience and Protection in the Classroom

Alice Monypenny

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

Department of Philosophy

March 2022

Abstract

This thesis offers an account of safety in the classroom and focuses on the relationship between safety and protection. Using a critical character epistemological lens, I argue that a student's ability to protect themselves from harm may vary in relation to their social positionality or carry significant costs.

I demonstrate how a range of epistemic character traits can aid students in protecting themselves from harm in hostile classroom environments. When theorising about epistemic character is done through a critical lens, the relationships between epistemic character and power structures in the social world are emphasised. I argue that the range of protective epistemic character traits available to a student will be affected by a range of factors, including how relationships of power and oppression affect the trajectory of their epistemic character development. Furthermore, the development and exercise of protective epistemic character traits, especially those which are conventionally classified as epistemic vices, can carry significant costs for students. Two such costs are (i) damage to epistemic character – the development of epistemic character in ways that are at odds with epistemic flourishing – and (ii) the constraint of epistemic character.

I then illustrate the relationship between protection and safety by presenting the following account of safety: a classroom is safe for some student S if being in and participating in the classroom does not put S at higher than tolerable level of risk of harm without S being required to protect themself in ways that cannot be reasonably expected of them. Means of self-protection which students cannot be expected to adopt include those which cause, or make students vulnerable to further harm, and those which significantly disadvantage them in other ways. I argue that unless educators have detailed knowledge of a student's ability to take on risk, then safety is required to make the classroom accessible to that student.

Parts of this thesis have been published in the following paper:

Monypenny, A. (2021) Between Vulnerability and Resilience: A Contextualist Picture of Protective Epistemic Character Traits. *Journal of Philosophy of Education,* 55(2) pp.358-370.

Acknowledgements

I have been very fortunate in having a large supervisory team. I am extremely grateful to Katie Monk, Aness Webster, Ian Kidd, Andy Fisher and Katharine Jenkins for all their encouragement, support and inspiration. Part of this thesis was written during a research visit to University of Connecticut: I thank Heather Battaly for all of her support during this visit.

I am grateful for the financial support I have received from Midlands 4 Cities Doctorial Training Partnership.

Thanks to Lilit Movsysian who first introduced me to vice epistemology when she invited me to join her on a trip to Durham in 2015 to attend a workshop on a topic which I knew nothing about. I would also like to thank all of the following for helpful comments and fruitful philosophical discussion: Lani Watson, Charlie Crerar, Heather Muraviov, Taylor Tate, Tracy Llanera, Taylor Mathews, Jon Robson, Jane Gatley, Simon Thunder, Riana Popat, the philosophy postgraduate community at the Univesity of Nottingham, the Nottingham and Sheffield Vice Squad and the Midlands Philosophy of Education postgraduates group.

I am very grateful for my friends and family for the support and encouragement including my parents, Fliss Radomska, Francis Myerscough and Snehal Shah. Finally, a very special thanks to Janset Çetinkaya and Andrea Raimondi for being there and helping me stay focused and smiling through the ups and downs of doing a PhD during a pandemic.

Table of Contents

Introduction **5**

0.1 Critical Character Theory **6**

0.2 Higher Education, Safe Spaces and Students' Unions **8**

0.3 Conceptual roots of safe spaces in liberatory and activism movements **12**

0.3.1 Spaces to meet specific needs 12
0.3.2 'Homeplace': Spaces for sanctuary and respite 16
0.3.3 A space for independence and counterculture 18

0.4 Overview of the thesis **21**

Chapter 1: Safety in the classroom: conceptual issues **23**

1.1 Student perspective on safety: two types of testimony **24**

1.2 Being safe and feeling safe: objective and subjective risk **28**

1.3 Safety, comfort and negative emotional responses **33**

1.4 The value of Safety in the Classroom **39**

1.5 Summary **44**

Chapter 2: Epistemic character and protection from harm **45**

2.1 Character epistemology and education **46**

2.2 Risks of harm in hostile classrooms **50**

2.3 Protective Epistemic Character Traits **51**

2.3.1 Strategy 1: 'Battle on through' 52
2.3.2 Strategy 2: 'Believe in yourself' 53
2.3.3 Strategy 3: 'Just don't listen' 53
2.3.4 Strategy 4: 'Keep your head down' 55

2.4 The shaping of character development **57**

2.4.1 Costs of developing and exercising PECTs 58
2.4.2 Oppression and the trajectory of epistemic character development 60
2.4.3 Failure to be protective 61

2.5 Implications for teaching epistemic virtue **62**

Chapter 3: Epistemic Character Damage and Normative Contextualism **66**

3.1 Epistemic Coping Strategies **68**

3.2 Epistemic Corruption 71
3.2.1 Epistemic Character Damage 72
3.2.2 Epistemic Coping Strategies as Epistemic Corruption 74
3.3 Normative Contextualism 76
3.4 (Resolving) the tension 79
3.4.1 Strong and weak normative contextualism 80
3.4.2 Dual-level normative contextualism 82
3.4.3 Resolving the tension 86
3.5 Conclusion 88

Chapter 4: Constraining Epistemic Agency: Costs of PECTs from the 'keep your head down' cluster 90

4.1 Epistemic Agency 90
4.2 The reliability of epistemic powers 95
4.2.1 Access to shared epistemic resources 96
4.2.2 Uptake and recognition 99
4.3 Exercise of powers 104
4.4 Intellectual Servility as a PECT 106
4.5 The impacts of intellectual servility on epistemic agency 108
4.6 Agency and Character 111
4.7 Conclusion 114

Chapter 5: Safety, Risk and Protection: Access Issues 116

5.1 Safety and Risk 118
5.1.1 Tolerable Risk 118
5.1.2 Safety and Protection 122
5.1.3 Protective measures which are not genuinely available to students 126
5.1.4 Protective measures which disadvantage students 128
5.1.5 Harmful protective measures 131
5.2 Access 132
5.2.1 *Actual-* and *Presumed-Accessibility* 132
5.2.2 Presumed-accessibility and safety 135
5.3 Two case studies 137
5.3.1 Content Warnings 137
5.3.2 Dignity harms 140
5.4 Summary 145

Conclusion 147

Bibliography 150

Introduction

In November 2015, the University of Bristol's Journalism Society invited the controversial journalist Milo Yiannopoulos to speak on campus. The running of the event was opposed by others at the university due to Yiannopoulos's previous transphobic comments (see for example Yiannopoulos, 2015) and his statements denying the reality of rape culture (Wise, 2015) and it was questioned whether it was appropriate to offer him a platform at the university. In particular, the university's Feminist Society argued that the event would violate the Student Union's Safe Space policy. A statement from the committee of the Feminist Society outlined their stance:

> "As a committee, we do not support the University of Bristol Journalism Society's decision to host Milo Yiannapoulos [sic] on November 20th, 2015. We feel that this violates the University's safe space policy, and we are not in favour of the kind of hate speech and vitriol that Yiannapoulos [sic] perpetuates.
>
> "We believe that all societies should uphold the safety and security of all students at the University, and we feel that the Journalism Society has failed to take this essential factor into consideration." (Quoted in Wise, 2015)

The event generated a great deal of discussion and there was a range of opinions within the student body (see for example Brar, 2015). In the end, the event went ahead but in a revised form: Yiannopoulos debated journalist Rebecca Reid on the question 'have we reached an age of gender equality?'.

I was an MA student at the University of Bristol at the time, and the debates preceding the event as well as the resulting backlash against the Student Union's Safe Space policy sparked my interest in the topic of safety in education. The debate around the event and the value of safe space policies more generally was oftentimes clouded by a lack of clarity on central

terms. Particularly, what was meant by the 'safety' which such policies were supposed to protect?

The creation of safe spaces or the implementation of safe space policies by universities, students' unions and student groups has generated a great deal of debate, garnering input from both within and beyond the institutions themselves, touching on topics such as freedom of speech, censorship and the role of the university in contemporary society. Since the early 2010s numerous commentaries on safe space policies and their effects (particularly their relationship to 'no-platforming' – the disinviting or banning of particular external speakers on the grounds of their endorsement of hateful views) have been published by academics, journalists and student activists.[1]

In this thesis, I focus on a subsection of that debate, relating specifically to safety in the classroom itself. One of the central questions which this thesis addresses is 'what does it mean for a classroom to be safe?' The proposed account offers a conceptual framework for understanding the relationship between safety, risk and self-protection. In order to make judgements as to whether a classroom is safe, or which measures must be put in place to make it safe in particular cases, significant amounts of contextual information must be inputted into the account. The aim of this thesis is to show what we must do with that information in order to answer questions such as, is the classroom safe for a particular student and what responsibilities do educators have regarding making the classroom safe for a particular student. In that regard, I do not attempt to give any concrete answers as to what a safe classroom looks like, or whether any particular classroom is safe (besides the stylised examples which I use to illustrate my account). I will explore the various different meanings of safety further in chapter 1, but it is worth mentioning here that I take safety to be broader than just physical safety and to also include emotional or psychological safety.

0.1 Critical Character Theory

Much of the pushback against the claim that events like the Yiannopolous debate contravene the requirements of safety centres around the idea that

[1] For examples of criticisms of safe spaces in universities in the media see Pells, (2017), Bergamini (2018) and Whitten (2018) and for responses to critics, see Bouattia, (2018), Wilson (2018) and Roth (2019). Examples of scholarly engagement with the topic include Palfrey (2017), Ramsay (2017) and Anderson (2021)

intellectual development requires the willingness to be exposed to and engage with views contrasting with one's own.[2] If education is too safe, then students will not be pushed beyond their comfort zone, learn to think critically or develop a variety of intellectual skills (see for example Ramsay, 2017). The university is, after all, supposed to be a place of learning and intellectual development. But such arguments take for granted the claim that being exposed to, and engaging with, expressions of hateful or otherwise harmful views contributes to intellectual development in a positive way.

The questioning of this assumption leads to the second central theme of this thesis – the way in which epistemic character[3] (roughly the set of dispositions, attitudes and ways of thinking which guide how an individual goes about engaging in epistemic activities) develops as a protection against and as a result of risk of harm in the classroom. My approach to theorising about epistemic character is largely informed by the principles which Kidd (2020, forthcoming) outlines in his 'critical character epistemology'. Understanding the development of epistemic character requires understanding its socially embedded nature. In particular, it requires attention to the ways in which character is shaped by oppression. Kidd argues that to understand epistemic character as shaped by and shaping the social world, our theorising must be sensitive to the source and origins of epistemic character traits and recognise that the value of particular character traits is contextually dependent (Kidd, 2020).[4]

One of the ways in which social environments impact the development of epistemic character is through the phenomenon which Kidd (2020, 2021) terms 'epistemic corruption': damage to an agent's epistemic character resulting from interactions with agents, environments, institutions, or structures that facilitate the development and exercise of epistemic vices.

---

[2] Another central theme is the opposition to censorship, protection of freedom of speech and academic freedom (see for example Hantzis and Looser, 1995; Ben-Porath, 2017; Palfrey, 2017; Shahvisi, 2018). While these issues intersect with considerations of safety, I will not address them directly in this thesis.

[3] The terms 'intellectual character' and 'epistemic character' have been used largely interchangeably in the literature. For the sake of simplicity, I will use 'epistemic character' (and correspondingly, 'epistemic virtue' and 'epistemic vice') throughout this thesis.

[4] A further feature of critical character epistemology which Kidd (2020) describes, but is less relevant for this thesis is 'axiological pluralism': the appeal to a range of different values in order to explain the goodness or badness of epistemic character traits (2020, pp.80-81).

In this thesis, I focus on the relationship between the risk of harms and the development of epistemic character, by examining the ways in which character may develop in response to, and as a means to protect oneself against the risk of harm.

0.2 Higher Education, Safe Spaces and Students' Unions

In this thesis, I limit my discussion primarily to higher education. When considering safety in relation to younger students there may be additional factors that come into play – for example, the safeguarding responsibilities of schools. I will not attempt to address these issues here. That being said, I present my account in chapter 5 in abstract terms so that it applies to a range of classrooms.

As noted above, I focus on a subset of the debate around safety at higher education institutions: safety in the classroom. However, it is worth situating this discussion within the wider context. I will give a brief overview of two areas which provide a useful backdrop to the topic. In this section I discuss safe spaces within the university but beyond the classroom (and particularly within students' unions), and in the next section, the creation of separate and protected spaces within activist and liberatory movements.

To begin, there is a distinction worth noting between the creation of particular safe spaces within the university (such as within the classroom, particular student services or support networks) and the notion that the university as a whole is a safe space[5] (Ramsay, 2017, p. 6). It is much easier to create and ensure safety in small, self-contained spaces – for example, by setting rules for participants or controlling who uses the space. Such spaces also are much more likely to have a defined aim or purpose. For example, a support group for LGBTQ+ students has an explicit aim of providing a space in which LGBTQ+ students can seek support from their peers and talk openly about their experiences.

On the other hand, the notion of the university as a whole being a safe space is much more complex. It is much more difficult to set standards and expectations for behaviour because of the vast variety of different spaces within it. Expectations regarding how students should conduct themselves in an LGBTQ+ support group will be different to how they can be expected

[5] I use the term 'space' in a loose way to include a range of physical and non-physical settings.

to conduct themselves at social events or at the students' union bar. Much of the university is a public space, in the sense that it is freely open to anyone to come and go as they please (in some cases, there is no clearly demarcated campus at all) and therefore it can be difficult to define the physical space which is covered by the notion of the university as a safe space and to whom the behavioural expectations apply.

Another complicating factor is the variety of purposes which the university campus serves. The university is an educational institution and therefore, as a space, it by and large serves educational purposes. However, very often, students (and some staff) live and socialise on campus. To this extent, it may be expected that some spaces within the university offer a place for respite. Even if students are expected to engage with views that challenge their own as part of their university education, they may be entitled to retreat from this in their leisure time, or expect to feel 'safe' in the place in which they live. For example, consider the very recent incident at Durham University in December 2021 in which right-wing columnist Rod Liddle gave an after-dinner speech at a college formal dinner. In the speech Liddle made transphobic comments and derogatory remarks regarding sex-workers. When students walked out of the speech in protest, the college principal derided them as 'pathetic' for being unwilling to listen to views which challenged their own (Wolfe-Robinson, 2021). However, the speech did not take place in a purely educational setting – it was a social event within the students' place of residence.[6] This difference in the purpose which the space serves may lead to differences in the way which students, speakers and other guests are expected to behave when compared to purely educational settings.

In the case of the Yiannopoulos debate at the University of Bristol, although the event took place on university property, it was organised by a student group and the safe space policy being allegedly violated belonged to the students' union rather than the university itself. Therefore, the decisions being made regarding the case were made within the students' union and

[6] It is reported that students were not informed that Liddle would be speaking when booking a place at the dinner (Cunliffe, 2021).

See also Ben-Porath's discussion of an incident at Yale in which students argued that the role of a professor with responsibilities overseeing student residences was 'not to foster intellectual debate and difficult conversations but rather to create and protect a safe space – a "home" environment – for students.' (2017, p. 80).

the relevant student groups rather than by any member of staff within the university.[7]

Within the debate regarding safe spaces in education, there are three interesting things to note about the role of Students' Unions. Firstly, many of the spaces which their safe space policies cover serve a quasi-educational purpose. Students' Unions host events, student groups, support networks and social spaces which are in many ways distinct and separate from the university classroom. They serve to support students during their university education, but their primary aim is not always educational. Secondly, because the unions are student led, the safe space policies are examples of expectations and demands of safety issued by (at least a sub-set of) students themselves.[8]

Thirdly, students' unions, and the groups contained within them often have explicit liberatory goals[9]. Students' Union safe space policies often state a commitment to diversity, equality and inclusion and note that there is a lot of work still to be done to achieve these things (see for example Edinburgh University Students' Union, no date; Kings College London Students' Union, no date; Bristol Students' Union, 2014). They note the discrimination and oppression which some members of the student body may face.

In general, these safe space policies are codes of conduct or statements of expectations of members' behaviour. Those which provide more detail list explicitly the kinds of behaviours which contribute to the exclusion or lack

[7] Other examples include the cancelling of an event involving Yiannopoulos and Germaine Greer as the University of Manchester for the reason that the invitation of either speaker contravened the Unions' safe space policy (The Mancunion, 2015).

[8] There is by no means unilateral support for safe space policies across the student body. For example, in a survey conducted by the Higher Education Policy Institute in 2016, just under half (48%) of full-time undergraduates surveyed answered 'yes' to the question 'Do you think universities should adopt safe space policies?' (Hillman, p.50-51).

[9] For example, several students' unions (including University of Nottingham Students' Union, Bristol Students' Union, Edinburgh University Students' Association, Royal Holloway Students' Union, and University of Manchester Students' Union) have elected 'liberation' or 'diversity' officers focused on addressing issues of liberation, equality, diversity and inclusion within the student body.

of safety of certain students. Others give general statements of their aims and leave the exact details open to interpretation.[10]

Several policies give some explicit indication of exactly what they take 'safe space' to mean: a 'space which is welcoming and safe and includes the prohibition of discriminatory language and actions.' (Students' Union University of Manchester, no date) but which is not equivalent to a space involving no debate (Sussex Students' Union, no date; Bristol Students' Union, 2014)

Some policies (for example Edinburgh University Students' Union, no date; Goldsmiths Students' Union, 2012) set out in concrete terms the kinds of expectations and regulations which they entail. This may include certain behaviours which are prohibited or met with zero tolerance. These behaviours include inciting hatred or establishing groups or societies which promote ideologies opposed to the constitutional aims of the union, creating an intimidating environment for other members or staff, making sexual comments or unwanted sexual invitations, unwanted physical contact, and denying access on the grounds of irrelevant characteristics.[11]

Furthermore, some (for example, University of Manchester Students' Union) extend their policies to invited speakers, such that organisers and groups have a responsibility to challenge speakers violating expectations set by the policy. Others (Sussex Students' Union, no date; Bristol Students' Union, 2014) require that safe spaces be identified and reaffirmed through verbal and physical notices. For instance, physical signs might be required in rooms on Students' Union premises and statements summarising the policy must be included on event descriptions and invitations, particularly on social media, and read aloud at the beginning of events. The reason for this is more than just to remind members and guests about the ways they are expected to behave, but also to reassure them that the policy will be upheld, helping people to *feel* that they are safe.

---

[10] The statements discussed below were collected from the policy documents and websites of Students Unions around the U.K. in 2018. Some policies are still in place but others may have been reviewed since.

[11] Some of these behaviours may be difficult to define and identify – for example it can easily become contentious whether a particular environment was really intimidating or not depending on whose perspective we take to be relevant. Those who are in positions to implement these policies must be careful not to reinforce power dynamics by ignoring the claims of marginalised members when they identify problematic behaviours.

There can be a range of different types of spaces which are created within the university beyond the classroom. These will serve a variety of goals and be adapted to suit the variety of different settings in which they are created. As can be seen in the students' union safe space policies discussed above, safety can be used to serve the aims of liberation, equality and inclusion. I will discuss further in chapter 1 whether, within the classroom, safety aids or hinders the achievement of these aims.

0.3 Conceptual roots of safe spaces in liberatory and activism movements

To illustrate how safety might serve liberatory aims, it is worth first considering some parallels with the creation of safe, protected and separate spaces within activist and liberatory movements. Whilst the term 'safe' was not always used to describe these spaces, some of the motivations behind the creation of safe spaces in education can be understood by comparison to these spaces. The examples discussed below come from a range of different contexts and many differ greatly from the classroom, but I use them to identify a number of themes which may be present in the background when thinking about safety in the classroom. In the examples discussed below, the space is 'protected' in a sense from oppressive structures, practices or individuals which can cause harm or damage or derail the liberatory projects for which the spaces are created.

Moira Kenney (2001) argues that the term safe spaces, as commonly used in US contexts today, has roots in the LGBTQ+ community in California during the 1960s (where certain venues were identified as safe in the sense that they were LGBTQ+ friendly) and the women's liberation movement of the 1960s and 1970s – where the term 'safe space' was used to signal spaces which attempted to distance themselves from and offer escape from patriarchal norms.

I will discuss three themes below: spaces in which the achievement of liberatory aims are promoted by protecting against oppressive forces which hinder the achievement of those aims; spaces for sanctuary and respite; and spaces which are sympathetic to marginalised voices and projects.

0.3.1 Spaces to meet specific needs

One sense in which a space can be safe for members of a particular social group is if the space is designed to be welcoming for and serve the needs of members of that group. Under conditions of oppression in which the voices of members of oppressed groups are not taken seriously – or worse, are

silenced – taking seriously and meeting these needs can be in itself a radical act. As a result, oppressive norms, structures and practices may make meeting these needs difficult. Spaces – such as support groups, community centres, activist meetings – therefore need to be protected and protect their members from the influence of such oppressive norms, structures and practices. This may be done in a variety of ways.

Within the US Womens' Liberation Movement of the 1960s and 70s, 'consciousness-raising groups'[12] created physical and discursive spaces for a form of collective reflection which was discouraged within patriarchal society. Pamela Allen (1973) – a member of San Francisco feminist group Sudsofloppen and co-founder of New York Radical Women – writes about the process of women getting together and talking about their experiences in domains of life often characterised as personal and non-political, to gain a better understanding of women's experiences of oppression.

But at the time, the action of women getting together and discussing their experiences of the everyday was in itself radical and was met with resistance (Sarachild, 1978, p. 145). The lack of acceptance and support for such an endeavour motivated the need for the creation of dedicated spaces – meetings which were designed and structured for this specific purpose. Allen describes the development of the group *Sudsofloppen*: 'We defined our group as a place in which to think: to think about our lives, our society, and our potential for being creative individuals and for building a women's movement.' (1973, p. 271).

They termed what they created 'Free Space'. Allen noted that it was important for the group to offer positive responses to, and affirmation of, individual women's experiences. This helped women to open up and talk about their experiences and feelings after being discouraged from doing so outside of the group. It is important, Allen emphasises, 'that someone listens and does not ridicule; someone listens and acknowledges the validity of another's view of her life.' (1973, p. 274).

Some tension arose in the running of consciousness-raising groups around the balance between policing behaviour to prevent internalised patterns of oppression from creeping into the room and allowing women to break free

[12] Similar practices also appeared under different names. For example, Pamela Allen (1973) uses the term 'Free Space' for an almost identical practice. Echols (1989, pp. 83–84) notes that although the term 'consciousness-raising' was original to the Women's Liberation Movement, early advocates for the practice were largely influenced by similar practices being used within the Civil Rights Movement.

from sexist stereotypes of 'correct' feminine behaviour. In particular, historian Alice Echols (1989, p. 88) describes a debate that arose in many groups and was labelled in one as 'let the quiet women talk'. An unstructured approach to the running of the groups allowed some more assertive members to dominate and the atmosphere became confrontational. Moves to make the groups more egalitarian were eyed with suspicion because they were seen as censoring women who did not conform to feminine stereotypes of politeness and quietness. Policing of the dominating behaviour was seen as enforcing the patriarchal idea that women should not be bold and assertive but meek and gentle (*ibid.*, pp.88-89). Often these assertive behaviours were developed specifically to help women get their voices heard in a patriarchal world, but in the context of the consciousness-raising group, where space was made for women's voices, they had the potential to silence others. This demonstrates some of the difficulties which can arise in meeting diverse needs even within a specific group. Sometimes the measures put in place to make the space welcoming to some will exclude others.

Consciousness-raising groups tended to exclude men on the understanding that women[13] would be able to speak more freely without men present. The presence of a male voice within the group would inevitably change the dynamics due to the socialised and learned behaviours developed within a male supremacist society (Allen, 1973, p. 40; Sarachild, 1978, p. 144). Even if men were present but did not speak – or were limited in the way that they could participate – the organisers worried that this could still impact the dynamic of the group.

In similar examples, the exclusionary nature of such groups has been criticised. Individuals are excluded on the grounds of gender, race or some other characteristic and if the example were reversed, it is sometimes argued, so that, for instance, a men-only space was created, it would be criticised on the grounds of being exclusionary. In 2015, Goldsmiths (University of London) Black and Minority Ethnic (BME) Network came under fire for asking those who do not self-identify as BME not to attend one of their events.[14] Hajeera Begum, BME officer at University College London at the time, spoke in solidarity with Goldsmiths BME network:

---

[13] Although the original consciousness-raising groups may not have framed it in this way, it is worth extending this – as is done by many contemporary feminist groups – to include all those of a marginalised gender, including those who are non-binary.

[14] It is worth noting that there was no exclusionary policy in place which could prevent non-BME students from attending. Instead, a request was made that they did not attend. For

> 'The importance of self-defining spaces comes from the power of self-organising. Only students who have experienced specific oppressions know what solutions they need and it is important that these students are given the space to direct this. Whilst allies in defeating oppression are appreciated, it is important in discussions and organisational meetings that the individuals that are experiencing oppression are the ones leading their liberation. This is not always possible when individuals with more privilege are present as the power imbalance can dominate and drown out oppressed voices.' (2015)

The broader context is relevant in determining whether a separate space is justified. In the case of both the consciousness-raising groups and the Goldsmiths event, the spaces were being created for a purpose which could not be achieved, or was difficult to achieve, in the broader community because of oppressive structures.

Another potential site of disagreement is the boundaries of who is excluded and who is not – such as the question of whether the inclusion of trans women in women-only spaces hinders the aim for which the space was created. For example, in the case of women-only spaces for survivors of sexual violence, the inclusion of trans women has been opposed on the grounds that the space is intended for women who have been victims of sexual violence perpetrated by men and the presence of those they perceive as men (regardless of how those individuals self-identify) in the space could be triggering (Manners, 2019, p. 8).[15] However, sexual violence and LGBTQ+ support practitioner, Paula Manners (*ibid.* pp. 8-9, 13), argues that this does not justify the exclusion of trans women. It would not be considered justified, she argues, to exclude someone with a particular regional accent from the space just because another user of the space was triggered by that accent (because, for example, their abuser had that

---

critics of this kind of separate space, the absence of exclusionary policies may alleviate some worries, but they may still criticise the atmosphere of exclusion which this supposedly creates.

[15] Some arguments (see for example Lawford-Smith, no date) go a step further and argue that such spaces are created for women who are victims of sexual assault, of which the perpetrators are mostly male. Those who are male should therefore be excluded from the space (on the grounds of safety) and trans women are considered, by those who offer this argument, to fall into this category. However, the claim that trans women are male is opposed by many (see for example Jenkins, 2016; Barnes, 2020).

accent). If such a case arose, the service would need to find a solution which addressed the needs of both users of the space (*ibid.*). Likewise, Manners argues, the exclusion of trans women cannot be justified by the concern that their presence could be trigger for other users of the space. Instead, should such a conflict occur, some other solution must be found which allows both to use the space. Furthermore, policies excluding trans women for such reasons would likely impact others, such as cis women who do not conform to societal standards of femininity (*ibid.*, p.8).

In women-only spaces more broadly, Manners cites the criticisms from authors such as Audre Lorde and Angela Davis that consciousness-raising groups should not assume homogeneity amongst women and their experiences of misogyny, since it would tend to privilege the voices of certain women – typically white, heterosexual, middle class women (see also Echols, 1989, pp. 89–91). Trans women's experiences may sometimes differ from those of cis women, but they still suffer oppression as a result of misogyny (Manners, 2019, p. 7). Therefore, a space created to allow women to speak about and come to understand their oppression should serve the needs of trans women just as much as it serves the needs of cis women.

0.3.2 'Homeplace': Spaces for sanctuary and respite

Another way in which a space may be safe is by offering sanctuary and respite – away from (as far as is possible) domination and oppression. While they have a less directly active aim, they are still sites of resistance because of the sanctuary and shelter from oppression which they offer for those who are constantly engaging in a battle for liberation.

Bell hooks (1990) speaks of the 'homeplace' as a site where black folk can shelter from racist oppression, where they matter and are not subjected to many forms of oppression and discrimination which they face in their everyday lives. This breathing, resting space can give people the chance to recover from their experiences, take stock amongst those who understand and share their experiences and build their strength back up.[16]

Whilst the homeplace is the paradigm which hooks describes, this kind of space can also exist within friendships and relationships.[17] Educator

[16] Further examples of this kind of space may include community groups which offer a form of extended support network – including religious or cultural organisations.

[17] Of course, a notable difference here is that whereas the homeplace generally involves a physical and private space, the spaces created among friends are metaphorical. More generally, a lot of the examples discussed here may not be fixed to a physical space, but

Elizabeth Ellsworth describes the way in which her students formed small informal affinity groups outside the classroom for the purpose of reflecting on the difficult conversation which took place in the classroom from a shared vantage point:

> Informal, overlapping affinity groups formed and met unofficially for the purpose of articulating and refining positions based on shared oppressions, ideological analyses, or interests. They shared grievances about the dynamics of the larger group and performed reality checks for each other. Because they were "unofficial" groups constituted on the spot in response to specific needs or simply as a result of casual encounters outside of the classroom, alliances could be shaped and reshaped as strategies in context.…
>
> They provided some participants with safer home bases from which they gained support, important understandings and a language for entering the larger classroom interactions each week. (1989, p. 317)

Unlike the spaces described in the previous section, these 'homeplaces' arise organically and informally amongst friends, family or community members who have things in common. Sometimes, though, spaces are created to help facilitate the formation of these community ties. For example, the creation of Black Cultural Centers on college campuses in the USA, beginning in the 1960s and 70s when the number of black students admitted to traditionally white institutions was increasing (Sanders, 2014, p. 31), aimed to create a space on campus where Black students could connect with each other and find support on majority white campuses. As well as organising formal activities, research and protests, these centres often acted as an informal community space – with common rooms and study spaces in which black students could find 'a place where they could talk to others and relate to them about having to put on a facade.' (Patton, 2006, pp. 643–644)

It is worth also noting that many spaces which are de facto separate were not created intentionally by the groups that they serve – many resulted from formal or informal segregation. Such spaces were not just safe spaces for marginalised groups (if they were safe at all), but were the only space available or open to them in a particular context. In *Teaching to Transgress*

---

rather they are spaces in the sense that they are venues for or offer opportunities for interactions among a community.

(1994), bell hooks recounts her experience of moving from segregated to desegregated schooling early in her education. She recalls her time at a segregated school with fondness, painting a picture of teachers who genuinely cared about their black students and gave them an empowering education. On the other hand, after desegregation, she describes being bussed across town to the 'white' schools where black students were treated with hostility and merely tolerated by teachers who took little interest in them (*ibid*., pp.2-4).

The relationship between these separate spaces and the broader aims of a liberatory movement can give rise to difficult tensions. On one hand, the continuation of separate spaces may further entrench segregation or may lead to less pressure to make other spaces accessible and welcoming to members of oppressed groups. But on the other hand, the expectation that members of oppressed groups give up separate spaces and enter spaces which are hostile for them or in which they face discrimination can be burdensome.

0.3.3 A space for independence and counterculture

Separate spaces also offer the opportunity for more freedom in scholarship and creative projects. The independence and autonomy offered by such spaces can allow for the development of epistemic practices and ways of theorising and meaning making which are marginalised or dismissed within the mainstream. For example, journalist Liv Little founded the online and print magazine *gal-dem* in 2015 – a platform for and by women of colour and non-binary people of colour. In response to the lack of representation within mainstream journalism, she felt that this space to express themselves through writing, art and photography, gave women and non-binary people of colour an autonomous space in which they met their own needs and created identities that need not be in relation to the white or male view-point. In an article for The Guardian, Little writes:

> 'Existing in white spaces can be alienating, and the creation of our site was a response to this. Having young women of colour from all over the globe reach out to us and express how much our work resonates with them is an incredible feeling. The expressions of black, intersectional feminism and womanism are integral parts of our self-care and we will not wait for the next wave of mainstream feminism to acknowledge our existence. We are here and have always been. Our experiences are just as important as any other.' (2016)

Whilst the work of Gal-Dem does have a wide audience extending beyond that of women and non-binary people of colour, the creation of the space for that particular demographic represents a refusal to make themselves and their narrative appealing and palatable to the mainstream media. This kind of space represents a form of safety against the hostility towards the voices of women and non-binary people of colour which can be found in mainstream journalism.

As well as the self-care and self-nurturing promoted by creating a separate space in which an oppressed group's needs are met, it can also function as a means of defiance. When a movement or group create for themselves a separate space in which they are the participants, audience and beneficiaries, they send out a message about the importance of the privileged gaze and participation. The language of the political lesbian movement reflects this:

> '[Our] time and space and energy is ours, and men would all too willingly steal it from us in any way possible… [T]aking time and energy away from them… means that we have more control over ourselves and our own lives… They simply cannot stand this. It drives them literally crazy not to get a reaction from us. Our reactions to them give them credulity and power.' (Leeds Revolutionary Feminist Group, 1981, p. 27)

By refusing to enact women's liberation in the male domain, political lesbians signalled the irrelevance and unimportance of men in women's liberation as well as their tendency to obstruct the purpose of the movement.

It is important to note a contrast between the 'separate' spaces described above and more radical *separatist* spaces. The spaces described above are created in response to a particular need. Given facts about the world and the ways in which oppression functions to control and silence members of oppressed groups, achieving liberation for (or even just meeting the needs of) members of those groups may require separate spaces. This is different from a stronger claim that the liberation of members of these groups necessarily involves separation and independence from others. This stronger claim can be seen embodied in the practices of lesbian separatist groups, such as The Furies and Radicalesbians, which broke away from the main women's liberation movement in the USA during the early 1970s. After feeling that lesbianism as a vital political choice was not being taken seriously by the women's movement, they required a space in which energy

would not be depleted in defending the legitimacy of lesbianism. The Furies initially created separate spaces in part to avoid antagonism and heterosexism within the radical feminist movement at large (Valk, 2002, p. 310). But this separatism was also motivated by the desire to build a women's counterculture and live and work entirely independently from men.

The belief was that, with men present (or in some cases, with heterosexual women present) such a radical creation would not be possible (Valk, 2002, p. 310). In *The Woman Identified Woman* (1970), the New York group Radicalesbians described how the female identity is created within male supremacy and conceived to be inherently less than human. To escape this and to find their true self and value women must create a new sense of self: 'As long as we cling to the idea of "being a woman," we will sense some conflict in that incipient self, the sense of I, the sense of a whole person.' (*ibid*., p.3). This new sense of identity must be created independently and not in relation to men. As a result of these ideological commitments, for some groups this required the absence of men (and sometimes heterosexual women) (*ibid*., p.303).

I mention these more radical separatist spaces to highlight the distinction between the exclusion of members of privileged groups because contingently liberational aims cannot be achieved without a space free from oppressive dynamics, and groups which conceive of their liberation as necessarily separate. Space with the latter kind of motivation is much harder to defend within education. Thus, it is important not to conflate the two. Whilst I will not defend at length any particular means of achieving safety (including separate spaces and the exclusion of, or limiting the participation of, members of privileged groups) in this thesis, I wish to flag that the motivation for such measures within education spaces lies in pragmatic needs for the achievement of liberatory aims and not in a separatist ideology.

The picture I have given is by no means exhaustive but merely serves to highlight several themes in the creation of spaces which are protected from the influence of oppressive norms, structures and practices in order to promote the achievement of liberatory goals. I mention the examples above partly to distinguish them from the kinds of safety which can be created within the classroom. Whilst the creation of safety in the classroom may serve similar liberatory aims, and to some extent the same methods may be used, the pedagogical aims of the classroom entail that it cannot, for instance, be a separate space for members of a particular marginalised group (in most cases at least) and is not always a space of rest. It is

important to recognise both that these kinds of spaces within activist and liberatory movements may inform much of the thinking behind the use of safety in the classroom to achieve liberatory goals, and that safety in the classroom is not equivalent to separate or sanctuary spaces.

0.4 Overview of the thesis

In chapter 1, I survey the existing literature on safety in the classroom. I compare a variety of different approaches and assumptions about safety to produce a taxonomy of various concepts. By doing this, I bring out some of the conceptual issues and ambiguities which are not always acknowledged and discussed within the literature.

I then defend the approach to defining safety which I will adopt in the remainder of the thesis. Drawing on the existing literature relating to safety in the classroom, I identify three approaches which can be taken to defining safety: defining safety in terms of (a) objective risk of harm, (b) subjective risk of harm, and (c) negative emotional responses. I argue that by defining safety in terms of objective risk of harm, we can account for a range of different ways in which a classroom can be made unsafe for students whilst avoiding the problems faced by the other two approaches. Finally, I propose that adopting this type of account of safety lends itself well to supporting the claim that safety is required for making the classroom accessible (a claim which I will defend further in chapter 5).

In chapters 2, 3 and 4, my focus turns to the relationship between epistemic character and protection. In chapter 2, I argue that a range of epistemic character traits can be developed as means of self-protection. Whilst some of these are epistemic virtues, some are not. A number of factors can influence the kind of epistemic character traits which a student develops as a means of self-protection, including: the costs involved in developing them; how the student's social context impacts the trajectory of their character development; and the extent to which different epistemic character traits are actually protective for them. I identify two protective strategies each of which are supported by (and thus encourage) the development of a cluster of epistemic vices. Then in chapters 3 and 4, I examine the costs which arise from their development of these vicious protective epistemic character traits.

In chapter 3, I argue that the development of vicious protective epistemic character traits is costly to students because it constitutes 'damage' to their epistemic character. By focusing on cases in which, given extreme hostile conditions, the development of epistemic vices makes an individual

epistemically better off than they would otherwise be, I identify a tension between the notion of epistemic character damage and the doctrine of normative contextualism. On the face of it, normative contextualism leads us to conclude that closed-mindedness is an epistemic virtue (for a particular agent) when it is a salient and effective means for that agent to respond to her particular epistemic predicament. One way it may do this, is if it is an effective protective epistemic character trait. However, I argue that this conclusion prevents us from identifying the development of closed-mindedness in these circumstances as epistemic character damage. I then propose a revised form of normative contextualism which avoids this tension. In doing so, I show that damage to epistemic character constitutes a change in epistemic character which takes an individual further from epistemic flourishing.

In chapter 4, I examine the constraint of epistemic agency through the case study of intellectual servility developed as part of the 'keep your head down' protective strategy. I identify two conditions of epistemic agency: the possession of reliable epistemic powers (including appropriate recognition within one's epistemic community) and the exercise of those powers. I argue that the exercise of epistemic powers requires certain self-regarding attitudes – particularly the recognition of oneself as a capable and worthy knower. I show that the relationship between intellectual servility and epistemic agency is complex but primarily it constrains epistemic agency through the entrenchment of negative self-regarding attitudes.

Finally, in chapter 5, I return to the theme of safety in the classroom by examining the relationship between safety and protection from harm. In this chapter I do two things. Firstly, I defend an account of safety in the classroom according to which a classroom is safe for some student when they are at a lower than tolerable level of risk of harm without having to protect themselves in ways that cannot be reasonably expected of them. Secondly, I argue that this form of safety is required in order to make the classroom accessible when educators lack detailed knowledge of students' abilities to take on high levels of risk of harm. I conclude this chapter by presenting two case studies in which I illustrate the implications of my account.

Chapter 1

Safety in the classroom: conceptual issues

Before developing an account of safety in the classroom and exploring the relationship between protection and epistemic character, I will begin with a brief survey of the existing literature on safety in education. Robert Boostrom (1998, p. 398) observed how the language of safety and safe spaces in the classroom became common even though the terms themselves had largely been neglected as a subject of educational inquiry. The terms were widely used, he argued, but were undertheorised – very rarely were explanations given as to what a safe space is or justification offered as to why creating a safe space is a good thing (*ibid.*).

If the concept of safety is becoming entrenched in our thinking, both in descriptions of how a classroom should be, and a criticism of the current trends in education, then there is important work to be done to establish what safety means in this context. Otherwise, we risk the concept being used in vague and polymorphous ways that make its practical application or debate about its value or lack of value in the classroom difficult. In particular, if the notion of safety in education has become entangled with larger, and often highly charged moral, cultural, and ideological issues, then care will need to be taken in analysing the concept.

Although there has been some definitional work relating to safety since the time of Boostrom's writing (for example, Holley and Steiner, 2005; Stengel, 2010; The Roestone Collective, 2014; Callan, 2016), it is striking that this work is not widely taken up beyond limited fields – such as philosophy of education and social work education. In this chapter, I will use this literature to draw out some important questions which must be answered when offering an account of safety in the classroom. In §1.1 I show that when interpreting student perspectives on safety it is important to distinguish between two kinds of testimony that they can offer: testimony regarding the nature of safety and testimony regarding the means by which safety is achieved. In §1.2 I then note the difference between objective and subjective risk and show how each can correspond to an account of safety

which appears to identify something valuable in the classroom. However, I argue that an account of safety defined in terms of objective risk can successfully account for both the value of a lack (or low level) of objective risk and the lack (or low level) of subjective risk in the classroom. Relatedly, in §1.3, I examine the relationship between safety and comfort and present a range of worries that making classrooms safe prevents critical discussion since such discussion requires a willingness to challenge one's deeply held beliefs or privilege – which for some students may be experienced as uncomfortable or even distressing. However, I argue that these worries can be avoided by adopting an account of safety based on objective risk of harm. Finally, in §1.4 I discuss two reasons for which safety may be valuable in the classroom – as a useful pedagogical tool or as a threshold condition for accessibility. I end this chapter by outlining another form of argument that safety in the classroom is not valuable – a position to which the remainder of the thesis responds. According to this argument, in safe classrooms students lack opportunities to develop the skills to cope with the kinds of adversity which they will inevitably face outside of the classroom.

1.1 Student perspective on safety: two types of testimony

The first issue which I will discuss is a distinction between questions regarding the nature of safety and questions regarding the means by which safety is achieved. Acknowledging this distinction is particularly important when interpreting students' perspectives on safety in the classroom. Existing work (Ellsworth, 1989; Winans, 2005; Arao and Clemens, 2013) examining the nature of safety in the classroom, often pays attention to students' perceptions and descriptions of safety. One often cited study from Lynn Holley and Sue Steiner in 2005 sets out to explore students' perceptions of safe and unsafe classrooms. There are many questions which can be raised about the extent to which student perspectives can be useful in understanding the concept of safety. For example, we might question whether students are well placed to give insight into the concept of safety, especially when they are not trained to think critically about the practice and aims of pedagogy.

In addition to questions about reliability, it is important to prise apart a distinction between two kinds of questions which we can ask (and correspondingly, two kinds of answers) regarding students' perceptions of safety. To illustrate, consider an anecdote from feminist educator and theorist Berenice Malka Fisher in which she reflects on her experiences of addressing the issue of safety in college classrooms in the USA. She

recounts an incident when she explicitly asked her students what 'safety' meant to them:

> 'After a go-around in which students talk about experiences with violence, Vera, a white and apparently middle-class woman, comments on how white, middle-class assumptions pervade several of the stories. Sarabeth, who is also white, becomes so agitated she can hardly stay in her seat. "I thought this was a safe space," she says, "that we were not supposed to judge each other." Vera leaps to her own defense: "I thought it was safe to say what we think!" I become anxious about my responsibility toward each of the students and toward the class as a whole. "What is 'safety' really about?" I wonder out loud. As we go around the class, Felice says that safety has to do with knowing that I as a teacher will not let things get "out of control." Lourdes comments that she always feels safer when people are speaking Spanish, as she does with her family and friends. Bruce remarks that as the only man in the room he would feel safer if there were "more of us." It is clear that "safety" does not mean the same thing to everyone.' (2000, p. 137)

Fisher asks her students a broad question – 'What is safety really about?' – and arrives at the conclusion that 'safety' means a different thing to each of her students. One way her question can be interpreted (and her reflection on the students' responses suggests that this is her interpretation) is as a question about the nature of safety. In other words, her question is about what it is to be safe. However, it is not so clear that we can interpret all of the students' answers in the same way. When they report that safety is about knowing that the teacher is in control or that they feel safer when Spanish is being spoken, they are describing conditions which create safety or make them feel safe.

We can distinguish between two different forms of relevant student testimony, corresponding to two different questions that can be asked about safety in the classroom. On the one hand, students can report on what makes them feel safe, what conditions they associate with safety or think are required for safety. This kind of student testimony corresponds to questions about *the means of achieving safety*: That is, questions about what is needed to achieve (the end) of safety in the classroom. On the other hand, students can also testify regarding how they understand the concept of safety itself. This form of testimony addresses questions about the nature of safety: what is the nature of the thing – safety – which we are trying to

achieve in the classroom? The worries noted above regarding the reliability and relevance of student testimony about safety can apply to both of these categories but may be particularly salient with regard to the latter as it concerns conceptual questions about the nature of safety.

It is possible that some or all of the students in Fisher's anecdote held the same understanding of the nature of safety and are reporting the conditions, which for them, appear most important and relevant to achieving that end. For instance, Felice, Lourdes and Bruce could all understand a safe classroom as one in which participation does not carry significant risk. The answers they give reflect the conditions which for them lower the stakes of participating in the classroom: knowing that the teacher will step in to defuse hostility; being able to speak in a language that they feel more comfortable and able to express themselves fully; and not feeling vulnerable to being singled out on the basis of their gender. The differences in the students' answers could reflect differences in their sensitivity and vulnerability to different sorts of risk.[18] Furthermore, we cannot assume that students themselves clearly distinguish the two questions. Their individual circumstances and vulnerabilities may make particular means of achieving safety especially salient to them and can motivate their conceptualisation of safety. For example, Felice, may have particular vulnerabilities or anxiety such that, the means required for her to feel safe is for her to sense that someone is fully in control of the situation. However, because she thinks of safety primarily in terms of control, she may begin to conceptualise the underlying *nature* of (not just the means of achieving) safety primarily in those terms. To assess these possibilities, we would need more information than Fisher gives about the students and their reasoning behind their answers.

Similar ambiguities can be found in Holley and Steiner's 2005 study investigating understandings and perceptions of safety in undergraduate and postgraduate level social work classrooms. Students were asked what they perceived to contribute to safe and unsafe classrooms. Among the

---

[18] Despite the worries noted regarding the unreliability or ambiguity of student testimonies, there may also be ways in which student testimonies regarding the means of achieving safety are more reliable than educator or third-person perspectives. In many contexts students are likely to be in a better epistemic position than educators to know and understand their own vulnerabilities and therefore may have better insight regarding what must be done to make the classroom safe *for them*. I will discuss students privileged epistemic position regarding their relationship to and ability to take on risk further in chapter 5.

factors which students listed as contributing to a safe classroom were: un-biased and non-judgmental students and instructors; respect for others' opinions, ideas and perspectives; an instructor who modelled good participation, set down ground rules and demonstrated caring; a sense of community amongst peers; a willingness from both students and instructors to share information and share their ideas honestly; and instructors who are comfortable with conflict and challenge students (Holley and Steiner, 2005, p. 56). The behaviours which contributed to a classroom being unsafe included: bias, closed-mindedness and judgement from either students or instructors; instructors' failure to consider other viewpoints; abruptness and rudeness from instructors; hostility, frustration and anger from students; feelings of worry, intimidation, insecurity and vulnerability; and a sense of apathy towards the course on the part of students (*ibid*., p.58). [19]

One of the stated aims of Holley and Steiner's study is to investigate the ways in which safe classrooms *are created* (*ibid*., p.50). However, Holley and Steiner also state that the study aims to investigate what a safe classroom *is* (*ibid*.), whilst also exploring the roles of instructors, students and physical classroom environments in 'creating an environment that encourages honest and open dialogue' (*ibid*.). By purporting to investigate the creation of environments that encourage honest and open dialogues, Holley and Steiner could be interpreted as making an assumption regarding the nature of safety. That is, they assume that for classroom to be safe is for it to encourage honest and open dialogue.

There are further methodological issues with the study which convey a background assumption about how safety is to be understood. As they report on their method of data collection, Holley and Steiner imply that the questionnaire used contained no explicit mention of 'safety'. Instead, questions were framed in terms of how willing and able students felt in raising controversial ideas or sharing personal experiences related to the course content. For example:

> 'The first section of the questionnaire asked students if they could think of a social work class they had taken in which they felt able and willing to raise controversial thoughts and opinions or share personal experiences that were related to

[19] Another noteworthy finding of the study was slight differences in the perceptions of safety between different groups of students, with men reporting more than women, and white students reporting more than students of colour, that they had been in safe classrooms (ibid., p.55).

> the course material. If they could think of such a class, they were asked to list up to six words or phrases that described what the instructor did, or the instructor's characteristics, that were important in creating this type of classroom environment. […]
>
> The third section of the questionnaire asked questions to determine the importance that students place on a safe classroom environment…. [T]hey were asked to rate the general importance of creating an environment where they could honestly express themselves. They were then asked to indicate… if such a classroom environment changed what they learned' (*ibid.*, p.53)

At most, the results of the questionnaire can tell us about the conditions that promote participants' feelings of safety when safety is understood as an environment in which students can honestly express themselves. There is no reason to assume that students would label this as safety. If the participants in Holley and Steiner's study are like the students in Fisher's classroom, we cannot assume that such a classroom would fall within their definition of safety. Therefore, we cannot draw any conclusions from the study about students' perceptions of whatever they understand as *safety* (see also Redmond, 2010, p. 4).

What these ambiguities demonstrate is the difficulty in interpreting student perspectives and the need to take care in distinguishing claims about the nature of safety on one hand, and the means of achieving it on the other. In chapter 5, I will attempt to answer the question: what is it for a classroom to be safe? That is, I offer an account of the nature of safety. In the following two sections, I will discuss three alternative approaches to defining safety: defining safety in terms of objective risk of harm; defining safety in terms of subjective risk of harm; and defining safety in terms of negative emotional responses. I will argue in favour of adopting the first of these approaches.

1.2 Being safe and feeling safe: objective and subjective risk

In this section, I will examine the difference between a classroom being unsafe because students are, as a matter of fact, at risk of harm and a classroom being unsafe because students perceive themselves to be at risk of harm. I will argue that this distinction is complicated by the fact that the feeling of being unsafe, even when the actual likelihood of students being harmed is low, can itself *cause* harm. Considering this, I defend an approach

in which safety is defined in terms of objective risk of harm. I will begin by illustrating the different approaches to defining safety through an example.

Elizabeth Ellsworth, an educational theorist working in the US higher education context, reflects on her experiences teaching an undergraduate course on Racism in the Media and the way in which her students challenged the assumptions of critical pedagogy. Ellsworth observes that her classroom, at least to begin with, was a not a safe space in which students felt able to speak openly and honestly about their experiences. She speculates about the reasons for this:

> 'Our classroom was not in fact a safe space for students to speak out or talk back about their experiences of oppression both inside and outside the classroom. […] Things were not being said for a number of reasons. These included fear of being misunderstood and/or disclosing too much and becoming too vulnerable; memories of bad experiences in other contexts of speaking out; resentment that other oppressions (sexism, heterosexism, fat oppression, classism, anti-Semitism) were being marginalized in the name of addressing racism – and guilt for feeling such resentment; confusion about levels of trust and commitment surrounding those who were allies to another group's struggle; resentment by some students of colour for feeling that they were expected to disclose "more" and once again take the burden of doing the pedagogical work of educating White students/professor about the consequences of White middle-class prejudice; and resentment by White students for feeling that they had to prove that they were not the enemy.' (1989, pp. 315–316)

Interpreting Ellsworth's comments here, we can separate out two different senses in which her classroom was unsafe. Firstly, her students believe, perhaps rightly or perhaps not, that should they speak, out they will face negative consequences – being misunderstood or becoming too vulnerable. They protect themselves by remaining silent. Should they speak out, they relinquish this barrier and so are forced into a state of vulnerability. Her classroom is unsafe *for speaking out* because there is no further protection that students can rely on when they relinquish their silence. Speaking out incurs too high a risk.

Secondly, the task of speaking out elicits a range of negative emotional responses from the students, including guilt, resentment, and confusion.[20] These emotional responses can be distressing or draining and use up mental resources. The classroom context, with its diverse students, low levels of trust and situation within a larger institution, enables or perhaps encourages the students to have these negative reactions when invited to speak openly and honestly. In this section, I will focus on the first of these two senses before returning to the second in §1.3.

The extent to which students are actually at risk of harm – what I will call *objective risk of harm* – and the extent to which they perceive themselves to be at risk of harm – what I will call *subjective risk of harm* – may come apart. For example, it is not difficult to imagine a scenario where students would not face any backlash were they to speak out. Their peers would not be hostile or, if they were, Ellsworth herself would be able to manage and diffuse the situation adequately. However, because of a lack of trust, the students still end up believing that this is not the case. In such a case, students are not at risk of harm but perceive themselves to be. On the other hand, it is also plausible that students could be at risk of hostile backlash when speaking out without being aware of this risk.

A parallel in the discussion of security is worth examining here. A common approach to security is to recognise at least part of its value as arising from freedom from fear. Emma Rothschild (1995, p. 62) defines security as comprising of two elements: freedom from the prospect of violation (where this is understood as the presence of particular kinds of threats (John, 2011, p. 71)) and freedom from fear. Rothschild assumes that the latter is a result of the former (*ibid.*). But such accounts face the problem that the subjective (freedom from fear) and objective (freedom from the prospect of violation) elements of security come apart (*ibid.*, p.72). It is possible for someone to be objectively secure in the sense that they are free from the prospect of violation, but still fear violation, or be subjectively secure in the sense of being free of fear but be unknowingly at risk of violation.

Furthermore, each element of security is valuable independently of the other. The objective element is valuable because of the reduced likelihood of the occurrence of harm. The subjective element can be shown to be valuable

[20] The burden which the students feel in educating the white students and educators is an example of what Nora Berenstain (2016) calls *epistemic exploitation*: the expectation for marginalised persons to undergo 'unrecognized, uncompensated, emotionally taxing, coerced epistemic labour' (ibid., p.57) in order to educate about and offer explanations of their oppression to those who are more privileged.

for a number of reasons including the disvalue of feeling fear (from a hedonistic perspective) or because an awareness of being insecure affects the ability to exercise agency because it prevents individuals from counting on, and planning on the basis of, their continued attainment of fundamental goods (*ibid*, p.71).

A similar point can be made in relation to the value of safety in the classroom. Given the distinction between objective and subjective risk there are two different conditions which are valuable in the classroom: a lack (or low level) of objective risk of harm; and a lack (or low level) of subjective risk of harm. Each of these is valuable independently of the other. A lack (or low level) of objective risk of harm is valuable due to the reduced likelihood that students will be harmed. On the flipside, a high level of objective risk of harm is disvaluable because of the increased likelihood of students being harmed. This value and disvalue is independent of students' subjective level of risk.

A high level of subjective risk of harm is disvaluable because of the effect which it has on students' ability to participate in and benefit from the learning activities in the classroom. In Ellsworth's classroom, the feeling of being at risk of harm prevented students from speaking openly and honestly. If students believe that there are at risk of harm then they may pre-emptively act to protect themselves, especially by withdrawing from activities or becoming defensive. If Ellsworth wants her students to be able to speak openly and honestly, then she needs to create a classroom environment in which students feel that they are safe.[21]

Since there are two different conditions of value in the classroom, we have three options when it comes to defining safety. Firstly, we could define

---

[21] Creating classrooms in which there is a lack of (or low level) of subjective risk of harm for students presents a wide range of challenges for educators. Firstly, there will be difficulties in predicting the threats and risks which students will perceive as very often this will be influenced by their past experiences or their particular vulnerabilities and anxieties. As a result, it can be incredibly difficult to design a strategy for achieving the subjective element of safety in the classroom because there is always the possibility that some students will not feel safe. This is of course a natural tension which arises between the particularities of individual's concerns and circumstances and the generality of policies and practices. Creating usable policies that offer concrete guidance to educators will inevitably involve some level of generalisation – if a policy responds to all possible student circumstances or concerns it will either be infinitely complex or abstract and dependent on educators' applications of the principles to individual scenarios. I explore the differences in students' perception and tolerance of risk further in chapter 5 and the difficulties of identifying and meeting different students' needs in the conclusion of the thesis.

safety in terms of objective risk of harm such that the term 'safety' captures the valuable condition of having a lack (or low level) of objective risk of harm. Let's call this an objective account of safety. Secondly, we could define safety in terms of subjective risk of harm such that 'safety' captures the valuable condition of having a lack (or low level) of subjective risk of harm: let's call this a subjective account of safety. Or thirdly, we could attempt to offer a 'unified' account of safety which allows the term 'safety' to refer both to a lack (or low level) of objective risk of harm and a lack (or low level) of subjective risk of harm.

I will defend the first option – an objective account of safety. I do this by arguing that it is not the case that a lack (or low level) of subjective risk of harm is always valuable in a sense properly described as safety. Furthermore, those cases in which it is valuable in a sense properly described as safety can be explained in terms of a lack (or low level) of objective risk of harm.

On the first of these points: the claim that a lack (or low level) of subjective risk should be understood as safety rests on the idea that 'S is safe' and 'S feels safe' are to some degree overlapping in their meaning – that is, that feeling safe is one way (if not the only way) of being safe. However, it is not clear that this thought is as plausible when we focus on safety as a property of the classroom rather than the individual. 'The classroom is safe for S' and 'S feels that the classroom is safe for her' are two different claims. The latter is an epistemic claim about S's perception of the classroom and the former is a claim about the nature of the classroom. Therefore, whilst a lack (or low level) of subjective safety may be valuable in the classroom in its own right, it is a mistake to assume that this value constitutes part of the value of safety.

There are *some* instances in which a high level of subjective risk is disvaluable in a way that seems appropriately described as a lack of safety. For example, imagine a student who believes without good reason that she will be subjected to harassment in the classroom and this fear causes her severe distress. She is not at risk of the harm of which she believes herself to be (the harm arising from harassment) but instead is at risk of harms arising from the distressing nature of her fear. Therefore, given the nature of her fear she is objectively at risk of harm.[22] In short, the high level of

[22] This does not entail that the appropriate response would be for an educator to try and reduce the risk of the student being harassed. Instead, the risk of harm would be reduced

subjective risk is disvaluable – and thus can be described as a lack of safety – because it leads to a high level of risk of harm. Therefore, the kind of disvalue which is appropriately described as a lack of safety can be explained purely in terms of a high level of objective risk. Therefore, safety can be defined in terms of objective risk of harm.

In the next section, I will compare an objective account of safety to one further approach – defining safety in terms of negative emotional responses – and argue that an objective account is able to explain the extent to which students' experiences of negative emotional responses in the classroom can constitute a lack of safety.

1.3 Safety, comfort and negative emotional responses

I now return to the second reason for which Ellsworth describes her classroom as unsafe for her students. When asked to speak openly and honestly, the students had a range of negative emotional responses – feelings of guilt or resentment, feeling burdened with the task of educating others, or anxious that they must 'prove they are not the enemy' (Ellsworth, 1989, p. 316).

Even if a student chooses not to speak openly and honestly and therefore feels protected from hostility, she may still encounter feelings of guilt and resentment. Such feelings can arise from the context of the classroom and the subterranean power structures and social patterns that operate within it. Simply being invited to speak openly about their experiences of oppression, in that context, may lead some students to engage in a costly mental process that involves weighing risks and benefits. Those students may, for instance, have to consider whether to and how to speak, who will benefit, who their testimony might reach and how it might be interpreted, while questioning whether it is fair that they take on the task of 'educating' white students – especially when doing so might open them up to harm and bring few, if any, positive effects. Furthermore, invitations can be experienced by those who are relatively powerless as demands when coming from those in positions of power. That is, whilst the educator assumes that students have the option to stay silent if speaking out is too risky for them, the students themselves may not perceive that they are genuinely being offered that option.

---

by addressing the unjustified fear itself or, if appropriate and possible, removing things which prompt the student to believe she is at risk of harassment.

An individual student may have this kind of response as a result of general or specific anxieties or worries. Perhaps they have had bad experiences in the past, or they are generally not trusting of others or their peers in particular. Perhaps they experience social anxiety and speaking in the classroom, especially speaking about their own experiences, heightens that anxiety. These factors seem largely, if not entirely, beyond the educator's control.

However, students may face these questions as a result of the set-up or context of the classroom. Take for example, the feeling of resentment from students of colour that they must educate their white peers about their experiences of racism whilst getting little benefit from the discussion themselves. The choice of whether to speak openly and honestly for students of colour is likely to be more fraught in classrooms in which many of their peers are relatively ignorant about the effects of racism and the educator does little to prevent ill-informed views from dominating the discussion.

The inclusion of students' experiences of discomfort and other negative emotional responses under the umbrella of things which constitute a lack of safety in the classroom has given rise to numerous worries that such a notion of safety detracts from students' learning and obstructs important pedagogical aims. For example, Brian Arao and Kristi Clemens (2013) recount an episode during which students were given an activity designed to demonstrate differences in privilege and opportunity. The students responded negatively to the activity and challenged its worth in the classroom. Those whose identities as members of dominant groups were made salient by the activity, 'stated that they felt persecuted, blamed, and negatively judged' (*ibid*., p.137). They reported feelings of guilt about their own positions as well as helplessness when recognising the lack of privilege and opportunity others experienced (*ibid*.).

In addition, the activity was challenged by those whose identities as members of oppressed and marginalised groups were made salient. They reported that the activity was a 'painful reminder of the oppression and marginalization they experience on a daily basis' and that they were placed in the role of educators for other students which caused them to feel angry and 'sorrowful' (*ibid*., p.137).

Arao and Clemens report the students making an explicit link between these responses to the activity and a lack of safety:

> 'Interestingly, a critique shared by many participants across target and agent group identities was that they experienced the activity as a violation of the safe space ground rules

> established with each participant at the outset of the module. The profound feelings of discomfort many of them experienced were, in their view, incongruent with the idea of safety.' (*ibid*.)

Arao and Clemens appear to accept their students' understanding of safety, comparing it with definitions from the literature which cast a safe space as '[a]n environment in which everyone feels comfortable expressing themselves and participating fully, without fear of attack, ridicule, or denial of experience.' (*ibid*., p.138). They conclude that such an understanding of safety is inconsistent with the challenging and risky task of engaging with and acknowledging one's own privilege or sharing experiences of oppression (*ibid*., pp.139-140) and for this reason Arao and Clemens question the value of safety when having honest discussions about issues of social justice (*ibid*., p.139).

It is worth separating out two distinct worries. The first concern relates to the use (and perhaps misuse) of calls for safety or allegations of a lack of safety by students. Calls for safety can be used by members of dominant groups as means to maintain their privilege by labelling uncomfortable or painful acknowledgements of their privilege as unsafe (Ludlow, 2004; Arao and Clemens, 2013, p. 140). Students may misidentify uncomfortable feelings and experiences with a lack of safety. Ludlow (2004; see also Winans, 2005) argues that feelings of safety can arise from assumptions of sameness. When students find this sameness challenged through the exploration of other identities, they may feel this as a decrease in safety (hooks, 1994, p. 82; Ludlow, 2004, p. 41). In some cases the classification of such feelings as a lack of safety may be used to silence or sanitise discussion which challenges or undermines privileged viewpoints.

Ludlow cites an example of one of her own students who objected to the idea that the classroom was a safe space because the class often discussed homosexuality and they 'hated homosexuals' (2004, p. 41). According to Ludlow, the student spoke out about the lack of safety as a form of protest against the disruption of privilege (*ibid*., p.45; see also Garran and Rasmussen, 2014, p. 407). For those in positions of privilege, there can be a lack of urgency and importance to the discussion, since it is not a vital step in reducing their own experiences of oppression. Therefore, with less to lose when the discussion is halted, members of privileged groups are in a strategic position of power to be able to set the norms for discussion. This threatens the ability of the classroom to be a place of liberation since members of oppressed groups are denied the ability to engage in discussion

on their own terms and instead must present themselves as 'safe' to members of privileged groups.[23]

Calls for safety can also provide students with a means to abdicate responsibility for the consequences of their speech. Being held accountable for the consequences of one's speech can be said to bring about negative emotional reactions such as guilt and embarrassment – which are said to be incongruent with safety. Recall Sarabeth, one of the students from Fisher's anecdote who said, "I thought this was a safe space… that we were not supposed to judge each other." Such an appeal to safety casts being 'judged' – that is, having her 'white, middle-class' assumptions pointed out – and the negative emotional response that it prompts as inimical to safety. In this way, Sarabeth seeks to avoid answering to criticism and challenges to her speech.

Low stakes are often important for encouraging students to take risks, particularly when discussing sensitive and complex topics such as oppression and social justice (Arao and Clemens, 2013, pp. 144–145). However, the notion that participants in a dialogue are entitled to distance themselves from the consequences of their speech overlooks the fact that students often have little control over how they feel – including how they respond to the speech of others. Students who are distressed, insulted or feel attacked by the speech of others often have no control over these responses. When a safe classroom is understood as one where students do not judge each other, such responses will still be felt but students will be prevented from expressing them. If they cannot be expressed, and therefore discussed and replied to, then the affected hearer is left to take the burden of dealing with their response silently (*ibid*., p.145). By classifying guilt and the feeling of being criticised as incongruent with safety, a speaker can absolve themselves of the responsibility to meaningfully reflect on and address the implications of their speech.

Given that my concern in this thesis is with the nature of safety in the classroom, I will largely put to one side the worries regarding the misuse of calls for safety. That being said, perhaps this practical issue can be addressed to some degree through more clarity and consensus on what it

---

[23] This need not always be done consciously and intentionally by members of privileged groups. Winans (2005) describes how her students at a rural, predominantly white college in the U.S. conflated being racist and acknowledging race. Therefore, their fear of being perceived to be racist motivated them to ignore race as a salient factor in their lives. Therefore, discussions in classroom which tackled issues of race could be experienced by these students as not safe from the risk of being identified as a racist (ibid., p.257).

means for a classroom to be safe. That is, it will be an advantage of an account if it can be used to demonstrate that such uses are misuses.

Secondly, there is the more fundamental concern that the kinds of negative emotional responses which are classed as unsafe cannot be excluded from a genuinely liberatory classroom. For example, as part of her 'pedagogy of discomfort' Megan Boler (1999) emphasises the need for students, as well as educators, to sit with and examine their emotional responses (including the difficult and painful ones) in order to uncover and challenge their deeply held beliefs, habits and values. What we see and pay attention to (and more importantly, what we do not see or pay attention to) is partly guided by emotions such as fear. Recognising what we do not see requires that we are willing to confront and sit with this fear (*ibid.*, p.182). Boler writes:

> The aim of discomfort is for each person, myself included, to explore beliefs and values; to examine when visual "habits" and emotional selectivity have become rigid and immune to flexibility; and to identify when and how our habits harm ourselves and others. (*ibid.*, p.185)

Boostrom (1998, p. 145) argues that the term 'safe space' when used to describe a classroom has the tendency to be conflated with a space which is without stress. Moreover, 'stress' itself is a vague concept which can incorporate both deep existential anxiety and vague discomfort. But such broad notion casts as unsafe a whole range of feelings which will inevitably arise if students are to be made aware of their own ignorance and pushed to critically reflect on their perspective. As Boostrom writes, 'Understood as the avoidance of stress, the 'safe space' metaphor drains from classroom life every impulse toward critical reflection.' (*ibid.*, p.406).

However, I will argue that these worries rest on the false assumption that students' experiences of *any* negative emotional responses have the potential to make the classroom unsafe. It is not clear that all kinds of negative emotional responses are incongruent with the classroom being safe. That is, describing a classroom in which students feel mild discomfort as 'unsafe' seems to be a misuse or non-literal use of the term. However, sometimes it does seem right to describe a classroom as unsafe because students experience significant negative emotional responses such as distress or severe discomfort.

For example, imagine a student who experiences some embarrassment and guilt when it is pointed out that he is ignorant of some of the ways in which women adapt their behaviour to protect themselves from male violence. This embarrassment and guilt may be to some degree unpleasant for the

student at the time but does not have any lasting negative impact on him (and hopefully has the positive effect of prompting him to reflect on his male privilege). Compare this to a student who feels humiliated and extremely guilty when she is aggressively 'called-out' by her peers for using incorrect terminology when discussing gender identity due to her lack of familiarity with the topic. This experience may be genuinely distressing and harmful to the student and have a lasting impact on her confidence in discussing this topic. In the former example, it is not appropriate to describe the classroom as unsafe on account of the embarrassment and guilt the student feels. However, in the latter example, it is plausible to describe the classroom as unsafe for the student since the effects of such a mistake and the emotional responses they cause result in genuine emotional harm to the student.

I propose that a suitable criterion for determining which negative emotional responses are rightly classified as making the classroom unsafe is offered by the objective account of safety discussed in the previous section. That is, only those negative emotional responses which cause harm to the individual have the potential to make the classroom unsafe. This is because the occurrence of harmful negative emotional responses leads to a high objective risk of harm to students. Those negative emotional responses which do not cause harm do not have the potential to make the classroom unsafe. Therefore, making a classroom safe only requires preventing negative emotional responses that cause harm. A safe classroom may still involve negative emotional responses such as discomfort or guilt so long as they do not cause students harm.[24]

Limiting the connection between negative emotional responses and a lack of safety in this way alleviates the second worry raised above. Even when the classroom is a place for critical discussion and challenging privilege, students should not be at genuine risk of harm. The kinds of negative emotional responses which cause harm should still be avoided. This is consistent with the implications of an objective account of safety. Negative emotional responses which cause students harm should be prevented in the classroom, but there is no need to prevent those which do not cause harm –

[24] It is possible that a negative emotional response of the same kind and severity could cause harm to one student but not another due to differences in students' emotional sensitivities and vulnerabilities. As will become clearer as I develop my account of safety more precisely in chapter 5, the safety of the classroom is always understood in relation to a particular student. It is possible that the very same conditions may make the classroom unsafe for one student but not another.

making possible critical discussion which challenges students' privilege and deeply held beliefs.

I have argued in favour of an account of safety defined in terms of objective risk of harm on the grounds that the subset of instances in which a high level of subjective risk of harm or negative emotional responses can appropriately be described as making the classroom unsafe can sufficiently be explained in terms of objective risk of harm. It is this approach which I will take up in chapter 5 when I present an account of safety in the classroom. In the final section of this chapter, I present one further argument that safety is not valuable in the classroom and will describe how I will respond to this argument in the remainder of the thesis.

1.4 The value of Safety in the Classroom

Intertwined with the question of exactly what we mean by safety in the classroom is the question of whether and why we think that it is valuable. To conclude my arguments in this chapter, I will now consider the implications of adopting an objective risk account of safety for the *value* of safety in the classroom.[25]

In §1.3, I responded to worries that safety in the classroom can obstruct the achievement of liberatory and pedagogical aims by arguing that when negative emotional responses that make the classroom unsafe are restricted to those which cause students harm, then the only things which are excluded from the classroom are those that would not be required by the liberatory and pedagogical aims in the classroom anyway. Even if it is important for students to address difficult questions and sit with discomfort in order to challenge their deeply held beliefs, as an educational setting, such challenges should stop short of causing harm to students.

However, there is a further argument a worth considering: safety (defined in terms of objective risk of harm) is not always valuable in the classroom because, by over-protecting students, it prevents them from developing the necessary skills to protect themselves from harm. According to a common line of thought within education, students' development should be guided in ways which will enable them to cope with many of the challenges they may encounter beyond the classroom. One way it may do this is by equipping students with the skills, capacities and resources that will build their *resilience*.

[25] Parts of this section have also appeared in Monypenny (2021).

Resilience arises in part from external conditions such as supportive environments and nurturing relationships, and in part from individual character, ways of thinking and behaviour (Russell, 2015).[26] These internal dimensions of resilience are emphasised within character education (see Earl and Arthur, no date; Walker, no date) and within educational psychologies inspired by Carol Dweck's (2006) influential growth mind-set model. Common to these approaches is a sense that students' character traits significantly affect the way in which they are disposed to respond to and cope with potential harms and adversity.

An interest in promoting resilience within education developed as a response to a growing focus on understanding the causes and impact of larger issues in terms of individual self-esteem – part of a wider cultural shift in the USA and UK sometimes termed the 'self-esteem movement' (Emler, 2001, pp. 2–3; Noble and McGrath, 2012, p. 611). The self-esteem movement had a marked impact on educational practices that included an increased focus on developing students' sense of achievement and self-expression in relatively non-competitive and failure-free learning environments (Noble and McGrath, 2012, p. 611).

Many critics were sceptical of the movement's view of students as inherently vulnerable and in need of professional therapeutic support in dealing with life's challenges (Furedi, 2003; Ecclestone, 2004, 2008). Educational environments in which students constantly receive praise and rarely, if ever, experience genuine failure, deprive them of valuable opportunities to develop a capacity to deal productively with frustration (Noble and McGrath, 2012, p. 611). Ecclestone (2004, p. 129) also criticised the movement for failing to address underlying structural inequalities when exclusion is treated purely as an individual psychological condition to be remedied by increased self-esteem.

It was against this backdrop that resilience education developed: a model of education – encapsulated by the Penn Resiliency Program in the early 2000s (see Freres *et al.*, 2002) – which aims to foster resilience by providing nurturing environments and cultivating individual skills and strengths (Gillham *et al.*, 2012, p. 660; Noble and McGrath, 2012). A large amount of this work occurred in the field of character education whose aim was to

[26] Alternatively, resilience could be understood as *constituted by* these (and other) factors, rather than *arising from* them. This would lead to a more interpersonal and socially embedded notion of resilience. I do not explore this approach to resilience here since I am interested in responding to approaches which treat it as a quality or trait of the individual.

develop students' character by facilitating the development of various excellences whilst mitigating or removing negative traits. The role of education is taken to be in part the development of students' agency so that they can overcome adversity and achieve success. This will not be done, proponents of resilience education argue, merely by improving self-esteem.

These concerns about preventing students from developing the skills to face adversity can be found echoed in other critiques of safety in education. For example, Catharine Ecclestone (2004) argues that the treatment of students as being 'at risk' and in need of protection encourages a diminished view of autonomy and can influence students to view themselves as vulnerable and fragile – discouraging the development of resilience. In a similar vein, Jonathan Haidt and Greg Lukianoff (2019) compare education to a form of 'exposure therapy'[27] in which students should learn positive psychological habits (inspired heavily by cognitive behavioural therapy) which equip them to overcome adversity and hardship. Excessive safety in education, Haidt and Lukianoff argue, leads students to view themselves as fragile and as a result become less resilient.[28]

One character profile through which resilience has been exemplified and valorised in the social imaginary is the gritty individual. Grittiness is an individual quality defined as 'working strenuously towards goals, maintaining effort and interest over years despite failure, adversity and plateaus in progress.' (Duckworth *et al*., 2007, pp.1087-1088; see also Stokas, 2015).

However, the focus on resilience and grit within education has been criticised on the grounds that it diverts attention and effort away from addressing structural inequalities and systemic injustices. Ariana Gonzalez Stokas (2015, p. 522) argues that grit is often needed when critical support for success is missing. Focus on grit can convey to students that hostile environments and failures in support are to be tolerated and that they should succeed despite structural issues remaining unaddressed (ibid., pp. 521–522). Arthur *et al*. (2016) find it telling that Duckworth suggests that

---

[27] It is not at all evident that what Haidt and Lukianoff describe as 'exposure therapy' is actually reflective of the therapeutic practice.

[28] It interesting to note that whilst there are similarities in their critiques of safety in education, Haidt and Lukianoff and Ecclestone have vastly differing views on the role of therapy within education. Ecclestone is highly critical of the influence of therapeutic practices and focus on self-esteem within education.

disadvantaged or underprivileged students in particular will need and benefit from education focused on the development of grit. They note that this starkly contrasts historical approaches to character development which saw the privileged and comfortable as the least experienced in enduring hardship and most lacking in grit. Students who are structurally disadvantaged most likely have already learned the need to endure adversity in education and other aspects of their lives (Stokas, 2015, p. 522).

These worries can be summarised in the following way: it is unnecessary, and perhaps damaging, to make classrooms safe for students since students should either possess or develop the skills and resources to protect themselves from harm. The remainder of this thesis will focus on responding to this claim. I will argue that we cannot assume that all students will be able to develop means to protect themselves from harm without exposing themselves to further harm or putting themselves at a significant disadvantage. Therefore, a lack a safety in the classroom leads to an unequal classroom – some students will be disadvantaged or exposed to further harm by the need to protect themselves. Rather than being a useful (or not) tool in the classroom, safety is instead a fundamental condition required for the classroom to be open and accessible to students.

One such view – of safety as a baseline condition for access – is defended by philosopher of education Eamonn Callan (2016) who argues that whilst intellectual safety – being protected from having one's unquestioned beliefs challenged – is not always conducive to good learning, another more fundamental kind of safety is required. To be 'dignity safe' is, 'to be free from any reasonable anxiety that others will treat one as having inferior social rank to theirs.' (*ibid.*, p.65). Dignity safety allows students in the classroom to participate in the knowledge that they will be treated with equal respect and not be humiliated. This, Callan argues, is a threshold condition which is generally taken for granted in the classroom (*ibid.*, p.68).

Sigal Ben-Porath (2016, p. 79) builds on Callan's account and argues that dignity safety is an issue of access. A lack of dignity safety can, wholly or partially, prevent students from being able to fully take advantage of learning opportunities (*ibid.*, p.80). A student who faces constant attacks on their dignity safety – by witnessing physical and verbal assaults on people like them or feeling vulnerable to violence and abuse, or being surrounded by messages that people like them do not belong in the institution or have equal status in the community – faces the wearing task of constantly responding to and resisting threats. This can be extremely draining and,

especially when it goes unacknowledged, can direct attention away from full and fruitful engagement in learning.

In contrast to Boostrom's concerns about safety being conflated with a lack of stress, Angela Carter (2015; see also Kessler, 2018, pp. 144–145) argues that the language of safety is important in characterising these kinds of access requirements because it emphasises the centrality of these issues to allowing all students enter the classroom. It demonstrates how such tools and practice are necessary because without them students are in a state of unsafety. They are not just preferences or things which make students feel more comfortable.

One example of this is the use of content warnings,[29] which are often cited as a being a tool for achieving safety in the classroom. Carter (*ibid.*) argues that content warnings are best understood as tools for access. Rather than creating a tendency to opt out, they enable students to opt into difficult topics or topics which personally affect them because they provide prior warning. For example, if a student has experienced sexual assault and graphic descriptions of sexual assault can trigger distressing anxiety or PTSD symptoms, the prospect of taking a feminist philosophy module where it is likely that the topic of sexual assault will be raised may be difficult for her – she knows that if the topic is brought up without prior warning a distressing response may be triggered. She may choose to avoid this risk or attend the module but be on edge. On the other hand, if she is confident that (where possible) prior warning will be given when this topic will be discussed or is likely to come up, she will be able to either avoid just those instances or will have the chance to mentally and emotionally prepare herself.

Here, I have offered some brief arguments that safety should be understood as a baseline access condition in the classroom. I will develop this point more extensively in chapter 5 and argue in more detail for the claim that, in the absence of highly detailed educator knowledge of a particular

[29] Content warnings (also referred to as content notes or trigger warnings) have come under fire as features of some types of safe space. In 2016, University of Chicago Dean of Students sent a letter to the cohort of new students informing them that the university's 'commitment to academic freedom means that we do not support so called "trigger warnings"… and we do not condone the creation of intellectual "safe spaces" where individuals can retreat from ideas and perspectives at odds with their own.' (Ellison, 2020). I will discuss the use of content warnings further in §5.3.1.

students' ability to take on risk of harm, making the classroom accessible for that student requires making it safe for them.

1.5 Summary

In this chapter, I have surveyed the existing literature on safety in the classroom and in doing so, identified a number of different approaches to defining safety. I argued that by defining safety in terms of objective risk of harm, we can account for a range of different ways in which a classroom can be made unsafe for students whilst avoiding the problems faced by accounts which define safety in terms of subjective risk of harm or negative emotional responses. Finally, I proposed that adopting this kind of account of safety lends itself well to supporting the claim that safety is required for making the classroom accessible (a claim which I will defend further in chapter 5).

This sets the groundwork for the overall task of this thesis – offering an account of safety in the classroom. I will proceed with this task in a roundabout way by first exploring the connection between epistemic character and the ability of students to protect themselves from harm.

Chapter 2

Epistemic character and protection from harm

So far, I have presented several approaches to defining safety and argued in favour of an objective approach – the approach which I will adopt in the remainder of this thesis. Before developing my account of safety in the classroom in chapter 5, the next three chapters examine the impact on students when they are not protected from harm by educators or educational institutions and must protect themselves (if they are able to do so). I will argue that, whilst in many circumstances, students may possess or be able to develop skills and traits which allow them to protect themselves effectively from harm, sometimes these self-protective measures will carry significant costs. This will form a basis for my argument in chapter 5 that a classroom is safe only when a student is not at a higher than tolerable level of risk of harm *without being required to protect herself from harm in ways that cannot reasonably be expected of her*.

In this chapter, I will explain the link between safety in the classroom and epistemic character. I will begin in §2.1 with a brief introduction to character epistemology and, in particular, critical character epistemology. Critical character epistemology is concerned with understanding how the development of epistemic character is shaped by the norms, power structures and practices of social environments (Kidd, 2020). I adopt this approach to epistemic character theorising by examining the ways in which exposure to risk of harm influences the development of epistemic character. In §2.3, I identify a number of epistemic character traits which can be developed as means of self-protection. I call these *protective epistemic character traits* (PECTs). I argue that PECTs can be either epistemic virtues or epistemic vices and (in §2.4) identify a number of factors which may influence students to develop vicious rather than virtuous PECTs.[30]

[30] A version of this chapter has also been published as Monypenny (2021)

2.1 Character epistemology and education

Character epistemology,[31] which developed first as virtue epistemology (for example Code, 1984; Sosa, 1985; Montmarquet, 1993; Zagzebski, 1996) and later as vice epistemology (for example Battaly, 2014; Cassam, 2019; Tanesini, 2021) subscribes to the general ontological thesis that epistemic agents have an epistemic character consisting of a set of epistemic dispositions[32] that guide the way we engage in epistemic activities. Being a knower and part of an epistemic community is active and dynamic (Baehr, 2011, p. 1) – it involves a range of different activities including: forming and holding beliefs, learning, discovering, investigating, teaching, listening, discussing, evaluating and many more. There are different ways in which these activities can be performed (both in terms of the way in which a particular activity is performed and in terms of which ones are performed when). How well we do in our epistemic projects is influenced by how we go about these activities (though it is an important question as to how significant the role of character is and how much external factors or matters of luck determine the outcomes of epistemic projects).

Epistemic virtues are excellences of epistemic character – dispositions which make us excellent epistemic agents. Paradigm examples include intellectual honesty, open-mindedness and intellectual perseverance. On the other hand, epistemic vices are flaws or failings of epistemic character[33] and often discussed examples include intellectual arrogance and closed-mindedness. Exactly how 'excellence' and 'failing' are understood varies. There are two main views within character epistemology – consequentialism and responsibilism. Consequentialists define virtues and vices in terms of their consequences. For example, for Sosa, 'an intellectual virtue is a quality bound to help maximise one's surplus of truth over error' (1985, p. 227).[34] As a result, epistemic virtues and vices are fairly broad

[31] I use this term, following Baehr (2006) and Kidd (2020, forthcoming) to refer to the combined fields of virtue and vice epistemology.

[32] There is some disagreement over what kind of things are correctly identified as virtues or vices. For example, Cassam (2019) adopts a 'pluralist' account of vice which allows that vices can include character traits, ways of thinking vices and attitudes.

[33] Crerar (2018) notes that virtue epistemologists have tended to assume that epistemic vices are just the mirror image of virtues and presents a challenge to this assumption.

[34] Rather than truth, others define virtues and vices in terms of their tendency to produce other forms of epistemic good, including knowledge (Cassam, 2019, p. 12) and 'cognitive contact with reality' (Zagzebski, 1996).

categories incorporating so-called 'hard-wired' traits such as good eyesight or poor memory.

On the other hand, virtue and vice responsibilists restrict the categories 'virtue' and 'vice' to those traits for which individuals are deserving of praise and blame (because they are responsible for those traits in the required way).[35] For responsibilists, traits such as good eyesight and poor memory would not count as epistemic virtues and vices, respectively, since they are not traits for which an individual can be praised or blamed. The most common form of responsibilism is motivationalism – the requirement that epistemic virtue requires a certain kind of motivation – usually understood as a love of truth, or some more specific motivation derived from a love of truth (for example Zagzebski, 1996; Baehr, 2011).[36] These couplings of virtue with praise and vice with blame have been challenged by critical character theorists on the grounds that in some instances, other evaluative responses are more appropriate – I will return to this point in my discussion of normative contextualism in chapter 3.

Recently, Heather Battaly (2019) proposed a third account – personalism – with regard to epistemic vices: traits for which the individual need not be wholly responsible but are still to some degree aspects of their character. This account is intended to capture traits which individuals develop as a result of indoctrination or upbringing (and therefore cannot wholly be held responsible for) but are still worthy of criticism (though are not appropriate candidates for blame).[37]

To be epistemic virtues or vices, traits must be stable aspects of epistemic character – over time and across different situations. To act in an open-minded way on a few occasions is not enough to possess the virtue of open-mindedness. Having the virtue requires having a stable disposition which consistently guides behaviour in a range of different circumstances.[38] For a

---

[35] Some responsibilists, such as Zagaebski (1996) and Cassam (2019), also require that virtues and vices have good and bad consequences, respectively.

[36] The strong motivational requirement has been challenged by, among others, Orozco (2013) and Crerar (2021)

[37] One of the motivations for the development of this view is that vice responsibilism rules out too many traits that we would want to consider vices on the grounds that individuals cannot genuinely be held responsible for their developement (Battaly, 2019, p. 121).

[38] Although exactly how consistently a disposition must guide behaviour (and therefore, how much inconsistent behaviour is consistent with possessing the trait) can vary depending on the trait. Alfano (2013, p. 241) distinguishes between virtues and vices which are high fidelity (and require that the individual almost always act in accordance with the

disposition to become a stable part of someone's character, it must be practiced. It is here that there is a natural connection between character epistemology and education (see for example Battaly, 2013; Baehr, 2016; Kotzee, 2017; Curren, 2019). A significant aim of education is to develop students into good knowers and good participants in intellectual inquiry. If being a good knower involves developing good stable traits of epistemic character, then helping students to become good knowers will involve providing them with opportunities to exercise and practice epistemically virtuous behaviour so that these behaviours will develop into stable dispositions.

However, epistemic character does not develop in a vacuum. An individual's circumstances, the challenges they face, the resources and opportunities they have available to them will impact the development of their epistemic character. Kidd (2020) argues that theorising about epistemic character should take seriously this fact and pay attention to the ways in which social structures, practices and institutions shape the development of character. The kind of theorising which Kidd describes (and terms 'critical character epistemology' (2020, forthcoming)) is largely influenced by Robin Dillon's (2012) proposals for 'critical character theory'. Dillon challenges some of the assumptions made in traditional virtue ethics - such as the assumption that if a trait is a virtue, it is a virtue for all agents in all contexts. She argues for character-focused theorising that seeks to understand 'how character is shaped by, supports, and resists domination and oppression' (*ibid.*, p.83).

Kidd (2020, pp. 77–78) notes several central features of Dillon's critical character theory, two of which are important for this thesis. Firstly, when thinking about individual character, critical character theory takes into consideration how those individuals are situated within social practices and institutions which contain within them relationships of power and oppression. Secondly, with roots in critical theory (Dillon, 2012, p. 85), it has liberatory aims. Dillon argues that oppression can operate on and through character (ibid.; see also Card, 1996; Tessman, 2001) and therefore, critical character theory aims to uncover this oppression and work towards amelioratory goals.

In the epistemic domain, critical character epistemology examines the relationship between epistemic character and social oppression. Different epistemic agents have different *epistemic predicaments* (Medina, 2013, p.

---

disposition) and others which are low fidelity (and can tolerate more behaviour that is not in accordance with the trait).

28; Kidd, 2020, pp. 73–4): each agent is uniquely socially and epistemically situated, facing different epistemic challenges and having access to different epistemic opportunities and resources.

To be effective in developing students to be good knowers, education must focus on not just promoting the development of epistemic virtues, but also avoiding or undoing the development of epistemic vices. Doing this effectively, requires an understanding of the conditions or factors which hinder the development of virtues and promote the development of vices. That is, the conditions or factors which are *epistemically corrupting* (Kidd, 2019a, 2020, 2021) environments, structures, agents, practices, or institutions which promote the development of vice and hinder the development of virtue. (I will discuss epistemic corruption further in chapter 3). Effectively equipping students to overcome and respond to the challenges they face will require aiding students in understanding their particular predicaments and which character profiles will allow them to thrive in their circumstances.

The commitment of critical character epistemology to uncovering the complexities of epistemic character and the context-sensitivity of the value of many traits can come into conflict with the practical aims of education. For example, designing curricula which promote the development of epistemic virtues is made considerably more difficult if it must take into consideration how the unique predicament of each student will affect their epistemic character development. Therefore, critical character epistemology may cast doubt on the tenability of traditional approaches to character education. I will return to this point in §2.5.

Another feature of conventional character theory which Dillon notes is the tendency to neglect the negative aspects of and influences on character. She notes how whilst there is a lot of work on moral virtues, there is comparatively little work on vices. Similarly, vice epistemology is a relatively new field in comparison to virtue epistemology. Critical character epistemology involves a somewhat more pessimistic outlook on the epistemic world than virtue epistemology[39] and pays attention to the harms, barriers and dangers that impact epistemic character development. I will focus on some of the harms that may arise when stereotypes, biases, prejudices, and discrimination infiltrate the classroom environment.

[39] In recent years there has also been a large amount of work on the vices of public discourse (see for example Lynch, 2018; Kidd, 2021; Watson, 2021). This work is in many ways also pessimistic about the general levels of virtues and vice in our epistemic communities.

2.2 Risks of harm in hostile classrooms

Now that I have given a brief introduction to critical character epistemology, I'll begin by describing some harms which can arise in the classroom as a result of what might be called a 'hostile environment'[40]. The harms which I describe will not arise in all classrooms, but may arise, in a variety of combinations, in many different classrooms for many different students. In §2.3, I will then offer examples of four different protective strategies (and the epistemic character traits[41] which support them) which students might adopt to protect themselves against the harms described below.

Consider the following case:

> Elena is in a discussion-based class and finds the climate within the classroom to be extremely hostile. She faces constant hostility, ridicule and belittlement from her peers. Her contributions are dismissed as unintelligent or irrational because she is unfairly perceived not to be knowledgeable about the topic, and incapable of making relevant or interesting contributions. She also comes up against viewpoints which challenge her equal status within the classroom on the basis of her identity; for example, that women are prone to emotional reasoning and therefore are less reliable in reasoning objectively.

There are a range of different ways in which Elena could potentially be harmed in this environment, including a range of entangled psychological, social and epistemic harms. Possible *psychological harms* include repeated exposure to degrading and discriminatory treatment which can affect her self-trust, self-esteem and self-respect and can cause what Sandra Bartky (1990) terms 'psychological oppression': the internalisation of discriminatory and prejudicial attitudes which cast one as inferior. As a result, her sense of and belief in her own abilities, potential and worth may

---

[40] Battaly uses the term 'epistemically hostile environments' to refer to environments which are 'extremely polluted' with information, informants, and sources of information which are 'false, unreliable, or aimed at misdirection' (2018, p. 39). I will consider cases of the type discussed by Battaly in chapter 3, but here I use the term more loosely to refer to environments in which individuals are treated with hostility in ways that are unconducive to their attainment of epistemic goods and positive epistemic development.

[41] I focus on epistemic character here but there may also be a similar relationship between moral character and protection.

be eroded – she may come to see herself as less capable than her peers, lower her aspirations, and accept discriminatory treatment.[42]

She could also suffer a range of *social harms* – including exclusion, isolation and being treated as lacking equal status in the classroom. Finally, she may also suffer harms which are more distinctly *epistemic*. For instance, as a result of a decreased level of trust in her abilities and potential, she may lose confidence in her beliefs, possibly losing some of her true beliefs. In turn, her epistemic development may also be restricted: she may not attempt more demanding epistemic tasks because she does not believe that she will succeed. Furthermore, the discriminatory treatment of her by her peers may deny her opportunities to develop and practice epistemic skills by not allowing her to take part in discussion as an equal participant.[43]

Whether, and how much, Elena will suffer these and other harms as a result of the hostile climate in the classroom will depend in part on whether her peers or the educator intervene to prevent or respond to these harms. If they do not, then whether she suffers these harms will depend on her own ability to predict, avoid, respond to or mitigate harms. This might be influenced by a range of factors including the resources and support networks she has available to her, and the skills and dispositions that she has developed. In particular, I will focus on the way in which Elena responds to her environment, and the harms she encounters as a result, is impacted by her epistemic character. As I will argue below, which of the many different epistemic character traits are relevant and important to her at a given time will depend on her specific interests, concerns, challenges, and needs—what José Medina calls her 'epistemic predicament' (2013, p. 28).

2.3 Protective Epistemic Character Traits

There are a number of different strategies which students like Elena might adopt to protect themselves from the harms described above. I will discuss four strategies below (although there may be many more): 'battle on through', 'believe in yourself', 'just don't listen', and 'keep your head down'. Each protective strategy can be made more effective through the possession

---

[42] See also Hookway (2010) for an example of how this may occur.

[43] As an example, see Fricker's (2007, pp. 47–48) argument that repeated instances of testimonial injustice can have negative impacts on epistemic development. The effects of oppression on the development of character are also discussed by, among others, Du Bois (2015)and Césaire (1972).

and exercise of *protective epistemic character traits* (PECTs): epistemic character traits which are developed and/or exercised as a means of enabling or disposing an individual to detect, assess or respond to actual or anticipated sources of harm in ways that tend to prevent, limit or mitigate those harms. For each of the protective strategies, I identify two related PECTs[44]. The first two strategies presented are supported by traits which are commonly valued in resilience and character education.[45]

2.3.1 Strategy 1: 'Battle on through'

The first protective strategy which students might adopt is to persistently battle through adversity and not let challenges hold them back from achieving important goals. This can be made easier through the possession and exercise of *intellectual courage* and *intellectual perseverance.* Suppose that Elena developed and exercised *intellectual courage* – roughly, the dispositions to correctly perceive potential harms and appropriately persist in pursuing epistemic goals despite potential harms or setbacks (Roberts and Wood, 2007; Battaly, 2017; Kidd, 2019b). Intellectual courage and perseverance can be protective for Elena because they help her to avoid and mitigate the long-term negative effects of repeated exclusion from discussion and dismissal on her learning and wider epistemic development because they motivate her to keep trying until she is taken seriously. It can also mitigate the harms of individual setbacks and dismissals by motivating her to find other opportunities for participation and engagement in discussion.

---

[44] My methodology is somewhat similar to that used by Baehr (2011) in his identification of six 'natural groupings' of epistemic virtues. Each grouping relates to a distinct 'inquiry-relevant' challenge and contains virtues which equip individuals to respond to and overcome these challenges. Like Baehr, I do not take the grouping I present to be a strict classification or taxonomy – they need not correspond to any structural connections or distinctions between epistemic character traits. The groupings of PECTs need not be considered the primary way of categorising epistemic character traits – it is simply the salient way of grouping them when thinking about epistemic character in the context of protection. There may be other ways of grouping epistemic character traits together which intersect with or are more fine-grained than the groupings offered here.

[45] For example, The Jubilee Centre for Character and Virtues (2012, p. 4) lists courage, confidence, determination, perseverance and resilience among the 'building blocks of character'.

2.3.2 Strategy 2: 'Believe in yourself'

PECTs such as *intellectual confidence* – the appropriate grasp of the strength of one's epistemic position and abilities (W. E. Jones, 2012, p. 697) – and *proper pride* in one's intellectual achievements (Tanesini, 2018, p. 26) aid self-protection by disposing students to have a robust sense of and trust in their own abilities and worth. These traits help some students to self-protect against psychological harms such as damage to self-esteem or loss of self-trust by disposing them to be appropriately[46] resistant to the negative messages about their competency and self-worth conveyed by others. By cultivating and exercising the virtues of intellectual confidence and proper pride, Elena may be less likely to take to heart the challenges and smears on her intelligence because she is disposed to appropriately trust her own intellectual abilities. Although she may concede to some of the challenges she encounters in the hostile classroom, a positive attitude towards her epistemic position and abilities will allow her to view these challenges proportionately and avoid taking them as signs of her inherent or irredeemable incompetency (see also Battaly, 2021).

I now turn to two further protective strategies which are supported by traits which, far from being included on standard lists of epistemic virtues taught in the classroom, are conventionally classified as epistemic vices.[47] These epistemic vices can have some benefit for their bearer under certain conditions.

2.3.3 Strategy 3: 'Just don't listen'

One way to respond to hostility, degradation or insults from others is to block them out. Battaly (2018) argues that *closed-mindedness* (an inability or unwillingness to engage with relevant intellectual alternatives) and *dogmatism* (an unwillingness to engage with alternatives to a belief one

---

[46] I include 'appropriate' because, the virtues of intellectual confidence, or proper pride, the attitudes should be to some degree well-calibrated to one's actual competency. The intellectually confident or properly proud individual's confidence in their own abilities would be properly responsive to genuine criticisms (see for example Kidd, 2016; Tanesini, 2018, pp. 25–26).

[47] As I will go on to discuss in chapter 3, normative contextualism challenges the acontextual classification of these traits as vices given the value they have for certain individuals in certain contexts. In this chapter, I use 'virtue' and 'vice' to mean traits conventionally (that is, by traditional virtue epistemology) recognised as virtues and traits conventionally recognised as vices respectively. I will return to this terminological point in chapter 3.

already holds) can be beneficial for individuals in certain epistemic environments heavily polluted with misinformation or propaganda. An open-minded individual in such an environment risks being taken in by some of the misinformation around them. Because it is difficult to differentiate between reliable information and misinformation, they give consideration to any pieces of misinformation they come across. At worst, they may be convinced by some of the misinformation and replace some true beliefs with false ones. Even if they were to avoid this, they would have to spend a lot of time and effort separating out reliable information from misinformation (*ibid.*, pp. 38-40).

On the other hand, a closed-minded individual does not waste time sifting through and giving consideration to misinformation – they dismiss it out of hand because they have a tendency not to give all relevant information due consideration. Because of this, Battaly argues that within environments rife with misinformation, closed-mindedness at the very least *minimises* bad epistemic effects for the individual (ibid, p. 41).

Battaly also notes how closed-mindedness can help members of marginalised groups avoid losing confidence in their intellectual abilities and overemphasising their limitations when their credibility is constantly doubted (*ibid.* p.40). If Elena becomes closed-minded, she will be disposed to dismiss the views of her peers when they challenge her competency, especially those views which deny her equal status in the discussion. In this way closed-mindedness can be a PECT for Elena as it allows her to avoid some of the potential psychological and epistemic harms which could arise from taking to heart those claims.

It could be objected that the trait which Elena develops and exercises is not actually closed-mindedness. Being open-minded does not require giving consideration to *all* intellectual alternatives. For example, in an astrophysics classroom, the teacher and students need not give consideration to the theory that the earth is flat. Being open-minded only requires giving appropriate consideration to *relevant* intellectual alternatives (Battaly, 2018, 2020). Therefore, it is not the case that any failure to give consideration to an intellectual alternative constitutes closed-minded behaviour – only the failure to give consideration to *relevant* intellectual options.

Therefore, if the views which Elena dismisses and fails to take seriously are not relevant intellectual options then she is not exercising closed-mindedness by dismissing them. Battaly (2018, p. 39) argues that one of the grounds by which an intellectual option is relevant is its prevalence within

the epistemic community. She argues that if a view is widely adopted within an epistemic community then it is due consideration. Therefore, the open-minded person would not dismiss it out of hand. In Elena's case, the view that she is not academically capable is widely shared amongst her peers. If it was just endorsed by a small but vocal minority then we might be inclined to say that it is not a relevant intellectual alternative, and Elena is not displaying closed-mindedness by dismissing it. However, in cases in which the view is widely shared, perhaps throughout the institution as a whole because of systematic prejudice against people from Elena's social group, it would be considered a relevant intellectual alternative. Therefore, by dismissing it out of hand and not giving it any consideration or attempting to refute it, Elena would be displaying closed-mindedness. It suffices for my point here that, in some cases, the trait developed as a means of self-protection is closed-mindedness, that is, it is enough to show that closed-mindedness *can* be developed as a PECT.

2.3.4 Strategy 4: 'Keep your head down'

The final strategy involves students protecting themselves by avoiding conflict and placating those who are hostile towards them. One way in which students may adopt this strategy is through the development of the epistemic vices of *intellectual servility* and *intellectual timidity*.

Alessandra Tanesini (2018, 2021) argues that intellectual servility (excessive deference to others rooted in a lack of pride in one's intellectual achievements, feelings of shame and a need for others' approval) can be developed as a survival strategy for members of oppressed groups. When prevailing stereotypes and expectations state that one should be quiet and deferent, then living up to these expectations can allow one to gain some acceptance from those in positions of power. As Tanesini describes, '[i]t trades off social acceptance for self-esteem, dignity and self-respect. It results in the acceptance of a life lived in shame' (2018, p. 30). Such individuals face a choice between having a healthy respect for their own competency and worth but being excluded from the epistemic community or being admitted only with a low status that denies them respect and recognition of their competency. For this reason, intellectual servility can be a PECT because it is a means of avoiding or reducing hostility, confrontation and the harms which can follow from them.

In a similar way, intellectual timidity (self-silencing and lack of active engagement in epistemic projects stemming from a desire to avoid being negatively judged by others) can be a means of ego protection (ibid., pp. 32–

33). When one is likely to be unfairly dismissed as incompetent when speaking up, one can avoid the harms of such humiliation by shying away from being noticed.

It might be objected that Elena is protected, not by exercise of intellectual servility or timidity, but instead by behaviour which *appears* timid or servile. That is, protecting herself by 'keeping her head down' does not entail that she internalises the negative attitudes towards her own competency and abilities. All she needs is to appear to conform to stereotypes or keep quiet to avoid negative judgement whilst she maintains quiet confidence in her own abilities. Tanesini considers this possibility but is sceptical that such behaviour could be performed over long periods of time without the attitudes being internalised. As a means of avoiding cognitive dissonance the attitudes will tend to adapt over time to become consistent with behaviours (ibid., pp. 30–31; 2021, p.114; see also Cooper, 2008). Figure 1 summarises the four protective strategies and the epistemic character traits which support them.



Figure 1. PECTs supporting the four protective strategies

The development of PECTs supporting the 'keep your head down' and 'just don't listen' strategies can carry large costs. In particular, the development of epistemic vices constitutes a form of damage to epistemic character (Kidd, 2019a, 2020). While students may develop and exercise these traits to avoid conflict, they do so at the cost of undermining their competency as inquirers. By developing intellectual timidity or intellectual servility and accepting or avoiding the negative assessments of their abilities by their peers, students become vulnerable to a loss of trust in their abilities and miss out on opportunities for positive development. By becoming closed-minded, they may dismiss genuine criticism and feedback and may miss out on valuable opportunities to revise their beliefs and improve their epistemic position.

It might be objected that, given these costs, vicious PECTs are not genuinely protective since they lead students to be no better off, or even cause them to be worse off, than they would otherwise be had they not developed the epistemic vice. However, protective strategies are generally adopted as short-term responses to an immediate threat of harm. Perhaps they are not protective in the strong sense of ensuring that the student endures less harm overall when she develops or exercises the PECT. But they are protective in a weaker sense that they offer an immediate protection from the harm or harms at hand.

2.4 The shaping of character development

So far, I have presented four different protective strategies – each of which is supported by the development and exercise of various PECTs – which a student like Elena might adopt as a means of self-protection in a hostile classroom. Now I turn to considering the factors which might influence which strategy such students adopt.

Suppose that Elena comes from a working-class background and has a strong regional accent which people tend to associate with her socio-economic background. She is also the first person in her family to attend university and she often does not recognise, and feels alienated by, some of the cultural references and expectations that her middle-class peers share with each other. At least some of the hostility she encounters is rooted in classist prejudice. Throughout her education she receives very little encouragement in academic pursuits partly because her teachers, family and peers have low expectations of her because of her socio-economic background.

For the reasons set out above, intellectual perseverance and intellectual courage may be protective for Elena. But this may come at considerable cost and involves a trade-off between protection from harm and other aims. For example, she may avoid and mitigate some potential harms by pushing forward despite obstacles, but as a result she has to put so much energy and effort into repeatedly trying to be taken seriously in the discussion, that she ends up feeling drained and demoralised and has less energy and time to spend on other aspects of her learning or other important projects.

Furthermore, internalisation of the attitudes of her peers and the lack of encouragement she has received in academic endeavours, may contribute to a lack of recognition of herself as someone who is able to succeed in the classroom. This can make it difficult for her to become perseverant: if she does not see the end goal as within her reach, then it can be hard to see struggling to overcome obstacles in its pursuit as worthwhile.[48]

Of course, the picture will be extremely complex and the devil is in the detail; there will be many factors affecting how each individual student's epistemic character develops and is exercised in response to hostile classrooms. For now, I highlight three relevant factors: (a) the costs which the development and exercise of PECTs may carry for students given their context; (b) the way in which context can push epistemic character development in a particular direction and make some forms of development more difficult; and (c) the context-dependence of an epistemic character trait's status as protective.

2.4.1 Costs of developing and exercising PECTs

A PECT may help a student to prevent, limit or mitigate the incurrence of some particular harm and yet incur a cost. One way is by making the student more vulnerable to other types of harm: the exercise of perseverance is costly to Elena when it results in exhaustion and burnout. Even though she is disposed to weather setbacks for the sake of epistemic goods, repeated exposure may still have a lasting impact on her. This may be especially true if Elena must exercise a great deal of courage and perseverance in other areas of her life and has little opportunity for respite.

---

[48] The effects of stereotype threat (see Steele and Aronson, 1995) may also undermine the development of an individual's confidence in their ability within a particular domain by providing apparent evidence of lack of ability.

The exercise and development of PECTs can also be costly in terms of the time and energy it drains from other projects. Practising intellectual perseverance when one is constantly facing obstacles requires focus and determination. Having to spend so much time, energy and focus on being taken seriously may prevent Elena from focusing on listening to others, thinking about the topic in question, or developing her understanding. The demandingness of this process, if often repeated, may have long term impacts on Elena's learning, slowing down her progress or leaving other skills under-developed because she lacks the time or energy to invest in them.[49]

The aforementioned PECTs may also be costly if they are often misinterpreted, devalued or even punished when exercised by members of certain groups. For example, women (especially women of colour) may be regarded as bossy or arrogant when they display intellectual confidence because what others deem to be an 'appropriate' level of confidence in their abilities (given prejudicial stereotypes about their lesser intelligence and expectations about feminine modesty) does not reflect their actual capabilities. Such traits may carry costs through the repercussions with transgressing such norms, for example, being passed by for opportunities because one is seen as having a difficult character or being seen to fulfil oppressive stereotypes.

Students' abilities to take on these costs vary. Those who incur high costs elsewhere or look to the classroom as a place of respite because they face hardship in other aspects of their lives may be disproportionately disadvantaged by the costs associated with developing and exercising PECTs.

Note however, that the costs of a PECT need not always prevent students from adopting it as a means of self-protection. I argued at the end of §2.3 that the development of vicious PECTs carries costs for students. However, this does not mean that students will never adopt them. For example, students may not anticipate future costs of developing vicious PECTs or those costs will be less salient to them given the scale and immediacy of the harms from which they are attempting to protect themselves. Whilst students may themselves appear willing to adopt protective strategies

[49] The relationship between discrimination and psychological health has been explored in literature on *minority stress* (see Meyer, 2003, 2007).

which carry significant costs, educators should not teach or expect students to adopt PECTs which carry significant costs for them.[50]

2.4.2 Oppression and the trajectory of epistemic character development

The second factor influencing students' responses to hostile climates in the classroom is the way in which the trajectory of their epistemic character development is directed or limited by background conditions. José Medina (2013), building on work from feminist and critical race theorists such as Claudia Card and W.E.B. Du Bois, examines how relationships of privilege and oppression create tendencies towards the development of particular epistemic character traits in differently privileged groups. He describes trends in character development amongst those who are privileged and those who are oppressed; charting patterns in the socio-genesis of many traits within relationships of privilege and oppression.[51]

Medina acknowledges that in many ways oppression hinders the development of certain forms of virtuous epistemic character:

> '[A]mong the consequences [of oppression] for marginalized subjects is the feeling of intellectual inferiority, a poor self-assessment of one's cognitive assets and capacities, which has many ramifications and an overall negative impact on one's daily life (the way one speaks or thinks; the way one envisions goals as legitimate and feasible, the way one responds to challenges, and so on).' (*ibid*. p.41)

Students who are systematically treated as inferior, less intelligent or less worthy of an educator's attention, are deprived of the encouragement, material resources, challenges, advice or constructive criticism that can build proper pride and intellectual confidence. For such students, cultivating these PECTs can require greater cognitive effort than for those who have received help and reinforcement along the way. If a student's general level of intellectual confidence has been eroded over time, it is less likely to be well-calibrated to their intellectual abilities (see also K. Jones, 2012). In this way, oppression and disadvantage can function as forces working against the development of 'believe in yourself' PECTs.

---

[50] I will expand on this point further in chapter 5.

[51] Medina does note, however, that these patterns of development are neither inevitable for, nor exclusive to, any particular group (2013, pp. 40, 43).

The same can be said regarding PECTs supporting the 'battle on through' strategy. If a student is not treated by others as a capable knower and participant in the classroom, that student is less likely to develop a view of themselves as a capable knower and participant. If they consider themselves to be someone who *cannot* achieve long term goals because they lack the necessary skills and abilities, they are unlikely to see short term adversity and challenges to be worth weathering for the sake of epistemic goals. Therefore, they are less likely to exercise intellectual courage and perseverance. On the other hand, it may be much easier for such students, whose environment promotes low self-confidence and self-esteem, to adopt the 'keep your head down' strategy.

2.4.3 Failure to be protective

Even when epistemic character traits supporting the 'believe in yourself' and 'battle on through' strategies can be developed, they will not be protective for all students who develop them. Different students will have different circumstances – with different challenges and resources. As such, the same trait may have different benefits and drawbacks for different students. Therefore, the value of an epistemic character trait as protective – and thus its status as a PECT – is contextual.[52] A trait which is a PECT for some student in some particular context may not be a PECT for another student in that context or the same student in a different context.

Imagine another student in a situation similar to Elena's but for whom the degrading and dismissive treatment from their peers is part of a much larger and systematic problem; the education system as a whole is biased against people from their social group and the discrimination they face is reinforced at an institutional level. A motivation to overcome obstacles in order to achieve a goal can only take them so far if the obstacles are simply too great for them to overcome or they lack support from their peers and teachers. This suggests that intellectual courage will not prevent them from being excluded from discussion and missing out on opportunities to learn and develop epistemically. Instead, intellectual courage may leave them exhausted and drained still with little chance of ultimate success.

---

[52] In chapter 3, I will discuss a related but stronger claim – that the status of an epistemic character trait as a virtue or as a vice is contextually dependent (see Daukas, 2019; Kidd, 2020). However, the claim I make here is merely that a trait's status as a PECT is contextually dependent.

It is worth emphasising the distinction between epistemic character traits which fail to be protective for a particular student in a particular context and PECTs that result in, or make the student more vulnerable to, harms other than those they protect against. We may still say that these latter traits are protective in some respect (and thus are PECTs for the students in question) even if they do not make them better off overall.

2.5 Implications for teaching epistemic virtue

It would be wrong to conclude from the existence of vicious PECTs that we should begin to teach them to students as means of protection. Besides an intuitive resistance to actively teaching epistemic vices, doing so fails to alleviate the burden placed on such students who take on additional harms by developing epistemic vices to protect themselves. (In chapters 3 and 4, I will further explore the costs of vicious epistemic character development.) Neither does it do anything to address the limitations that are placed on students' epistemic character development which cause them to develop vicious rather than virtuous PECTs.

Through the recognition of certain epistemic vices as PECTs, we can identify the aetiological root of those epistemic vices for some individuals (see Kidd, 2020, pp. 78–80). Namely, we can recognise that they are developed as means of protection when the development of epistemic virtues as a means of self-protection is not a viable option for students. Given the practical aim of positively developing students' epistemic character through education, understanding why epistemic vices are developed is an important step in avoiding and counteracting vicious epistemic character development. If we fail to recognise the protective role of these epistemic vices and the conditions which influences students towards their development, then we run the risk of unwittingly perpetuating their development by creating conditions ordinarily thought to facilitate the development of virtuous PECTs.

Looking through a critical lens, we can see that the relationship between epistemic character and protection is complex. Using the tools of critical character epistemology, we can approach the teaching of virtue so that it reflects how the social context and positions within oppressive systems will affect students. This change in our approach results in thinking about when an epistemic character trait counts as protective as well as the cost of developing such a trait and what obstacles there are to this development. This approach requires sensitivity to the underlying structures to realise the kind of protection that is necessary for some students. One result of this

contextualism is a challenge to the idea that there are some standard or even paradigmatic virtues which should be taught to all students as a means of protection because the epistemic virtues on these standard lists will not be protective for all students.

It might be objected that the standard character epistemic approach is already context-sensitive because having the relevant epistemic virtues to a high degree requires having the practical wisdom to appropriately judge in which contexts to exercise them (see for example, Zagzebski, 1996; Johnson, 2021). If one genuinely possesses intellectual courage, then one would not expose oneself to harms for the sake of epistemic goals *regardless* of the severity of the potential harm and the likelihood and value of success. The intellectually courageous individual is not reckless – they use practical judgement to assess whether the risk of harm to which they expose themselves is not too great and is worthwhile. Intellectual courage would not demand of the student facing systemic discrimination that they stand up against these obstacles alone if doing so would be futile. Likewise, intellectual perseverance will not dispose individuals to endlessly pursue hopeless projects or ones whose end goal does not justify the time and energy which must be put into them (Battaly, 2017, p. 684).

Furthermore, a single epistemic virtue – such as intellectual courage – can take a wide variety of forms. Kidd argues that we should avoid the tendency to think of intellectual courage as manifested only in positive actions such as taking a stand or defending a belief (2019b, p. 249). A whole range of activities, including terminating inquiry which will not lead to epistemic goods and abandoning deeply held beliefs, can be exercises of intellectual courage for certain individuals under certain conditions (*ibid.*). If education is to effectively promote the development of intellectual courage to a range of students, then it must lead students to recognise the range of possible forms which the virtue can take under different contexts and in response to different harms. The problem with typical educational approaches, it could be argued, is not the focus on intellectual courage per se, but rather the focus on a narrow picture of the forms which courage can take. If students were presented with and educated to develop a greater variety of forms of intellectual courage, then more students would be able to develop intellectual courage as a PECT.

I accept that standard character epistemology can accommodate some context-sensitivity in this way, however, I raise three points as to why this is not sufficient. Firstly, the contextualism provided by the role of practical wisdom in epistemic character has substantial implications for the teaching

of epistemic virtues. Consider the paradigmatic picture of the intellectually courageous individual as being undeterred in their epistemic projects by threats of harm. They are not easily overwhelmed by challenges and adversity, but for the sake of avoiding foolhardiness they do recognise that in some scenarios backing down is the practically wise thing to do. Such a model has limited usefulness for teaching students who are systematically vulnerable, because for them the majority of the threats and challenges they face are ones that they cannot stand up to alone. If they possess intellectual courage, it will likely look very different.

Secondly, practical wisdom in exercising intellectual courage is displayed by those who have developed the epistemic virtue to a relatively high degree – those who have practice in exercising (or not) the virtue in a range of different contexts. Students who are just beginning to develop intellectual courage will not be so skilled in judging when its exercise is appropriate. As such, they may be prone to exercise intellectual courage in contexts which do them more harm than good. Therefore, even if intellectual courage is protective when developed to a high degree, the process of learning to be intellectually courageous may result in some students being prone to exercise intellectual courage in harmful ways – especially when the standard picture of what an intellectually courageous individual looks like does not match up to their reality.

Thirdly, and most crucially, when practical wisdom dictates that intellectual courage should not be exercised, it does not offer any protection. In scenarios in which the battles are just too great for students to overcome, the virtue of intellectual courage (along with practical wisdom) tells them that now is *not* the time to battle on through. Intellectual courage therefore does not aid them in adopting a 'battle on through' strategy for avoiding potential long-term harms.

Rather than accept the status quo and further entrench it by encouraging students to develop vicious PECTs, there are two possible ways of responding. Firstly, the structural reasons for which the trajectory of students' character development is pushed towards the development of vicious PECTs or for which virtuous PECTs carry significant costs or fail to be protective. However, this is a long-term solution which cannot be achieved overnight. Therefore, secondly, by avoiding the need for students to protect themselves when doing so would lead to vicious epistemic character development, the harms arising from this can be reduced in the short term. In chapter 5, I will argue that educators and educational institutions have a duty to protect students from harm when failing to do so

would entail that students' only means of self-protection would cause them or make the vulnerable to other harms. In the next two chapters, I will examine in more detail the costs arising from the development of vicious PECT

Chapter 3

Epistemic Character Damage and Normative Contextualism

In the previous chapter, I identified a number of epistemic vices which can be developed and exercised by students in order to aid them in protecting themselves against harm. However, the development of epistemic vice as a form of self-protection is not desirable. Most intuitively, it is undesirable because the development of such epistemic vices will have negative consequences in their epistemic endeavours more broadly – they are likely to exercise the vices in other contexts where they will have bad effects and are no longer needed as means of self-protection. There is also a deeper sense in which students are made worse off in terms of the *damage* done to their epistemic character. To isolate this deeper sense in which character is damaged, I consider cases in which an individual's character develops in ways which are conventionally recognised as epistemically vicious in order to cope within a ubiquitously hostile environment (such that there are very few opportunities for them to escape the hostile environment) and in which their vicious epistemic character development makes them epistemically better off than they otherwise would have been. In such cases, where vicious epistemic character development makes an individual better off than they otherwise would be, how should we understand the nature of the damage which the development of the vice does to epistemic character? In other words, in what sense does vicious epistemic character development required for coping in a hostile environment make an individual worse off?

In the process of answering this question, I identify a tension between two features of Kidd's critical character epistemology. In character theory both within and beyond epistemology, an attempt to pay attention to the socially embedded nature of epistemic character has been made through the development of critical character theory (Dillon, 2012). One of the central features of critical character theory – as developed by Robin Dillon (2012) and recently applied to character epistemology by Ian James Kidd (2020) – is the doctrine of *normative contextualism*. The doctrine's core claim is that

the normative status of a character trait is dependent on the specific context of its bearer (Kidd, 2020, p. 81). The normative status of a trait is its classification as a virtue or as a vice. Therefore, whether a character trait is classified as a virtue or as a vice will depend on the context of the particular agent by whom it is exercised. Differently situated agents face different challenges and as a result, which habits and dispositions will be valuable to them given their specific contexts will differ. If a character trait (even one conventionally considered a vice) enables a particular agent to survive under adverse conditions (including surviving and resisting conditions of oppression), then according to normative contextualism, it is a virtue for that agent in that context (*ibid.*).

Another aim of critical character theory is to explore how character is shaped by the social world – particularly relationships of power, privilege and oppression (Dillon, 2012, p. 84). This includes identifying environments, structures, institutions or agents, which corrupt character (Kidd, 2020, p. 71). Identifying corruption of character and understanding how and why it occurs is an important step in creating conditions in which the development of vices can be avoided, discouraged, or reversed.

In this chapter, I identify a tension between two aims of critical character epistemology. This tension arises when we consider a particular species of epistemic corruption: the development of what I will call 'epistemic coping strategies'. Under certain oppressive conditions, epistemic character traits conventionally understood as epistemic vices can be developed as a means of coping with the challenges in the environment. Epistemic coping strategies are strategies which allow agents to do better epistemically under such conditions than they would otherwise by developing traits not normally endorsed. In such cases, I will argue that there is a tension between the two claims following, respectively, from the desire to classify the development of epistemic coping strategies as epistemic corruption and the endorsement of normative contextualism:

1) As products of epistemic corruption, the traits developed as part of an epistemic coping strategy must be epistemic vices or hinder the development of epistemic virtues.

2) According to normative contextualism, as traits which help agents respond to and cope with their circumstances, traits developed as part of an epistemic coping strategy are epistemic virtues.

Therefore, traits developed as part of epistemic coping strategies must simultaneously be epistemic virtues and hinder the development of epistemic virtues (or be epistemic vices).

I will argue that we can avoid this tension by modifying normative contextualism. The position I develop in this chapter – *dual-level normative contextualism* – involves the evaluation of epistemic character traits on two separate levels. One level of evaluation reflects the salience and effectiveness of the epistemic character trait as a means for the individual to respond to the particular challenges of their epistemic predicament. The second level of evaluation reflects the relationship between the epistemic character trait and the non-contextual ideal of epistemic flourishing. This allows us to simultaneously and consistently evaluate the same traits both negatively (and thus class their development as epistemic corruption) and positively (in accordance with normative contextualism).

In section 1, I will define and give an example of an epistemic coping strategy. In section 2, I present Kidd's account of epistemic corruption and show that it requires changes to epistemic character involving the development of epistemic vice or which hinder the development of epistemic virtue. I argue that epistemic coping strategies are a form of epistemic corruption and therefore the traits developed as part of them are epistemic vices or hinder the development of epistemic virtues. In section 3, I set out the motivations for and main claims of normative contextualism. I then argue that according to normative contextualism, traits developed as part of epistemic coping strategies are epistemic virtues because they are effective and salient means for agents to respond to the challenges of their particular contexts. In section 4, I make explicit the conflict between the conclusions arrived at in sections 2 and 3: the traits developed as part of epistemic coping strategies cannot be both epistemic virtues and hinder the development of epistemic virtues (or be epistemic vices). I then propose a resolution to the tension by developing *dual-level normative contextualism*.

3.1 Epistemic Coping Strategies

As seen in the previous chapter, epistemic character traits (including epistemic vices) can be developed as means of self-protection. PECTs fall under a broader category of changes to epistemic character which occur as a means of coping with an oppressive, hostile or otherwise difficult environment.

Under certain conditions agents can be forced to adopt epistemic character traits, which under normal conditions are defects of epistemic character, in

order to cope with the challenges of their environment. In some cases, this can involve the sacrifice of epistemic goods for the sake of other goods. Consider the characters in George Orwell's (1949) novel *Nineteen Eighty-Four* who face persecution for challenging the complete epistemic authority of the totalitarian Party. Many of them develop epistemic vices such as extreme gullibility or closed-mindedness in order to keep their heads down and protect themselves.[53]

The vicious[54] PECTs identified in the previous chapter can function in this way. For example, the development of intellectual servility in a hostile environment can be a means of protecting goods such as social inclusion at the cost of epistemic goods.[55]

However, it need not always be the case that the development of traits conventionally considered to be epistemic vices involves the sacrifice of epistemic goods for the sake of other goods. Under some extreme hostile conditions, adopting epistemic character traits which typically have bad epistemic effects for the individual and would not normally be endorsed or admired, actually help agents do epistemically better than they otherwise would. This kind of case is discussed by Heather Battaly (2018) when she considers the possibility that closed-mindedness can have good epistemic effects (and be an 'effects-virtue') in 'epistemically hostile' environments heavily polluted with propaganda and misinformation. For ease of reference, I will refer to these cases as 'epistemic coping strategies'. For something to be an epistemic coping strategy in the sense I have in mind, it must, as a matter of fact, make the individual epistemically better off than they would otherwise be, were they not to adopt the strategy.[56]

---

[53] Heather Battaly (2018, p.39-40) discusses the utility of closed-mindedness in relation to this example.

[54] For terminological ease, I will continue to use the term vicious PECTs to refer to PECTs which are conventionally identified as epistemic vices (those in the 'keep your head down' and 'just don't listen' clusters), though according to normative contextualism, these PECTs may, for some individuals in some contexts, be virtues.

[55] In chapter 4, I will argue that in some extreme cases, the development of intellectual servility may lead to an individual's epistemic agency being less constrained than it would be, were she not to develop intellectual servility. In such cases it is possible that the individual is better able to attain epistemic goods by developing intellectual servility than she would otherwise.

[56] For any particular epistemic coping strategy, there might be genuine disagreement about whether it does genuinely make individuals epistemic better off than they would otherwise be (and therefore whether it qualifies as an epistemic coping strategy). It suffices for my

Consider the following example. The voices of people with marginalised racial identities are often suppressed, overlooked and silenced in structurally racist societies. Basic conversations about the existence and nature of racism can be difficult to have with those who do not experience it. Often, honest and frank accounts of the ways in which people's lives are affected by structural racism do not fit with dominant narratives of a fair, post-racist world which people are keen to maintain. As a result, these conversations can be draining, frustrating or even dangerous, leading some people of colour to avoid them.[57] As the writer and journalist Reni Eddo-Lodge writes:

> I can't talk to white people about race anymore because of the consequent denials, awkward cartwheels and mental acrobatics that they display when this is brought to their attention […]
>
> I cannot continue to emotionally exhaust myself trying to get this message across, while also toeing a very precarious line that tries not to implicate any one white person in their role of perpetuating structural racism, lest they character assassinate me. (Eddo-Lodge, 2017, pp. x–xii)

Like the example from *Nineteen Eighty-Four,* this is a coping strategy. Eddo-Lodge protects herself from burnout[58], 'character assassination' and other hostile or aggressive responses, should her interlocutors wish to silence or discredit her to avoid addressing the issues she is highlighting.

In addition, this approach can also be adopted to give epistemic benefits and avoid epistemic harms in at least two ways. Firstly, it allows more room for deeper and more fruitful discussion within marginalised communities (and amongst genuine allies) because time and energy are not wasted

---

point in this chapter that there are at least some epistemic coping strategies which make individuals better off than they would otherwise be.

[57] This is similar to a practice which Kristie Dotson (2011, p. 244) terms 'testimonial smothering', wherein a speaker pre-emptively edits or withholds certain kinds of testimony because they believe that offering such testimony would put them at risk of hostile reprisal, or they avoid expressing themselves in certain ways, because they believe their audience will be unable and unwilling to understand.

[58]This burnout may be caused by epistemic exploitation (Berenstain, 2016; see fn.20). Berenstain argues that the demand for marginalised persons to do the uncompensated and emotionally taxing epistemic labour of educating others about the nature of their oppression is epistemically exploitative.

establishing foundational claims like the reality of racism. Secondly, by avoiding engagement in conversations about racism, it protects against the epistemic harms arising from the systematic dismissal, testimonial injustice (Fricker, 2007) and smears on credibility that can be used as tactics to avoid facing uncomfortable truths and conversations. As Miranda Fricker (2007, pp. 47–48) notes, repeatedly being on the receiving end of such dismissal can undermine speakers' faith in themselves as competent epistemic agents.[59] Testimonial injustice can also cause harm by limiting agents' status within the epistemic community – treating them as 'derivatized' subjects (Pohlhaus Jr, 2014, pp. 106–107). Therefore, the approach adopted by Eddo-Lodge can be described as an epistemic coping strategy: a strategy developed to allow her to avoid some of the potential epistemic harms in her environment and direct her time and energy towards more fruitful and worthwhile epistemic endeavours.

3.2 Epistemic Corruption

There are many things which can affect the development of epistemic character and in particular many things which can divert character development away from the development of virtue and towards the development of vice. We often talk of environments, people, institutions, practices and educational systems having a corrupting influence on individuals. Often the pressures resulting from oppressive environments incentivise practices which wouldn't normally be endorsed or lead individuals to face moral dilemmas which have a lasting impact on their moral character (Tessman, 2005). This is just one of the ways in which oppression causes harm.

As part of his critical character epistemology, Kidd (2020, 2021, forthcoming) explores in depth the ways in which a range of things can corrupt epistemic character and the different forms which this corruption can take. For Kidd, *epistemic corruption* occurs when an individual's epistemic character comes to be damaged as a result of interactions with agents, environments, institutions, or structures that facilitate the development and exercise of epistemic vices or prevents and erodes the development of epistemic virtues (*ibid.*, p.71). [60] I will explore what is meant by damage to epistemic character in §3.2.1. Epistemic corruption can occur

[59] See also Bartky (1990) on internalised or 'psychological oppression'.

[60] Kidd distinguishes between two modes of corruption: 'active' corruption (vice-promoting) and 'passive' corruption (virtue-stifling) (2020, p.71).

through a number of different means, including the absence or derision of virtuous exemplars, the valorisation of vice and vicious exemplars, establishment of policies which necessitate the exercise of vice, and incentivisation of the exercise of vice (*ibid.*, pp.75-76). Recognising epistemic corruption and understanding its causes, processes and effects is an important part of the ameliorative project of vice epistemology since promoting the development of virtuous rather than vicious epistemic character will require preventing, mitigating and reversing the effects of epistemic corruption.

In §3.2.1, I argue that for epistemic corruption to take place, it must result in a change to epistemic character that either involves the development of epistemic vice or hinders the development of epistemic virtue. Then, in §3.2.2, I argue that necessitation of epistemic coping strategies is a form of epistemic corruption. Therefore, as instances of epistemic corruption, epistemic coping strategies must involve the development of epistemic vice or changes to epistemic character which hinder the development of epistemic virtues.

3.2.1 Epistemic Character Damage

For an environment, institution or practice to be epistemically corrupting, it must be liable to cause *damage* to epistemic character. The idea of damage to character has been explored in the moral domain in the work of feminist ethicists Claudia Card (1996) and Lisa Tessman (2001, 2005).[61] Card explores the *constitutive moral luck* involved in the development of the self. In particular, she focuses on the ways in which the development of moral agency can be limited, or prevented entirely, by the 'damaging nature of oppression' (1996, p.4). Tessman (2005) develops this idea within the framework of Aristotelian virtue ethics. The possession of virtue is not sufficient for one's life to go well - certain external factors also affect our ability to lead a good life: ill-health, misfortune or poverty can prevent one from living a good life even when one is perfectly virtuous.[62] Oppression can impact these external factors: it can deprive one of material resources, deny

[61] Focus on the impacts of oppression on character has a long precedent in the work of, for instance, Du Bois (2015), Cesaire (1972), Wollstonecraft (2004), and Astell (1999).

[62] For example, Aristotle writes that external goods are required for happiness, either because they are required for virtuous action or their lack is 'like a stain on happiness' (Nicomachean Ethics, I.8, 1099a31-b3). See also Nussbaum (1986).

opportunities or impact health and well-being such that one cannot live a good life.[63]

But, as Tessman argues, oppression can also negatively affect one's ability to flourish by impacting who one is, how one thinks, and by affecting one's values and motivations.[64] *Moral damage* occurs when an agent is 'prevented from exercising or developing some of the virtues' or 'when there is a certain sort of self that one ought to be, but the unconducive conditions of oppression bar one from cultivating this self' (*ibid*., p.4)

Kidd notes that the attention paid to character damage within feminist and critical race theories have tended to focus on damage to *moral* character. He speculates that such work has 'useful resources to offer those who want to understand the character-epistemic dimensions of social oppression'. (2020, p. 69). Transferring this notion across to the epistemic domain, Kidd describes damage to epistemic character as:

> the deterioration of any pre-existing virtues and integrity already present in the subject's character, or the failure of the subject to develop an epistemic character characterised by virtues and integrity. (Kidd, 2020, p. 71)

Damage to epistemic character need not always be the result of epistemic corruption. It can also be the result of bad luck or failing cognitive abilities for which no one is culpable.

Two things must occur in order for epistemic character to be damaged. Firstly, epistemic character damage requires *change* to epistemic character. Something has not been damaged if nothing about it has changed. There are ways in which changes to an individual's circumstances may change the *prospects* of her virtuous epistemic character development without making an actual change to her epistemic character. For example, imagine a young student who moves from a school which encourages intellectual humility to one with a culture in which intellectual arrogance is encouraged and rewarded. On day one in her new school, her prospect for developing the virtue of intellectual humility – the trajectory of her epistemic character development – has changed, although no epistemic character changes have

---

[63] Tessman argues that we do not need to give a fully fleshed out account of the good life in order to understand the concept of moral damage. All we need is an understanding of the way in which lives go better, for example, in the absence of overwhelming loss, when one believes that one is deserving of love, and when one does not fear violence. (2001, p. 79).

[64] Tessman gives 'psychological oppression' (Bartky, 1990) – the internalisation of prejudicial attitudes towards oneself – as one example of how this can occur.

yet taken place. Therefore, her character has not been damaged at the point where she first joins the school (although it may be damaged at a later time).

Secondly, to damage something is to change it for the worse. The changes which take place as a result of epistemic corruption must be negative. For Kidd, the damage done to the epistemic character of epistemically corrupted agents is the deterioration of epistemic virtues – this includes the loss, weakening, or destabilising of virtues – as well as the failure to develop epistemic virtues (*ibid.*). However, it is not the case that epistemic character is damaged whenever epistemic virtues are not being developed: epistemic agents, even those who are virtuous, need not be constantly developing epistemic virtues whenever they engage in epistemic activities. Therefore, when Kidd states that damage to epistemic character includes the 'failure to develop epistemic character characterised by virtues and integrity' (*ibid.*), we should interpret this to mean changes to epistemic character which *hinder* the development of virtuous epistemic character – that is, decrease the agent's prospects or likelihood of developing epistemic virtues.

3.2.2 Epistemic Coping Strategies as Epistemic Corruption

One way in which environments can be epistemically corrupting is by incentivising the development and exercise of epistemic vice and disincentivising the development and exercise of epistemic virtue (Kidd, 2020, pp. 76–77). In such environments individuals who develop and exercise epistemic vices may, for instance, benefit from social goods such as esteem and respect from others or accrue material goods such as greater success in their careers than those who are epistemically virtuous. In some cases, the development and exercise of epistemic vices is incentivised because it is the only or most viable response to threats of harm. The environments which give rise to vicious PECTs and epistemic coping strategies fall into this category: epistemic vices are incentivised because they lead to better outcomes and prevent harms while epistemic virtues are disincentivised because they lead to worse outcomes and do not prevent harms.

In the previous section, I established that epistemic character damage (and therefore epistemic corruption) requires either (a) the development of epistemic vice or (b) changes to epistemic character that hinder the development of epistemic virtue. Therefore, it follows that, as a form of epistemic corruption, the traits developed as part of epistemic coping strategies will either be epistemic vices or will hinder the development of epistemic virtues. In short, we can argue the following:

(A1) For epistemic character damage (and thus epistemic corruption) to occur there must be changes to epistemic character which involve the development of epistemic vice and/or hinder the development of epistemic virtue.

(A2) The development of an epistemic coping strategy is an instance of epistemic corruption.

(A3) Therefore, the development of an epistemic coping strategy must include changes to epistemic character which involve the development of epistemic vice and/or hinder the development of epistemic virtue.

The traits developed in Eddo-Lodge's strategy are conventionally recognised as epistemic vices, or at the very least are contrary to the kind of character profiles conventionally evaluated as epistemically virtuous. The tendency to truncate one's testimony for the reasons which Eddo-Lodge describes is putatively at odds with the virtue of intellectual honesty since it involves avoiding or leaving out certain truths pertinent to the topic at hand or altering one's testimony so that it is more palatable than telling the truth. Holding an attitude of hostility and suspicion towards others and making a general assumption of bias and prejudice in others seems opposed to the virtue of intellectual charity – the disposition to interpret others charitably. There may also be aspects of closed-mindedness at play within this coping strategy if the agent holds firmly to her experiential beliefs regarding the existence and nature of racism and is unwilling to engage with any alternative viewpoints, even if just to argue against them.

The concept of epistemic corruption is useful as a means of evaluating environments, practices and institutions. It highlights the negative changes that they make to epistemic character. Just as we explain this harm in terms of epistemic corruption in cases like *Nineteen Eighty-Four*, it is desirable that Kidd's account of epistemic corruption can play this explanatory role with regard to the adoption of Eddo-Lodge's strategy.

Perhaps the problematic nature of these environments can also be located elsewhere – in, for example, injustices, inequality or the objectification of members of marginalised groups. But, whilst these often coincide with epistemic corruption, we should not neglect the harm that is done through the impact on individual character. The best possible epistemic life that can be lived under oppressive conditions – achieved through the adoption of an epistemic coping strategy – is not comparable to living the best possible epistemic life under non-oppressive conditions. Recognising the changes to agent's epistemic character as damage (and thus the result of epistemic

corruption) allows us to recognise how oppression causes harm by limiting and distorting character development (Tessman, 2001, 2005, p. 19).

3.3 Normative Contextualism

I will now turn to normative contextualism and demonstrate how its application to cases of epistemic coping strategies brings us to a conclusion that conflicts with (A3). I will explain the motivations for and define normative contextualism and then argue that the traits which are developed as part of epistemic coping strategies are classified as epistemic virtues according to normative contextualism because they are effective and salient means for their bearer to respond to the particular challenges of their epistemic predicament.

In §2.4.3, I noted that the status of an epistemic character trait as protective is contextually dependent. No trait is universally protective – they are only protective for a particular individual in a particular context. The normative contextualism which I discuss below is a stronger version of this position: that the status of a trait as a virtue or as a vice is context dependent. The defence of this position aids the project of critical character epistemology in the same way as recognition that the protectiveness of a trait is context dependent. That is, recognising the complex relationships between epistemic character and social positionality requires acknowledgement that the value of an epistemic character trait may vary depending on the specific context of its bearer.

Normative contextualism has been defended and endorsed in a number of different forms within character epistemology. Heather Battaly (2018) considers whether closed-mindedness can be an 'effects-virtue' given its potential positive effects under certain extreme conditions. Nancy Daukas (2019, pp. 380–381) argues that acontextually characterising certain epistemic character traits as epistemic vices by definition can obscure the ways in which they may be helpful to members of marginalised groups as means of surviving under conditions of oppression.

Different epistemic agents have different *epistemic predicaments* (Medina, 2013, p. 28; Kidd, 2020, pp. 73–4). That is, each agent is uniquely socially and epistemically situated, facing different epistemic challenges and having access to different epistemic opportunities and resources. As a consequence, different agents require different epistemic character profiles in order to achieve epistemic excellence (Daukas, 2019, p. 381). For example, developing epistemic humility may help agents prone to arrogance achieve epistemic excellence, but its development may not be so beneficial to those

who are socialised to be prone to doubt their epistemic capabilities or are generally regarded by others as less capable.

It is important to note that the value of a trait for an individual, which determines whether it is a virtue for them, should not be understood simply as the value which the trait has for the individual in their immediate context. The normative contextualist does not claim that the normative status of a trait can vary from situation to situation for an individual. For example, it is not the case that closed-mindedness becomes an epistemic virtue for an individual when they are, for a short time, in a setting where they can easily be misled by misinformation, but then becomes an epistemic vice from them when they are in a setting that is not heavily polluted with misinformation. Rather, the normative status of a trait for an individual is dependent on the value that it has for them in their epistemic life more broadly.[65]

A consequence of normative contextualism is a weakening of the connection between virtue and praise: whilst the normative contextualist might agree that, say, intellectual arrogance can be a virtue for some individual given their epistemic predicament, they may want to resist the claim that we should praise that individual for their intellectual arrogance. Thus, endorsement of normative contextualism should go hand-in-hand with an expansion of our evaluative repertoire to also include evaluations such as endorsement of a trait, but regret that it must be developed. I will explain in §3.4.2, how my proposal for dual-level normative contextualism can support such evaluations of epistemic character traits.

It could be argued that the use of the terms 'virtue' and 'vice' does not seem appropriate when discussing these traits whose value is contextually dependent. Normative contextualism is posed in direct opposition to conventional character theories which take virtues and vices to be defined and valuable acontextually. By using the terminology of virtues and vices the normative contextualist is therefore positing that the categories of

---

[65] It is most natural to spell this out in consequentialist terms: epistemic virtues are traits which, for that individual, given their epistemic predicament, lead to the greatest acquisition, retention and sharing of epistemic goods overall. However, I leave open the possibility that it might be developed in a non-consequentialist fashion. For example, the normative contextualist might challenge the claim that certain motivations – such as a 'love of truth' (Zagzebski, 1996) – have intrinsic value. Rather, different motivations may be valuable for different individuals given their epistemic predicaments. If such a challenge could be made, then the normative contextualist need not define epistemic virtues and vices in purely consequentialist terms.

virtue and vice as understood by conventional (acontextual) character theories are misunderstood and should be replaced with the notion of traits whose value is contextually dependent. For those who accept normative contextualism, it may seem that the terminology of virtues and vices is less helpful, since it carries connotations of acontextually valuable traits. Doing without this terminology is in principle acceptable for the normative contextualist, however, I will continue to use it to demonstrate how the view is posed as a direct challenge to acontextual theories of virtue and vice. (I will return to this terminological point briefly in fn.71 in relation to my proposed revision of normative contextualism.)

Kidd (2020, p. 82) makes a distinction between *strong* and *weak* normative contextualism. According to *weak normative contextualism,* some or all epistemic character traits have a default normative status but under certain conditions this default status is overridden. *Strong* normative contextualism denies this and holds that some or all epistemic character traits only have a normative status in relation to particular agents in particular contexts. I will return to the significance of this distinction in section 4.1, but for now I will offer a general picture of normative contextualism.

Once we accept the fundamental claim of normative contextualism – that the normative statuses of some or all epistemic character traits are context-dependent – we can then ask what determines the normative status of epistemic character traits for particular agents and in particular contexts. Our answer to this question should be motivated by critical character theory's concern with understanding the social nature of character and, in particular, how it allows agents to respond to conditions of oppression.

Therefore, the normative status of an epistemic character trait, for some agent in some context, will be determined by how it allows a particular agent to respond to the particular challenges and opportunities that make up their epistemic predicament. They have a positive normative status (as epistemic virtues) if they are effective and salient ways for the agent to respond to their epistemic predicament and will lead them to do well epistemically given the circumstances.

We can therefore define normative contextualism as follows:

(B1) Normative contextualism: Traits which are salient and effective means for some particular agent to respond to the particular challenges of their particular epistemic predicament are, for them given their epistemic

predicament, epistemic virtues (and not epistemic vices)

The epistemic character traits developed as part of epistemic coping strategies enable their bearers to do better epistemically given their epistemic predicaments than they would otherwise. Therefore, according to normative contextualism, these traits have a positive normative status as epistemic virtues.

In other words, we can argue the following:

(B2) Epistemic character traits developed as part of epistemic coping strategies in hostile or oppressive environments are salient and effective ways for a particular agent to respond to the challenges they face, given their particular epistemic predicament.

(B3) Therefore, the epistemic character traits developed as part of epistemic coping strategies are epistemic virtues (and not epistemic vices) for that particular agent in that context.

It could be objected that the 'some or all' clause in Kidd's definition of both strong and weak normative contextualism leaves open the possibility that the traits developed as part of the epistemic coping strategies are not the kind of epistemic character traits whose normative status is contextually dependent. Kidd does not offer any suggestion as to a criterion by which we can distinguish those traits with a contextually dependent normative status and those without. However, given the examples that Dillon uses, the traits developed as part of epistemic coping strategies appear to be a paradigm example of the kind of traits whose normative status is contextually dependent.

3.4 (Resolving) the tension

In the remainder of the paper, I will pinpoint the tension and present my solution. The two arguments given above – (A1) to (A3) and (B1) to (B3) – give rise to seemingly conflicting conclusions: (A3) and (B3). The changes to epistemic character that occur in epistemic coping strategy cases cannot be both the loss or hinderance of epistemic virtue and the development of epistemic virtue.

The classification of epistemic coping strategies as a form of epistemic corruption rests upon the recognition that the epistemic character traits to which they give rise change epistemic character for the worse. But this

comes into conflict with the implications of normative contextualism: those same epistemic character traits are epistemic virtues in light of their value for their bearers in coping with their environments. If the normative status of epistemic character traits is determined solely by their value for a particular individual given their epistemic predicament, then we lose the ability to criticise the environments which necessitated the development of epistemic coping strategies on the grounds of being epistemically corrupting.

In section 4.1, I argue that weak normative contextualism has the potential to provide a solution to the tension because it distinguishes between the contextualised and default status of an epistemic character trait. Then in section 4.2, I argue that for this solution to work we must understand the default status as measured against a non-contextual ideal like epistemic flourishing. Finally, in section 4.3 I demonstrate how my proposed dual-level normative contextualism resolves the tension.

3.4.1 Strong and weak normative contextualism

Strong normative contextualism rejects the notion of a default normative status for some or all epistemic character traits. It is the claim that character traits must be defined in value-neutral terms and only within a particular context and in relation to a particular agent can they be assigned a normative status as an epistemic virtue or vice. (Kidd, 2020 p.83). For example, Nancy Daukas (2019, pp. 380–381) argues that an inherently value laden definition of the trait of vanity leaves no room for recognition that for members of marginalised groups who must be vigilant about how they are perceived, taking special care to ensure that one is seen positively by others, can be a virtue. According to strong normative contextualism there is no context or set of conditions assumed to be 'normal' in the sense that it sets the default normative status of the trait. The only way in which the normative status of a trait can be determined is by looking at its relationship to the epistemic excellence of a particular agent given their epistemic predicament.

Strong normative contextualism gives rise to the tension outlined above. If the only normative status of a trait is the one determined by the value of the trait for an individual given their epistemic predicament, then traits conventionally understood as epistemic vices that are developed as effective coping strategies in hostile environments would be evaluated in a positive way because of their effectiveness and salience as responses to that environment for that agent. Because strong normative contextualism holds that the normative status of a trait is determined by its value for an agent

given their epistemic predicament, it does not classify traits developed in epistemic coping strategies as epistemic vices and therefore the result of epistemic corruption.

The alternative, *weak* normative contextualism, holds that some or all epistemic traits may have a default normative status – one that features in the way in which the trait is defined - but under certain conditions this may be revised (Kidd, 2020, p. 83). The status will be revised because, under certain circumstances and given the agents particular epistemic predicament, the trait is salient and effective for responding to the particular challenges faced by the agent when it was not under normal conditions or, if it is normally a salient and effective response, it is not under certain conditions.[66] The key difference between weak and strong normative contextualism is that, according to the former but not the latter, a normative evaluation (the default normative status) is built into the definition of the epistemic character trait – although it may be overwritten in certain contexts.

If we adopt weak normative contextualism, then the normative status referred to in (B1) and (B3) will be the contextualised status of the traits as epistemic virtues for those agents given their epistemic predicaments. On the other hand, the normative status referred to in (A1) and (A3) is the default normative status of the traits as epistemic vices. Therefore, weak normative contextualism offers a potential means to resolve the tension between (A3) and (B3). To defend this solution, I will say more about how the default status of an epistemic character trait arises.

One way in which a trait may have a default status is because we make general assumptions about the context in which it is being developed and exercised. That is, under 'normal' conditions a trait has certain effects and the default status is the status accrued in light of these effects. For example, under normal conditions closed-mindedness has bad epistemic effects, such as causing agents to dismiss relevant information or fail to revise their false beliefs and undermining the justification for their true beliefs (Cassam, 2019, p. 35; Battaly, 2020, pp. 26–31). Thus, it has the default status of being an epistemic vice – though this default status is overridden when the context differs from the range of 'normal' conditions in certain ways. For example, in an extremely hostile epistemic environment dominated by vast

[66] An example of weak normative contextualism is Battaly's (2018) argument that closed-mindedness is *standardly* an epistemic vice but under certain conditions is an epistemic virtue.

amounts of propaganda and misinformation, closed-minded agents may avoid being misled by unreliable sources or being overwhelmed by the task of sifting through information (Battaly, 2018, pp. 39–40). Therefore, for the agent under these conditions, closed-mindedness is an epistemic virtue (*ibid.*).

To the extent that the 'abnormal' conditions are temporary or limited in their scope we might still consider the default status to be relevant. Our recognition of closed-mindedness as an epistemic virtue in the propaganda-filled environment may be tempered by the consideration that, were the agent to develop closed-mindedness as a stable and entrenched part of her character, then when she finds herself in better epistemic environments, it would not serve her well epistemically.

But this does not help us in cases like Eddo-Lodge's where the hostile conditions in which she finds herself are so ubiquitous and entrenched that prospects for substantive change any time soon are slim. Racism is structurally and historically entrenched in British society and so the conditions which give rise to Eddo-Lodge's epistemic coping strategy are not 'abnormal'. Were we to imagine a world in which those conditions did not obtain, it would be sufficiently different from the actual world to make it no longer a relevant comparison. Having epistemic character traits that are unsuited to contexts where epistemic coping strategies are not needed does not make her worse off since it is highly unlikely that she would find herself in such a context. Therefore, we have no reason to consider the development of these traits to be damage to her epistemic character. The default status of traits arising from their effects under 'normal' conditions cannot be used to explain why epistemic coping strategies constitute a form of epistemic corruption.

3.4.2 Dual-level normative contextualism

The default status of a trait may still hold sway over our judgements if it is derived not from the effects of the trait under 'normal' conditions, but from its relation to the non-contextualised ideal of a flourishing life. Such an approach, which I call 'dual-level normative contextualism', involves evaluating epistemic character traits on two levels: a contextualised level reflecting the salience and effectiveness of the trait in responding to the agent's particular epistemic predicament and a non-contextualised level reflecting the relationship between the trait and the living of a 'flourishing' epistemic life.

The contextualised level is the level on which the default status of epistemic character traits can be revised if they play a role in allowing their bearers to cope with extreme or hostile environments. But on the non-contextualised level, on which epistemic character traits are assigned their default normative status, they are measured against the ideal of an epistemically flourishing life.

The two levels of evaluation play different roles in our thinking about epistemic character traits. We appeal to the contextualised level in order to understand the forces at play in character development. It helps us to understand why individuals choose or are influenced to develop certain traits in particular environments. It is unhelpful to assume that, for all individuals in their differing contexts, the same epistemic character trait will be equally as valuable for them. If we are to understand what it means to be an epistemic agent who does well given one's epistemic predicament, we must look to the contextualised level of normative evaluation.

Furthermore, we do on occasion appeal to the contextualised level of evaluation to appraise agents' epistemic character. For example, we might commend the development of a coping strategy. We also critique the futility of developing traits which are not helpful in an agent's epistemic predicament. These appraisals need not be as strong as praise and blame – we might recognise the necessity and utility of developing a coping strategy and in some sense commend the agent but would hesitate to describe this as praising them for developing the coping strategy.

The non-contextualised level informs our thinking about the kind of lives which we hope for agents to be able to live. The regret we may feel at the development of traits which are necessary given the context is derived from our non-contextualised evaluation of the traits.

It is this non-contextualised level, I propose, to which we refer when describing the development of a trait as character damage. In the case of epistemic coping strategies, we label them as epistemic character damage because they deviate from the kind of life that we hope for agents to lead. This is not a consideration of the kind of life which agents could or should lead given their particular circumstances, but the best kind of life which would be possible were circumstances conducive to it. After all, we deem environments to be corrupting because they lead epistemic character development away from the ideal.

To illustrate this dual-levelled approach to the evaluation of character traits it is useful to consider Lisa Tessman's (2005) account of *burdened virtues* – virtues which are 'disjointed from their bearer's own flourishing' (*ibid*., p.4).

Tessman's account itself does not involve a commitment to normative contextualism[67] but does demonstrate a two-fold evaluation of moral character traits on which dual-level normative contextualism can be modelled.

According to Tessman, burdened virtues are *virtues* because they are morally necessary for resisting oppression or practically necessary for surviving oppression. In the understanding of what it means to eradicate oppression there is implicit some understanding of what it means for all to have the ability, freedom and potential to live a good or flourishing life (*ibid*., p.108). Burdened virtues are *burdened* because they are incompatible with the conception of a good life which liberatory projects aim to make possible for all, either because they have no place in a post-oppression world, or because their exercise under conditions of oppression leaves lasting psychological scars which impede flourishing (*ibid*., p.108).

Among the burdened virtues are traits which have conventionally been classified as vices and are not part of a character profile constitutive of flourishing as implicit in liberatory projects but, nonetheless, are necessary for surviving and resisting oppression (*ibid*., p.114). Examples include traits which form part of a 'hard resolve against the oppression' and 'encourag[e] or cultivat[e] anger and rage in addition to withholding more sympathetic forms of attention.' (*ibid*., p.115).

Tessman's two-fold evaluation of traits can be combined with normative contextualism as follows. The status of burdened virtues as *virtues* occurs on the contextualised level of evaluation: it reflects the salience and effectiveness of a trait as a response to an oppressed agents' context. But their status as *burdened* occurs on the non-contextualised level of evaluation – when they are measured against the ideal of a flourishing life.

To parallel Tessman's terminology, we might say that the traits developed in epistemic coping strategies like Eddo-Lodge's are *burdened epistemic*

---

[67] Tessman (2005) identifies two types of burdened virtues: ordinary virtues which, under conditions of oppression become burdened (de-coupled from their bearer's flourishing); and virtues (such as virtuous anger) which are required only under conditions of oppression and are by nature burdened. In neither of these cases, does the status of the trait as a virtue or a vice differ depending on whether it is developed under conditions of oppression or not. Therefore, her account is consistent with the claim that the normative status of traits is not contextually dependent, and therefore she need not endorse normative contextualism.

virtues which are inconsistent with *epistemic flourishing*.[68] These epistemic character traits are commendable only in light of their role in allowing members of marginalised racial groups to avoid the depletion of their time and energy and the erosion of their agency through engagement in fruitless or draining conversations and their role in allowing for the creation of more productive dialogues and inquiry regarding racialised oppression.

On the contextualised level, the epistemic character traits have a positive normative status – as epistemic virtues – due to their value in resisting oppression and working towards liberatory goals. This allows for a much richer understanding than simply recognising their conventional status as vices. It recognises them as understandable and predictable responses to oppressive conditions, underpinned by commendable motivations such as a desire to protect one's epistemic agency and gain a deep understanding of systems of oppression and how to resist and overcome them.[69]

On the non-contextualised level, we also recognise the negative impact of these traits. They involve a sacrifice made by those resisting oppression and pursuing liberatory goals whereby, in the fight to enable all to flourish epistemically, they develop or are psychologically scarred in ways that get in the way of their own flourishing. By appealing to their default normative status on this non-contextualised level, we can understand the development of such epistemic virtues as epistemic character damage. Their development damages epistemic character in the sense that it prevents their bearer from developing the kind of epistemic character necessary for living an epistemically flourishing life.

Before moving on, it is worth saying a little about epistemic flourishing – although I will not attempt to fully expound the concept here. There has been very little discussion of flourishing in relation to epistemic virtues and vices (with the exception of Brogaard, 2012; Baril, 2019). Berit Brogaard describes epistemic flourishing as the 'epistemic equivalent of Aristotle's eudaimonia (well-being, flourishing, happiness)' (2012, p. 97). For Brogaard, there is much more to epistemic flourishing than just the acquisition of knowledge.[70] It is to live a life that is epistemically fulfilling.

---

[68] The possibility of burdened epistemic virtues has also been discussed by Alfano and Robinson (2017) who argue that gossip is one such example.

[69] Nancy Daukas calls epistemic character traits that contribute to resistance to oppression and achievement of liberatory goals 'liberatory epistemic virtues' (2019, p. 387).

[70] For a critique of virtue epistemology's focus on purely the acquisition of knowledge, see Manson (2012)

An important part of eudaimonia is membership in a community. In parallel, epistemic flourishing requires having good epistemic relationships with others such that they enable you to share knowledge back and forth, embark on successful shared epistemic projects, and support each other's epistemic flourishing.

There will be many different ways to live a good epistemic life, but in general, to epistemically flourish is to be oriented towards the world such that one is able to direct one's epistemic projects and engage in them in a way that is fulfilling and meaningful. This fulfilment might include an enjoyment of learning, appreciation of complex questions and satisfaction in solving them, wonder at the natural world, and curiosity about and interest in the world around (Baril, 2019, p. 78).

Being cynical of others, refusing to engage in enquiry and truncating one's testimony prevents one from fully engaging in successful and fulfilling inquiry. Though Eddo-Lodge's approach may direct energy towards more fruitful inquiry with those who share experiences of racism (which may be meaningful and fulfilling) such an agents' epistemic life as a whole will not be characterised by good epistemic relationships with others and full participation in the epistemic community. Therefore, the traits developed as part of the epistemic coping strategy are inconsistent with living an epistemic life which fully embodies this flourishing.

3.4.3 Resolving the tension

I will now argue that adopting dual-level normative contextualism resolves the tension because it allows us to hold both (A3) and (B3) simultaneously and consistently. This is because they each refer to different levels of normative evaluation.

As it stands, Kidd's definition of epistemic character damage does not specify on which level of evaluation the changes to epistemic character must involve the development of epistemic vice or hinder the development of epistemic virtues. That is, epistemic character damage could involve changes to character which erode or prevent the individual from developing either: (a) traits which are virtues when evaluated on the *contextualised* level; or (b) traits which are virtues when evaluated on the *non-contextualised* level.

If we want to resolve the tension, then it cannot be the former: we would need the traits to have both positive and negative normative statuses on the contextualised level of evaluation. Therefore, it must be the latter. So (A1)

(and our definition of epistemic character damage) should be revised as follows:

> (A1)* For epistemic character damage (and thus epistemic corruption) to occur there must be changes to epistemic character which, when evaluated on the non-contextualised level, involve the development of epistemic vice and/or hinder the development of epistemic virtue.[71]

Given this reformulation of (A1) to (A1)*, we can revise the conclusion (A3) as follows:

> (A3*) Therefore, the development of an epistemic coping strategy must include changes to epistemic character which, when evaluated on the non-contextualised level, involve the development of epistemic vice and/or hinder the development of epistemic virtue.

When we adopt dual-level normative contextualism we must revise (B1) slightly:

> (B1)* Dual-level normative contextualism: traits which are salient and effective means for some particular agent to respond to the particular challenges of their epistemic predicament are, for them given their epistemic predicament, epistemic virtues (and not epistemic vices) when evaluated on the contextualised level.

We therefore must revise the conclusion to the argument, (B3):

> (B3)* When evaluated on the contextualised level, the epistemic character traits developed as part of epistemic coping strategies are epistemic virtues (and

---

[71] Whether or not we use the term 'virtue' both to describe traits that are constitutive of epistemic flourishing and to describe traits that contribute to the achievement of liberatory goals is, I think, a terminological point regarding which I have no strong feelings either way. On one hand, we might reserve the terms 'epistemic virtue' and 'epistemic vice' for terms which are evaluated positively or negative, respectively, on the contextualised level and define epistemic character damage purely in terms of epistemic character traits which are consistent or inconsistent with epistemic flourishing. On the other hand, as I have done here, we may use 'epistemic virtue' and 'epistemic vice' for evaluations on both levels but specify in each instance to which level of evaluation we are referring.

not epistemic vices) for that particular agent in that context.

(B3)* refers only to the normative status of the traits developed as part of epistemic coping strategies on the contextualised level. Therefore, we no longer have any conflict between (A3)* and (B3)* and we can consistently hold them in relation to the same epistemic character traits.

3.5 Conclusion

In this chapter, I have demonstrated that there is a potential tension between two features of critical character epistemology. On one hand, according to normative contextualism, traits which are salient and effective ways for an agent to respond to their epistemic predicament are epistemic virtues. On the other hand, epistemic corruption requires either the development of epistemic vice or changes to character that hinder the development of epistemic virtue.

To avoid this tension, I have argued that we must adopt a particular form of normative contextualism – dual-level normative contextualism – according to which epistemic character traits are evaluated on two levels. On the contextualised level, traits are evaluated according to their effectiveness and salience for aiding agents in responding to the challenges of their particular epistemic predicament. This satisfies the motivation behind normative contextualism – the need to recognise the dependence of an epistemic character trait's value on the particular agent's epistemic predicament. On the non-contextualised level, epistemic character traits are measured against the ideal of an epistemically flourishing life. This ideal does not vary according to context. When epistemic character traits are inconsistent with or detract from epistemic flourishing, we regret their development and consider it damage to epistemic character. It is this negative normative evaluation on the non-contextualised level that justifies the classification of epistemic coping strategies as a form of epistemic corruption. As such, to develop a critical character epistemology which endorses both normative contextualism and Kidd's account of epistemic corruption – our normative contextualism must allow for dual-levelled normative evaluations. It serves to explain the uncomfortable ambivalence held towards may epistemic character traits given the complexities of surviving and coping under conditions of oppression.

My proposal of dual-level normative contextualism allows us to retain critical character epistemology's commitment to normative contextualism whilst still being able to isolate the distinctive damage done to epistemic

character when epistemic vices are developed. With regard to PECTs, the account demonstrates two things. Firstly, the development of vicious PECTs causes damage to epistemic character in the sense that their development takes students further from the kind of epistemic character profile which we would hope for them to develop – one which allows them to epistemically flourish. Secondly, this damage is independent of the value of the trait for the particular student given their epistemic predicament. Although the examples discussed in chapter 2 did not display this feature, a PECT may be valuable for a student, both in terms of aiding self-protection from harm and leading to greater attainment of epistemic goods, whilst still causing damage to epistemic character. In such a case, when evaluated on the contextual level, such traits may not in fact be epistemic vices, though when evaluated on the non-contextual level, they will still be epistemic vices. (For the sake of clarity, I will continue to refer to traits in the 'keep your head down' and 'just don't listen' clusters as 'vicious' PECTs.)

The development of any vicious PECT carries the cost of damage to epistemic character. In the next chapter, I will identify a second cost which can arise from the development of PECTs in the 'keep your head down' cluster.

Chapter 4

Constraining Epistemic Agency: Costs of PECTs from the 'keep your head down' cluster

In this chapter, I continue to examine the costs of developing vicious PECTs. In the previous chapter I focused on the costs associated with damage to epistemic character. I now turn to a different dimension along which the development of vicious PECTs can make an individual worse off – the constraint of epistemic agency. This is a cost which arises from the development of some, but not all vicious PECTs, and so in this chapter I will focus particularly on a PECT from the 'keep your head down' cluster – namely, intellectual servility.

I'll begin in §1 by presenting an account of epistemic agency according to which it is the possession and exercise of epistemic powers. I will argue (in §2) that many epistemic powers are interpersonal in nature and therefore, their possession is dependent on access to shared epistemic resources and appropriate uptake and recognition from one's epistemic community. In §3, I show why the ability to exercise epistemic powers is also required for epistemic agency and, as such, systematic barriers to the exercise of epistemic powers can constrain epistemic agency.

In §4, I offer a more detailed picture of intellectual servility and how it can be developed as a PECT. Then in §5, I argue that the development of intellectual servility constrains epistemic agency primarily through the creation of psychological barriers to the exercise of epistemic powers. Finally, in §6, I will argue that the constraint of epistemic agency and epistemic character damage constitute related but distinct costs, since each can occur independently of the other.

4.1 Epistemic Agency

As I will demonstrate below, there are a range of ways to understand epistemic agency depending on the explanatory role which the concept is

intended to play. This thesis aims to examine the development of students as knowers in a way that takes seriously the impacts of their social environments on their development. As such, a conception of epistemic agency that is most useful for this project is one that reflects the socially embedded nature of agency and allows us to view the epistemic agent as someone who acts within, affects, and is affected by her epistemic community. Accordingly, an adequate account of epistemic agency should reflect the fact that being an epistemic agent is to be a socially situated being, whose interactions with others are infused with relationships of differential levels of power.[72]

Epistemic agency is a subcategory of agency. Therefore, to define epistemic agency, I begin with a general notion of agency and then narrow it down to identify *epistemic* agency. To be an agent in general, is to be 'the kind of being who acts' and has the power to 'bring about certain effects through the exercise of that power' (Zagzebski, 1996, p. 142). Acting as an agent involves exerting power and exerting a power requires a certain level of reliability in bringing about the right kind of effect. For example, I cannot be said to have the power to move my limbs if I cannot reliably bring about the effect that my limbs move when I exercise the power. Possessing a power requires that my intentional exercise of that power reliably brings about in the right way the intended effect.[73] Were I to regularly fail to move my limbs when I intended to and attempted to bring about the movement of my limbs as a result of this intention, then I could not be said to possess the power to move my limbs (*ibid.* p.145).

Given this general definition of agency, let us consider what it means to be an *epistemic* agent. In general terms, to be an epistemic agent is to be the kind of being who actively and intentionally undertakes epistemic activities. For the purposes of understanding epistemic agency (in the context of this project), I propose that consider an epistemic activity as one that aims at producing an epistemic good. For instance, since knowledge is an epistemic good, an activity that aims at producing some piece of knowledge is an epistemic activity. Since knowing requires believing, believing is an epistemic activity.[74] But there are other activities that count

---

[72] Such an approach has been defended in relation to agency more broadly (see for example Mackenzie, 2008; Webster, 2021).

[73] For discussion of what it means for a power to bring about an intended effect in the right way and deviant causation in action theory, see for example (Bishop, 1989; Mele, 1997).

[74] There is some disagreement as to whether believing is genuine action – that is, the kind of action that can be freely performed by an agent as an exercise of their agency (Zagzebski,

as epistemic activities that go beyond believing. These include, but are not limited to: considering, evaluating, discovering, hypothesising, learning, teaching, investigating, testifying, persuading, challenging, and so on. Thus, central to epistemic agency is engaging in these (and other) epistemic activities. Since engaging in epistemic activities requires the powers to perform the actions involved, being an epistemic agent requires both the possession of these powers and the ability to exercise these powers. For each type of epistemic action, there is a corresponding epistemic power to perform the actions, such as the power to form a belief and the power to consider a proposition.

Given the broad range of epistemic activities, it follows, that the epistemic powers involved in epistemic agency are not limited to powers to form, suspend, and withhold beliefs. Focusing on the act of believing gives rise to interesting questions regarding the nature of believing as a mental action, but limiting our attention to the act of believing does not serve my project. If a concept of epistemic agency is to be useful in helping us to understand the development of students as knowers under conditions of oppression, then it must capture the whole range of ways in which they exercise epistemic agency through engagement in shared epistemic activities as well as their relationship to others in their epistemic community. These dimensions of epistemic agency extend beyond just forming and holding beliefs. For instance, one way to achieve the epistemic good of knowledge is via testimony. This means the epistemic power to testify and the power to consider a piece of testimony both count as epistemic powers (Lackey, 2021). In sum, I propose that epistemic agency is exercised when we exercise epistemic powers that are involved in engaging in epistemic activities where all activities that are aimed at producing epistemic goods count as epistemic activities.

Although 'epistemic activities' are defined here as activities in terms of epistemic goods, we should note that being an epistemic agent does *not* require that the epistemic activities one engages in are *successful* at producing epistemic goods.[75] That is, I am engaging in epistemic activity if

---

2001; Sosa, 2013; Olson, 2015) – action given that is not clear that we can believe at will or have control over which beliefs we form (see, for example, Williams, 1973; Bennett, 1990; Hieronymi, 2008; McHugh, 2013).

[75] This is contra Zagzebski's (2001) claim that powers relevant to epistemic agency are those powers to attain epistemic goods. Given the claim (which Zagzebski endorses) that having a power to *x* requires a certain degree of reliability in x-ing, then such a proposal

I am exercising my epistemic powers (say, to assess the sincerity of a piece of testimony) even though such exercise does not yield knowledge (because I fail to form a true belief because I judge the testifier to be insincere). I am also engaging in epistemic activity if I am exercising my epistemic powers even though I am unreliable in producing epistemic goods by exercising such powers. Suppose that I often fail to form true beliefs and instead often form false beliefs when hearing testimonies from members of particular groups because I wrongly judge them to be insincere. I would often fail to produce epistemic goods, but my act of assessing the sincerity of testifiers still is an exercise of epistemic agency, albeit a problematic or unfruitful one. I am, after all, engaging, actively and intentionally in an epistemic activity.

We should also note that epistemic powers are not limited to powers to perform actions which aim at their bearer's acquisition or retention of epistemic goods. For example, a teacher acts in support of the learning and inquiry of others. The way in which she conducts herself and exercises cognitive powers to facilitate her student's learning need not involve acquiring knowledge herself or reflectively endorsing any of her students' beliefs. The teacher is an epistemic agent regardless of whether she performs actions which aim at *her* attainment or retention of epistemic goods. Hence, epistemic agency is exercised when we exercise epistemic powers that are involved in engaging in activities that are aimed at *producing* epistemic goods whether those goods accrue to the agent or someone else.

One further qualification is that epistemic agency comes in degrees. There is a huge range of different epistemic activities in which we can engage and these can vary vastly in their complexity and the skills that are required to perform them. Being an epistemic agent does not require that an individual possess the powers to perform any and all epistemic actions. Different individuals will have different combinations and ranges of epistemic powers. A young child, for example, has the powers to form beliefs on the basis of others' testimony and to request information by asking questions, but they may lack the power to critically evaluate testimony for reliability and relevance because they have not yet developed the necessary cognitive skills. This does not mean that the young child is not an epistemic agent. They are an epistemic agent to the extent that they possess some epistemic

---

entails that being an epistemic agent would require a certain degree of reliability in attaining epistemic goods.

powers, but they have a lesser degree of agency than someone with a greater range of epistemic powers. This point is important since many of the ways in which epistemic agency is constrained are very often partial or limited to particular contexts or domains.

As we have seen, engaging in epistemic activities involves performing mental actions that can be done by the individual independent of others. These include reasoning, evaluating arguments, drawing conclusions, formulating and comparing ideas, hypothesising and so on. However, although an isolated individual who performs these activities is an epistemic agent, we should note that we are not born with these epistemic powers. Indeed, we are taught these powers either directly (say, in a critical thinking or logic classroom) or, much more pervasively, indirectly. The development of these powers relies on other epistemic agents as well as access to and use of shared epistemic resources (such as languages, shared concepts and communally recognised patterns of reasoning and inference).

Given that some of our epistemic activities take place in concert with others, they require epistemic interactions with others. These epistemic interactions include, but are not limited to: asking and answering questions, testifying and evaluating others' testimony, providing evidence, directing a shared inquiry, introducing or modifying shared concepts and definitions and more generally adding to or making use of shared knowledge.

But this is not the only social dimension of epistemic agency. This is because it is not only epistemic powers that are distributed unevenly in our communities; social powers are distributed unevenly in our communities given structures of social hierarchy (privilege and oppression).[76] Hence, our ability to participate in epistemic activities can be affected by our social powers. That is, our epistemic agency can be constrained by oppression.

Finally, let me note that I do not assume that all constraint of epistemic agency constitutes a wrong or injustice and I do not attempt to distinguish those which do from those that do not (beside some brief comments in §4.2.2). My task in this chapter is to identify a cost (in the form of a constraint on epistemic agency) that results from the development of some vicious PECTs. Doing so does not require determining which constraints of epistemic character constitute wrongs or injustices. Rather, it suffices to show that the constraint of epistemic agency (at the very least when it is caused by changes to epistemic character) is a harm to the individual

[76] See, for instance, Haslanger (2012).

because it makes them worse off than they would be otherwise. Epistemic agency, I take it, is good for an individual. The more epistemic agency an individual has the better, although very few, if any, individuals do not have their epistemic agency constrained in various ways. If *y* causes an individual's epistemic character to be constrained to a greater degree than it would be were *y* not the case and all else being equal, then *y* makes the individual worse off.

Given that the possession and exercise of epistemic powers is necessary to engage in epistemic activities, our epistemic agency can be constrained by limiting which epistemic powers we possess or by systematic barriers to the exercise of these powers. In §4.2, I will show how an individual's agency can be constrained when they lack epistemic powers because they cannot reliably bring about the intended effects of the exercise of that power. Then in §4.3, I will show how, even when individuals possess epistemic powers, their epistemic agency can be constrained by systematic barriers to the exercise of those powers.

4.2 The reliability of epistemic powers

A putative power would be unreliable if when attempting to exercise it and attempting to engage in epistemic activities, the individual regularly fails to do so in the way she intended.

Often individuals will have multiple intentions when attempting to perform an epistemic action. For example, when a speaker presents her testimony to an audience, she may very well intend that her audience be convinced by her testimony and form a belief on its basis. There is of course a power associated with this intention: the power to convince an audience to form a belief on the basis of your testimony. To possess the power to reliably convince audiences on the basis of one's testimony is not something which we would expect of individuals with a 'normal' degree of epistemic agency (since many testifiers do not convince their audience). But she also has a less ambitious intention simply to offer her testimony to her audience so that they may, if they are convinced, form a belief on its basis. It is the former power which is most relevant for understanding the extent to which her epistemic agency is constrained since it is an important power which is required for many everyday epistemic interactions.

I propose that being an epistemic agent (in the important sense) corresponds to the power to perform the action rather than the power to succeed in bringing about the desired effects of the action. The latter would be a high bar: a large proportion of people do not reliably succeed in

convincing people on the basis of their testimony. But in order to understand the role which epistemic agency plays in the social dynamics of epistemic communities, we must also be interested in who is able to offer their testimony (that is not to say that whose testimony is convincing is not an important issue – but even those who are not believed have a certain power to 'put something on the table'). The power to offer testimony is still a power to make changes to the epistemic community (as I will show in §4.2.2, it requires eliciting the appropriate uptake and recognition from members of the epistemic community) but is a lower bar than the power to convince others on the basis of testimony.

In this section, I will explore some of the conditions which must be met in order to possess the powers required for epistemic agency – including conditions regarding the recognition and respect they receive from other members of their epistemic community. Because of the interpersonal nature of many epistemic activities, the reliability of these powers is dependent to a large degree on one's relationships to others in the epistemic community. I identify two necessary conditions for possessing such epistemic powers: access to shared epistemic resources; and appropriate recognition and uptake from other members of the epistemic community.

4.2.1 Access to shared epistemic resources

Many epistemic activities are undertaken along with others and require membership in an epistemic community. Even if they don't involve direct interaction with others, the epistemic resources which they make use of – information, languages, conceptual tools – are shared resources to which we have access as members of an epistemic community (Dotson, 2014). In order to participate in shared epistemic activities – that is to receive and transmit knowledge and collaborate with others, we need to be able to communicate with others in our epistemic community. This requires access to shared epistemic resources and that those epistemic resources are adequate for communication.

The resources shared within an epistemic community include, shared knowledge, language and conceptual schema, mutually recognised patterns of inference and modes of inquiry. These might be concepts which are recognised and understood by all members (such as single-mother, woman or holiday); the language(s) which members use to communicate; knowledge that is shared by most, if not all, members of the community and which can be taken for granted in exchanges (for example, that Christmas is celebrated on 25th December, that there are 100 cm in a metre); and forms of inference or reasoning that are recognised as acceptable within the

community and can be used in exchanges (for example, that inductive reasoning based on empirical evidence is recognised to be a valid way of arriving at a conclusion), and the 'rules' for engaging in these types of reasoning.

For example, if someone finds themselves in an epistemic community whose language they do not know and they are unable to find an interpreter, they are significantly hindered in participating in shared epistemic activities because they cannot understand and make themselves understood. As much as they have the capacities in themselves to perform a wide range of epistemic activities such as testifying, receiving, and evaluating evidence, those capabilities will not correspond to reliable powers to perform those actions within that community if language barriers present a systematic obstacle to being understood and therefore performing such tasks successfully. For example, they cannot testify if they cannot make themselves understood when they speak.[77]

The individual need not lack epistemic agency entirely. They may find ways of communicating certain things for example through gestures – but the range and complexity of the epistemic actions that can be performed will be limited. As I noted above, epistemic agency comes in degrees. The individual who does not know the language of the epistemic community in which they find themselves will have a lower degree of epistemic agency in that context than someone who is able to communicate fluently.[78] Similar restrictions of agency can also occur to a lesser degree when some but not all epistemic resources are shared. For example, if an individual lacks access (perhaps through limited access to education) to some knowledge which is shared by the rest of the epistemic community, then they may struggle to communicate effectively with others when that knowledge is assumed.

---

[77] Kristie Dotson (2014, p. 115) takes epistemic agency to require the ability to make use of shared epistemic resources *persuasively* in the sense that one must be able to make themselves understood through the use of those resources. Dotson includes this requirement in order to rule out cases in which an individual is able to make use of shared resources but is deemed (due to prejudice) to be unintelligible. It should not be interpreted to mean that epistemic agents must be able to use shared resources to successfully persuade others, since this would place too high a bar on epistemic agency for the reasons discussed above.

[78] Their low level of epistemic agency in this context is also consistent with their possession of a high level of epistemic agency in the context of an epistemic community whose language they know.

However, having access to shared resources does not entail that individuals are able to communicate effectively. Sometimes, the shared resources to which individuals have access will be inadequate for communicating some important aspect of their experience. This occurs in cases of what Fricker (2007) terms 'hermeneutical injustice'. Hermeneutical injustice occurs when a significant area of an individual's experience is obscured from collective understanding because of gaps in the shared hermeneutical resources (arising from identity prejudice in the production of those resources) (Fricker, 2007, pp. 154–155).

Fricker illustrates this in the case of Carmita Wood who experiences what we now call 'sexual harassment' from her boss and as a result quits her job. When she applied for unemployment benefits, she was asked to give a reason for leaving her job, but she found herself unable to explain to the investigator what had happened. At the time, the term 'sexual harassment' did not exist as a recognised means to describe what had happened to Wood. The conceptual resources which Wood had at her disposal to communicate to the investigator what she had experienced were limited to concepts such as 'bullying' or 'flirting', none of which were adequate. The reason was listed as 'personal' and her claim was denied as a result (*ibid*., p.150).

A gap in the concepts and language shared by Wood and the investigator (as well as other members of their epistemic community) meant that there were no adequate resources on which Wood could draw in order to communicate the nature of her experience. This need not entail that she herself lacked a name or concept to describe her experience (Dotson, 2014, p. 127), but any terms or concepts which she did possess were not recognised or known by the epistemic community at large[79]. This further impacts her ability to contribute to shared knowledge production. For example, she would struggle to share with others and make it common knowledge that her former boss acted inappropriately, because she cannot make others understand what happened to her (*ibid*.).

A lack of access to adequate shared resources can make individuals unreliable in performing the epistemic actions they intend to perform because they are misunderstood, misunderstand others or are otherwise

[79] In some cases, adequate epistemic resources may exist and be shared within epistemic sub-communities, but not by other sub-communities or the epistemic community as a whole. In these cases, the individual's epistemic agency may be constrained in relation to the communities which lack the adequate shared epistemic resources and not in relation to the epistemic sub-communities which possess them.

hindered in their communication. Because they are unreliable in achieving their intended effects, they lack the powers to perform these actions and their epistemic agency is constrained. Therefore, the lack of access to adequate shared epistemic resources constrains epistemic agency.

The extent to which epistemic agency is constrained by inadequacies in the shared epistemic resources also depends on the ease with which those inadequacies are remedied. Had Wood (and other women in similar situations) been able to easily introduce a term which filled the gap, then the constraint on her epistemic agency would have been relatively short-lived. However, if there is resistance to such changes from the epistemic community (through, for example, active ignorance (see for example Tuana, 2006; Mills, 2007)) then the constraint of epistemic agency would be more extensive.[80] As Kristie Dotson (2014, p. 131) notes, epistemic communities may be resistant to changes which go beyond shared social imaginaries, particularly if they present a challenge to dominant and privileged perspectives.

4.2.2 Uptake and recognition

Even when individuals have access to adequate shared epistemic resources, their epistemic agency may still be constrained because they are not reliably able to perform epistemic actions as they intend. Having the power to perform an action such as offering testimony relies on receiving the appropriate uptake from one's audience[81]. Consider the case of Marge Sherwood, a character from Anthony Minghella's screenplay, *The Talented Mr Ripley*:

---

[80] As Fricker argues in her account of hermeneutical injustice, a gap in the hermeneutical resources is not itself sufficient for hermeneutical injustice. Hermeneutical injustice occurs when an individual is disadvantaged by a gap in the hermeneutical resources which comes about because members of a particular group are marginalised in the process by which shared hermeneutical resources are developed (2007, p. 152ff.).

[81] I borrow the notion of uptake from the work of J.L. Austin (1975) and as developed by Langton and Hornsby in their work on illocutionary silencing. The performance of an illocutionary act requires uptake from an audience. This uptake consists in the speaker 'being taken to be performing the very illocutionary act that, in being so taken, she (the speaker) is performing'. (Hornsby and Langton, 1998, p. 25). Similarly, the performance of an epistemic action of the kind which involves interaction and communication with others, requires uptake from other members of the epistemic community. I take it that this uptake requires more than just that members of the epistemic community understands that the individual is attempting to perform the intended act, but also that they acknowledge the act as such.

SHERWOOD: Marge Sherwood's fiancé Dickie has gone missing and she suspects that his friend, Tom Ripley, is involved. As his family attempt to find out what has happened to him, Marge is asked for information such as when she last saw Dickie and how he was acting on the days preceding his disappearance. She attempts to get more involved by expressing and explaining her suspicion of Ripley. However, she is repeatedly dismissed and her suspicion is regarded as unjustified and inconsequential. At one point Dickie's father, Herbert Greenleaf tells her, "Marge, there's female intuition, and then there are facts" (Fricker, 2007, p. 9)

Sherwood attempts to participate in the investigation into Dickie's disappearance by attempting to direct the course of the investigation and suggest a line of inquiry. However, her attempt to do this is unsuccessful because she is not given the right kind of uptake from Greenleaf. Rather than recognise her to be suggesting a line of inquiry and attempting to guide the direction of the investigation, he responds to her as though she has merely expressed a 'bad feeling' about Ripley.[82] Regardless of Sherwood's intellectual capabilities and access to the shared epistemic resources required for effective communication, she is not able to successfully perform the action she intends to perform.

If Sherwood is repeatedly unable to perform this action, then she lacks the epistemic power to attempt to direct the course of the investigation and suggest a line of inquiry. Thus, her epistemic agency is constrained by the lack of this epistemic power. It could be objected that Sherwood does not lack the power – since she is able to perform the right kind of action despite failing to gain the correct kind of uptake. That is, she is able to do the requisite intellectual work and speak the right words in a way and context in which in principle, they could be understood and constitute an attempt to direct the course of the investigation. It is true that we can define such a power which Sherwood does in fact possess. However, as epistemic agents do not function in isolation from others, the powers which are relevant for a rich understanding of epistemic agency are those powers to engage in epistemic activities that involve engagement with their epistemic

[82] I do not think that it matters whether Greenleaf realises what Sherwood's intention is and undermines her intentionally, or whether he genuinely believes that she is merely expressing a bad feeling about Ripley. In both cases, Sherwood fails to get the appropriate uptake.

community. To say that Sherwood does possess the power to perform the right kinds of actions to suggest a line of inquiry regardless of whether she gets the correct uptake tells us very little about her standing (or lack of) as an epistemic agent. Therefore, Sherwood can be said to lack a power relevant to epistemic agency when she repeatedly fails to get the appropriate uptake for her attempts to direct the course of the investigation.

Therefore, the ability to perform a certain communicative action – such as asking a question or making a suggestion about the direction an inquiry should take – is directly related to the kinds of epistemic actions which they can perform because, as noted above, the majority of epistemic actions are interpersonal and therefore require communication. If the individual fails to communicate in the way intended, then they cannot perform the intended epistemic action. So far as uptake is required for communicative acts, it is also required for interpersonal epistemic actions.

One reason why uptake may not be received is because the epistemic action being attempted falls under the purview of a particular role within inquiry and the audience do not consider the individual as able or entitled to occupy that role. There are variety of roles that an individual can play within inquiry. Some involve being a source of information from which enquirers can gain information like a witness giving a statement. Others involve actively acquiring, possessing and transmitting knowledge: for example, the detective who asks questions, compiles and reviews evidence and tries to find answers. The role an individual occupies within a particular exchange will determine what they can or are expected to do within that exchange – including the demands they can make of others, the demands that can be made of them and the kind of uptake they can expect to receive from their interlocutors.

Different roles will be appropriate for different individuals at different times and in different contexts. Often one's role will be determined by one's epistemic standing in relation to the subject of inquiry – that is, how well one is positioned to perform certain kinds of epistemic actions. In other cases, individuals occupy or are restricted to certain roles for pragmatic reasons – to aid the division of epistemic labour. Consider for example the assignment of roles within a courtroom. Witnesses are expected to provide factual testimony for consideration by the jury. Any attempt from them to evaluate the evidence presented by others and arrive at a verdict will be dismissed or ignored because it falls outside of their prescribed role.

To some extent the witness's epistemic agency is constrained in the courtroom because she is unable to ask questions of others in the courtroom

(any attempts to do so would fail to get the correct uptake). The extent to which this constitutes a constraint of her epistemic agency may depend on whether she has the means to perform equivalent actions in alternative settings. Furthermore, as noted above, not all constraints of epistemic agency need be wrongs or injustices.

Medina argues that, to be truly equal, the relationships between members of the epistemic community must be reciprocal. This requires that the roles occupied by each are in principle reversible (Medina, 2013, p. 93).[83] Roles will not be in principle reversible when an individual is perceived not to be able or entitled to occupy a particular role on the basis of prejudicial identity stereotypes. This is what happens in Sherwood's case. Given her involvement in the whole affair, Sherwood is in a good epistemic position (or at least as good a position as Greenleaf) to be able to direct the course of the investigation into Dickie's disappearance. Therefore, it is appropriate for her to occupy the role of inquirer: a role in which she plays an active part in directing the course of the inquiry and suggesting new avenues for investigation based on the available evidence.

However, Greenleaf perceives her as limited to the role of informant: contributing to the inquiry only through the offering of factual testimony (such as when she last saw Dickie) because he believes that occupying the role of inquirer is beyond her capabilities given her gender. Greenleaf does not recognise Sherwood as able and entitled to occupy the same roles as him because she is a woman. He views her as limited to the role of being an informant – providing factual information about Dickie. He does not see her as able to occupy a role in which she makes conjectures and directs the course of the investigation. Therefore, the roles are not in principle reversible.

As a result, she gains appropriate uptake for only some of her attempted epistemic actions. When she provides information and answers questions about where and when she last saw Dickie, her interlocutors give her the appropriate recognition and she succeeds in performing the actions she intended. It is only when she exceeds the role expected of her because of

[83] Examples such as the courtroom clearly fail to meet this condition since the roles are not in principle reversible. The role which the witness occupies is prescribed by the institutional norms. Even if she is in a good epistemic position to guide the course of the inquiry, any attempt to do so would not be given the correct uptake. I take it that the restriction of agency in this case is purported to be justified as a means to enable the process of the trial.

prejudices that she does not get recognition from her interlocutors (Pohlhaus Jr, 2014, pp. 105–16).

Because the possession of the powers to perform interpersonal epistemic actions is dependent on receiving appropriate uptake from one's audience or interlocutors, the possession of these powers will be relative to context. It is possible to receive appropriate recognition from one audience or within one particular epistemic sub-community whilst lacking it from other audiences or epistemic sub-communities. For example, Sherwood may share her suspicions and evidence with her friends and be taken seriously such that she is able to direct their conversation regarding Dickie's disappearance towards Ripley. Therefore, in that context, she does possess the power to perform these actions even though she does not in the context of her conversation with Greenleaf.

Similarly, the same individual may get the appropriate uptake when performing a certain type of epistemic action in one domain or context but not when they perform the same act in another domain or context. When speaking with different audiences or in different settings, the same individual may get different levels of uptake and recognition.

Often, epistemic injustice (Fricker, 2007) and epistemic oppression (Dotson, 2014) are said to restrict or undermine epistemic agency in virtue of the failure to attain appropriate recognition and uptake. Such statements are often made without reference to degree or context. But it is very rarely the case that epistemic oppression and epistemic injustice restrict agency completely, in every single context or in relation to every epistemic sub-community. Instead, the non-relativised claims about epistemic oppression and injustice constraining agency can be seen as generalisations over a whole range of different contexts.

Sometimes these generalisations may weight certain contexts more than others because of their importance. For example, imagine a woman who generally does not encounter gender-based prejudice in her daily life and is taken seriously and respected as a knower in a range of domains, including her personal relationships and her career. However, when seeking treatment for a long-term health condition, she finds that she is not taken seriously. She has developed a great deal of understanding of her condition and attempts to steer the doctors' attention to certain avenues for diagnosis, but they regard her as being limited to the role of providing information regarding her physical symptoms. Therefore, in this context, she lacks the power to occupy the role of investigator because she lacks recognition when she attempts to occupy the role. This context is just one among many, but it

has a great deal of importance. Firstly, in terms of the practical implications of whether she receives the correct diagnosis and treatment, but also in terms of her sense of autonomy, control over and knowledge of her own body.

4.3 Exercise of powers

I now turn to a different way in which agency can be constrained. Even when an individual possesses an epistemic power in the sense that, were they to exercise it, it would be reliably successful in bringing about the intended effect, there may still be barriers to the exercise of that power. That is, the individuals possess the power but are not able to exercise it.

The difference between failing to possess a power and possessing it but being unable to exercise it, is that the latter is not a result of a lack of reliability. Possessing a power to *x* does not require that the power be exercised – that is that an attempt to *x* be made. Instead, possessing the power to *x* requires that when attempts to *x* are made, they are reliably successful in performing *x* as intended. If the power to *x* is not exercised, no attempt is made to *x*. When a power to x is possessed but cannot be exercised, there are systematic barriers which prevent the individual from making an attempt to x. I will give some examples below of what these systematic barriers may be.

To possess epistemic agency an individual need not constantly be exercising epistemic powers. I do not cease to be an epistemic agent when I am asleep or I am not participating in some kind of epistemic activity. However, an individual who is systematically unable to exercise epistemic powers does not possess epistemic agency in any meaningful sense. For example, consider the fact that education aims for not just students' possession of epistemic powers, but also that students recognise that they possess those powers and have the confidence to exercise them. A student who possesses a range of epistemic powers but would never exercise them because they did not recognise that they were capable of exercising them has a lesser degree of epistemic agency than a student who possesses the same epistemic powers and exercises them. Therefore, epistemic agency is constrained by the presence of systematic barriers to the exercise of epistemic powers.

To illustrate and further motivate this claim, consider the following variation on SHERWOOD:

SHERWOOD*: Sherwood takes to heart Herbert Greenleaf's comments about her being guided only by intuition and not being as capable as him (and men in general) or being objective and rational. She comes to believe

> that it is not appropriate for her to attempt to direct the course of the investigation by drawing attention to Ripley's involvement because she believes that she lacks the abilities and it is not her place to do so. She gets into a conversation with Dickie's brother Bob about Dickie's disappearance. Bob does not share his father's prejudice and recognises Sherwood as being in a good position to have valuable insights into what happened and suggest lines of inquiry. Were she to draw his attention to Ripley's suspicious behaviour he would take this suggestion seriously and give it the appropriate uptake. However, Sherwood no longer sees this as something she is in a position to do and refrains from attempting to suggest this line of inquiry.

Because Bob would give Sherwood the appropriate uptake and recognition, she does have the power to attempt to direct the course of the inquiry in this context. However, her attitudes towards herself prevent her from exercising that power. These negative attitudes towards her competency as a knower create a systematic barrier to the exercise of the power by causing her to conceive of it as beyond her abilities. This makes it very unlikely that she would exercise the power.[84]

There are often reasons for which we do not exercise our epistemic powers in some instances. For example, it may not be the appropriate time, we may not want to, or it may be the wrong setting. This is consistent with epistemic agency. But epistemic agency is constrained when there are *systematic* barriers which prevent the individual from attempting to perform some epistemic action. These barriers are systematic when they are persistent and occur in a wide range of contexts and domains. Such systematic barriers will make it very unlikely that the individual will ever perform the epistemic action because they will be prevented from *attempting* to perform the action. The action is not part of their agential repertoire.

In *Sherwood** the systematic barrier to the exercise of certain powers relates to Sherwood's attitudes towards her competency and status as a

---

[84] See also Paul Benson's (1994) discussion of an example from the 1944 film *Gaslight*. Discussing agency more broadly, Benson uses this case to argue that agency can be lost when an individual loses sense of herself as a worthy agent and does not trust herself to govern her conduct (ibid., p.657).

knower. Whilst these barriers are 'psychological', their source or cause need not be purely individual – they can have external causes. I will return to the relationship between psychological barriers to the exercise of powers and the internalisation of prejudice and oppression later when I discuss the impact of intellectual servility on epistemic agency in §4.5.

The particular attitudes that create systematic barriers to the exercise of epistemic powers are Sherwood's lack of self-trust and her lack of self-respect. They each prevent her from exercising the relevant power in slightly different ways. She lacks self-trust because she believes that failings of her cognitive abilities make her unreliable as a source of knowledge. For example, she might come to echo Greenleaf's sentiment that her suspicion regarding Ripley is based on unsupported intuition. She comes to believe that she is unreliable as a knower and therefore that by attempting to direct the course of inquiry she is more likely to hinder the investigation than aid it. Therefore, given her epistemic goal of wanting to find out what happened to Dickie, she chooses not to participate in this way.

She also lacks self-respect as a knower because she does not recognise herself to be entitled to perform this kind of action. To lack self-respect is to fail to regard oneself as equal to others and entitled to hold them to account (Mackenzie, 2014, p. 62). Lacking self-respect as a knower involves failing to regard oneself as having the appropriate standing to perform certain kinds of epistemic actions.

The erosion of epistemic agency through systematic barriers to the exercise of powers also comes in degrees. Sherwood's agency is limited because her lack of self-trust and self-respect prevent her from exercising her power to attempt to direct the course of the investigation, but it does not prevent her from exercising other powers such as the power to offer up factual testimony regarding when she last saw Dickie. The degree of epistemic agency which an individual possesses varies in relation to the number and range of epistemic powers which they are able to exercise.

I have established two conditions of epistemic agency: the possession of epistemic powers and the ability to exercise those powers. The lack of either of these two things constitutes a constraint of epistemic agency. I will now apply my account to show how the development of PECTs from the 'keep your head down' cluster can constrain epistemic agency.

4.4 Intellectual Servility as a PECT

In chapter 2 I identified a cluster of PECTs which protected students by helping them to 'keep their head down' and avoid harms which arise from

confrontation and hostility. One of the PECTs within this cluster is the vice of intellectual servility. In this section, I will begin by offering a more detailed picture of this vice and how it can be developed and function as a PECT. Then I argue that the development of intellectual servility can constrain epistemic agency primarily through the creation of psychological barriers to the exercise of epistemic powers.

The intellectually servile person is one who doubts and tends to belittle their own competencies and abilities. They tend to emphasise their weaknesses and minimise their strengths – attributing their successes to luck whilst attributing their failures to personal faults and failings. This sense of incompetency or weakness manifests in a tendency to compare themselves unfavourably to others – regarding others as their intellectual superiors and therefore as epistemically authoritative (Tanesini, 2018, pp. 28–29, 2021, p. 111).

Central to the vice of intellectual servility is a sense of shame for being less able than others. Therefore, the intellectually servile person is often deferent and feels a sense of shame for their lack of ability compared to others (*ibid*.). This can result in a feeling of anxiousness and worthlessness at the thought that they fail to meet the standards for membership and participation in the epistemic community. The intellectually servile person will accept and think themselves deserving of inferior status within the epistemic community.

Rather than resenting negative and discriminatory treatment, they attempt to ingratiate themselves with those in positions of power by conforming to expectations and suffering gladly. Within inquiry they are much more likely to defer to the views of those they regard as being intellectually authoritative, parrot their views and refrain from offering challenges to accepted views because they do not consider themselves qualified to do so. Intellectually servile individuals will lack assertiveness and will refrain from taking a directive role in steering the direction of inquiry, challenging paradigms, contributing to, developing or sharing subversive or norm-challenging knowledge or perspectives. They do not consider themselves as capable of adding anything new to inquiry and so see their role as largely limited to learning from others.

Intellectual servility is often adopted as a survival strategy under conditions of oppression (Tanesini, 2018, pp. 29–30, 2021, pp. 112–113). Members of oppressed groups who are at constant risk of harm, further exclusion or hostility, may mitigate those harms by ingratiating themselves with those in positions of power.

> It is a strategy aimed at minimizing harm in conditions of oppression which, however, achieves its goal at a high psychological cost since it promotes the development of a vicious trait of character. It is the strategy of accepting one's lower social status and seeking the approval of those who subordinate one by humiliating oneself, whilst praising, and parroting them. It trades off social acceptance for self-esteem, dignity, and self-respect. (Tanesini, 2018, p. 30)

As such, intellectual servility can function as a PECT by providing a strategy for avoiding harms which may arise from challenging oppressive norms, speaking out and 'putting one's head above the parapet'. This may be done intentionally, as a conscious trade-off between self-respect and some respite from hostility. Or it could be developed unconsciously, as a learned behaviour that is developed after attempts to stand up for oneself and challenge stereotypes end badly and result in harm.

4.5 The impacts of intellectual servility on epistemic agency

As I noted in chapter 2, the development of vicious PECTs is costly. One of the ways in which it is costly (as discussed in chapter 3) is that it constitutes epistemic character damage – it involves a change in character that takes the individual further from epistemic flourishing. I will argue that there is a second cost – distinct from epistemic character damage – that arises from the development of intellectual servility as a PECT. This further cost is the constraint of epistemic agency. In the rest of the chapter I will show how the development of intellectual servility constrains epistemic agency. Then I will show how the constraint of epistemic agency and epistemic character damage are distinct and can occur independently of each other.

Whilst I will argue that the development of intellectual servility as a PECT can constrain epistemic agency, the connection between the two is by no means straightforward. I will argue that under certain conditions the development of intellectual servility can *protect* epistemic agency because developing the trait prevents individuals from being entirely excluded from the epistemic community (and thereby having their epistemic agency heavily constrained).

There are two ways in which the development of intellectual servility as a PECT can constrain epistemic agency. The first – the creation of systematic barriers to the exercise of epistemic powers – is a relatively straightforward relationship and occurs in most, if not all, instances where intellectual servility is developed as a PECT.

I argued above that epistemic agency is constrained by psychological barriers to the exercise of epistemic powers such as a lack of self-trust and lack of self-respect. Intellectual servility involves negative attitudes towards one's cognitive capacities and competency as a knower (Tanesini, 2018, p. 30). These are precisely the kinds of self-regarding attitudes that act as systematic barriers to the exercise of some epistemic powers. Given their lack of trust in their reliability or their perception of themselves as 'unworthy', the intellectually servile person will hold back from engaging in certain kinds of epistemic actions or occupying certain roles within inquiry. This constitutes a psychological barrier to the exercise of the powers to perform the epistemic actions or occupy the epistemic roles in question. Therefore, by creating these psychological barriers to the exercise of epistemic powers, the development of intellectual servility constrains epistemic agency.

The second way in which the development of intellectual servility as a PECT can constrain epistemic agency is through its tendency to perpetuate oppressive norms which deny individuals the ability to make use of shared epistemic resources and gain the appropriate recognition and uptake from their epistemic community. However, this relationship is not straightforward, and in some cases, the development of intellectual servility as a PECT will increase, rather than constrain epistemic agency.

One set of conditions under which intellectual servility may be developed as a PECT is one in which oppressive norms and stereotypes cast the individual as being unable or unworthy to occupy certain roles within inquiry or perform certain kinds of epistemic actions. These stereotypes and norms may go hand in hand with denial of access to epistemic resources such as access to education, access to information, and a place at the table in certain conversations. Such conditions themselves will tend to constrain epistemic agency both by denying individuals the appropriate uptake and recognition they require to perform certain epistemic actions and by denying them access to shared epistemic resources.

The development of intellectual servility can further entrench this constraint of epistemic agency. By conforming to oppressive stereotypes and norms, they may become reinforced or more widespread. Individuals are more widely perceived to be incompetent or passive knowers if they present themselves in such a way as to appear to conform to stereotypes and set low expectations.

Intellectually servile individuals can also self-enforce these norms and choose not to access educational opportunities or information which they

don't perceive as 'for them'.[85] These choices will then impact which shared epistemic resources they have at their disposal and therefore what they can make use of when they participate in shared epistemic activities. For the reasons set out in §2.1, limited access to these resources can prevent them from possessing epistemic powers and therefore constrain their epistemic agency.

This is not a necessary effect of the development of intellectual servility as a PECT. In fact, in some cases, it could result in the individual's epistemic agency being less constrained than it otherwise would have been were she not intellectually servile. This is because, by conforming to norms and stereotypes, she gains greater acceptance into the epistemic community. Were she not to conform, she may face total or near total exclusion from the epistemic community, such that she has very little access to epistemic resources and gets little or no recognition and uptake from others. This would severely limit the ways in which she could participate in shared epistemic activities and therefore would severely constrain her epistemic agency. However, by developing intellectual servility as a means to conform to oppressive norms and stereotypes, she may avoid this kind of exclusion. Her epistemic agency will still be constrained because she will be limited in the range of epistemic actions for which she can gain appropriate uptake from her community or have access only to a limited range of epistemic resources. Nonetheless, she will be able to gain appropriate uptake for some epistemic actions and have access to some epistemic resources. In this way, she may possess greater range of epistemic powers if she develops intellectual servility as a PECT than she would if she did not and was excluded from the epistemic community. However, this picture will be further complicated by the fact that, given her intellectual servility, there may be systematic barriers to her exercise of these epistemic powers (for the reasons argued above).

The extent to which intellectual servility can protect epistemic agency in the way described will depend on whether failing to adopt intellectual servility would lead the individual to be excluded from the epistemic community in such a way that would prevent her from possessing epistemic powers. For example, in the case of Elena discussed in chapter 2, the hostility which she faces within the classroom (or perhaps even multiple educational settings) may not amount to exclusion from the epistemic

---

[85] When developed under conditions of oppression, intellectual servility is a form of *psychological* or *internalised oppression* (Bartky, 1990).

community. The hostility she faces if she does not conform to oppressive stereotypes and norms by developing intellectual servility need not make her unable to gain correct uptake for her epistemic actions or access shared epistemic resources to a large enough extent that she fails to possess epistemic powers and therefore, has her epistemic agency constrained. This does not mean that intellectual servility is not still salient as a PECT for Elena because it can protect her from other harms arising from this hostility. If her epistemic agency is not constrained when she does not develop intellectual servility as a PECT, then it is not the case that developing intellectual servility protects her epistemic agency. What I have shown, however, is that under certain circumstances the development of intellectual servility can lead an individual to have a greater degree of epistemic agency than they would have otherwise.

4.6 Agency and Character

I have identified two necessary conditions for epistemic agency and shown how the development of intellectual servility as a PECT can constrain epistemic agency. In this section, I now change my focus and will examine the relationship between the constraint of epistemic agency and epistemic character damage. I will argue that although epistemic character damage and the constraint of epistemic agency often go hand in hand, they are conceptually distinct. I will do this by showing that (a) damage to agency can occur without character damage and (b) character damage can occur without damage to agency.

Firstly, an individuals' agency can be constrained whilst their epistemic character remains unchanged. For example, many forms of epistemic oppression can, in principle, constrain agency via their relationships to other members of the epistemic community – through a failure of recognition or the affordance of a less than equal status in the epistemic community – without impacting their epistemic character. If an individual has already developed a virtuous epistemic character but then has her epistemic agency constrained through systematic epistemic oppression, her virtuous epistemic character could remain unchanged despite the constraint of epistemic agency. She may continue to possess whatever epistemic virtues she had despite being unable to exercise those virtues.

For example, a woman who starts a new job and is excluded from all important conversations and decision-making processes by her colleagues because they view women as less capable in doing their job, has her epistemic agency damaged in this context. However, it is possible for her to retain the epistemic character she had developed beforehand, at least for a

short period. She may continue to be intellectually diligent and persevering in the sense either that, were she be given the opportunity to participate, she would exercise these virtues, or she still attempts to exercise them despite the way in which she is treated.

Of course, in the actual world, the long-term effects of significant constraints on epistemic agency often impact epistemic character. This is one of the harms of testimonial injustice noted by Fricker (2007, pp. 47–48). When individuals are prevented from participating fully in their epistemic communities, they are less likely to develop and strengthen the skills and habits that make them epistemically virtuous. The negative stereotypes and oppressive norms underlying epistemic oppression can also be internalised, promoting the development of traits such as intellectual servility and timidity. I will return to the role of epistemic agency in the development of virtuous epistemic character below.

On the other hand, damage to epistemic character does not necessarily entail the constraint of epistemic agency. There are a number of vices which, when exercised, require a high degree of epistemic agency. One example is the vice of epistemic arrogance – an over confidence in one's beliefs and epistemic abilities (often coupled with an under-confidence in others' beliefs and abilities) which leads to a tendency to ignore one's weaknesses, rely too heavily on oneself, put down others, underestimate their abilities and exaggerate their weaknesses (Tanesini, 2016, 2020). This will involve a large amount of active engagement in epistemic activities such as giving and evaluating testimony, collecting evidence, and directing inquiry.

The ability to engage in such activities demonstrates the exercise of powers – even though these powers may not succeed in getting to the truth (since the arrogant may tend to stubbornly hold onto their own false beliefs rather than consult the expertise of others) they succeed in performing a range of epistemic actions. They also succeed in making use of shared epistemic resources. As Tanesini (2020) and Medina (2013) both argue, intellectual arrogance is very often a vice of the privileged. Privileged intellectually arrogant individuals will likely be the kind of person who is recognised to be a member of the epistemic community and granted access to the resources, given appropriate recognition and able to occupy the full range of roles within inquiry. They also possess the necessary self-regarding attitudes: they have a high (perhaps too high) level of self-trust and regard themselves as being worthy members of the epistemic community capable of occupying all roles within inquiry.

There may be cases in which epistemically vicious individuals are excluded from the epistemic community or are denied appropriate recognition and uptake because other members of their community recognise and deplore their vices. For example, an extremely arrogant individual may be dismissed, ignored and excluded from certain shared epistemic activities because people either recognise that their arrogance makes them highly unreliable as knowers and likely to hinder rather than aid inquiry, or they simply find them insufferable. In principle, this does have the potential to constrain epistemic agency if the arrogant individual is unable to perform certain kinds of epistemic actions or occupy certain roles within inquiry. However, the likelihood that such exclusion will be widespread and systematic enough is often relatively low. The constraints on epistemic agency for such an arrogant individual are likely to be relative to very localised contexts (particular sub-communities in which their arrogance is not tolerated). Overall, it is much less likely that the development of intellectual arrogance would lead to the constraint of epistemic agency than the development of intellectual servility.

I have shown that damage to epistemic character and the constraint of epistemic agency are distinct in that each can occur independently of the other. That being said, there are important connections between epistemic character and epistemic agency. I will discuss two such connections here: the need for epistemic agency in order to develop virtuous epistemic character, and the effect of internalised oppression on both agency and character.

Firstly, epistemic agency is required for the development of virtuous epistemic character. Epistemic character relates to the way in which an individual engages in epistemic activities. The epistemically virtuous person engages in them in an effective way. But epistemic virtues must be practiced in order to become entrenched features of epistemic character. It is not possible to practice virtuous traits, and therefore develop epistemic virtues, if one does not or is unable to engage in epistemic activities. If epistemic oppression heavily constrains epistemic agency early on, virtuous epistemic character may not have a chance to develop.

Even when epistemic oppression occurs at a later stage when agency has been developed to some extent, the inability to engage in a wider range of epistemic activities may limit the extent or range of virtues that are developed. For example, if an individual lacks the ability to immerse themselves in, direct, and reap the rewards of in-depth inquiry, they may not have the opportunity to develop the epistemic virtue of curiosity to a high degree. Epistemic virtues may also degrade over time if not practiced.

However, this does not necessarily follow from the constraint of epistemic agency – as argued above, it is possible for epistemic character to remain undamaged when epistemic agency is constrained.

Secondly, epistemic character damage and epistemic agency often have the same cause and therefore may often occur together. The internalisation of oppressive norms is one such case. The erosion of self-trust and self-respect resulting from the internalisation of oppressive norms can cause both damage to epistemic agency as well as causing epistemic character damage by promoting the development of vices such as intellectual servility.

4.7 Conclusion

In this chapter, I have argued that the development of intellectual servility as a PECT constrains epistemic agency primarily by creating psychological barriers to the exercise of epistemic powers. It can also, in some but not all cases, further entrench conditions wherein individuals lack access to shared epistemic resources or fail to gain appropriate recognition and uptake from other members of their epistemic community. Both things are required in order to reliably succeed in performing a range of epistemic actions and therefore their absence constrains epistemic agency.

This cost occurs in addition to epistemic character damage. It does not result from the development of all vicious PECTs. For example, the development of closed-mindedness does not tend to constrain epistemic agency. Here I have focused on the role of negative attitudes towards the self – such as a lack of self-trust and lack of self-respect – in creating psychological barriers to the exercise of epistemic powers. These attitudes are specific to PECTs from the 'keep your head down' cluster. But there may be other forms of psychological barrier which arise from other PECTs which have not been explored here.

I have now identified two ways in which the development of vicious PECTs can carry costs for students: through damage to character and the constraint of epistemic agency. These are the kinds of costs that must be taken into account when considering whether it is reasonable to expect students to protect themselves from harm. If protecting themselves from harm requires developing vicious PECTs then the incurrence of these costs may disadvantage students or cause them significant harm. Such an outcome is clearly not a positive one for students' epistemic character development. Thus, when individuals are more likely to develop vicious PECTs rather than virtuous ones (for the reasons discussed in chapter 2)

then the exposure to potential harm in the classroom will be detrimental to students' development rather than aid it.

In the final chapter, I show how the costs of self-protection are related to safety. I will argue that, for a classroom to be safe, students must not be required to protect themselves in ways that cannot be reasonably expected of them. Costs such as damage to epistemic character and the constraint of epistemic agency make a means of self-protection unreasonable. Thus, using the account developed in chapter 5, I will show that a classroom in which students must develop vicious PECTs like closed-mindedness and intellectual servility to protect themselves from harm is not safe for those students.

Chapter 5
Safety, Risk and Protection: Access Issues

In this chapter, I will tie together some of the threads that have developed over the course of the thesis and return to the central topic: safety in the classroom. I have focused in chapters 2, 3 and 4 on epistemic character traits which can protect their bearers from various harms arising in the classroom but which may themselves cause other kinds of harm. In this chapter, I will demonstrate how such costs associated with means of self-protection feature are related to safety in the classroom.

At the end of chapter 1, I briefly presented the claim that safety is valuable in the classroom because it is a baseline condition for access. In this chapter, I will further develop and defend this claim by arguing that high levels of objective risk of harm can make a classroom inaccessible for students. In the absence of detailed knowledge about the extent to which students are able and willing to expose themselves to high levels of risk, educators can ensure that classrooms are accessible to students by ensuring that those classrooms are safe according to the definition of safety which I will defend below. As a result, certain responsibilities for making the classroom safe fall on educators and educational institutions arising from their commitments to making the classroom accessible.

If educators and educational institutions are committed to the principle that no students should be discriminated against or unfairly disadvantaged, then they will be committed to the following assumptions:

(1) Educators and educational institutions should ensure that the institution, classroom, materials, and teaching are accessible, wherever possible, to the students to whom they purport to provide education.
(2) Educators and educational institutions should ensure that no student is unfairly disadvantaged by the way in which the education is provided or the services, resources, facilities, or benefits to which they are afforded access.

It is worth making a brief comment on the limitations of assumption (1). The responsibility to educators and educational institutions does not commit them to making the classroom and institution accessible to students for whom this would require measures or changes which place too great an imposition. Whilst assumption (1) refers to, I take it, a moral rather than legal obligation, it is useful to compare it to the language of 'reasonable adjustments' used in the Equality Act 2010 in relation to discrimination against people with disabilities. (It should be noted that the moral obligations may extend beyond the legal obligations set out in the act.) Those covered by the Equality Act 2010 (including educational institutions) are required to 'take reasonable steps to avoid putting disabled people at a substantial disadvantage by changing a provision, criterion, or practice, by removing or modifying physical barriers, and by providing auxiliary aids and services' (Wadham *et al.*, 2017, p. 44). Considerations for determining whether an adjustment is reasonable include: whether the adjustments are effective in overcoming substantial disadvantage; whether implementing them is practicable; the financial or other costs of implementing them and (in limited circumstances) the interests of others, including other students (Equality and Human Rights Commission, 2014, pp. 106–107). It is also important to note the requirement to make adjustments does not include adjustments made to 'competence standards' which are 'academic, medical, or other standards applied for the purpose of determining whether or not a person has a particular level or competence or ability' such as 'a requirement that a person has a particular level of relevant knowledge of a subject'. (*ibid.*, pp.98-99)

In §1, I will begin by developing my account of safety. I demonstrate the relationship between safety and protection, arguing that a classroom can be safe when students must protect themselves from harm but only when such self-protection can reasonably be expected of them.

In §2, I argue that higher levels of risk of harm may make a classroom inaccessible to a student. Whilst in particular instances a classroom may be accessible to a student despite a high risk of harm because the student is willing and able to take on that risk, without detailed knowledge of the student's circumstances educators cannot *presume* that a classroom with a high level of risk is accessible. If a classroom is safe for a student then the classroom can be presumed accessible for them (in the absence of other kinds of barriers to access). Finally, in §3, I apply my account to two case studies: content warnings and harmful speech in the classroom.

5.1 Safety and Risk

Safety is often posed in opposition to risk: roughly, a safe classroom is one that is not risky. It is from this basic intuitive understanding of safety that I begin to build my account. However, it need not be the case that a safe classroom is one which is entirely free from risk. The risk of harm is omnipresent but when the level of risk is low enough it can be ignored and remains in the background. This 'background level' of risk is one which individuals do not factor into decision making. For want of a better term, I will refer to this level of risk as a 'tolerable level of risk' although this involves a non-standard use of 'tolerable'. When risk exceeds this level, it begins to factor into individual's decision making but they may well be willing to take on that level of risk.[86]

Let us begin with the following account of safety:

> *Safety*$_1$: A classroom C is safe for some student S if being in and participating in C does not put S at higher than tolerable level of risk of harm.

There are several points which must be explored in more depth to make this account more precise: what determines the tolerable level of risk and how does the availability of protective measures (including protective equipment, behaviours that reduce the risk of harm, or 'psychological' protective mechanisms such as the PECTs identified in chapter 2) impact the safety of the classroom?

5.1.1 Tolerable Risk

Determining what is required to make the classroom safe will depend in part on the level of risk which can be tolerated. There are many ways in which low levels of risk are present in our lives but, because they are so low, we do not pay much attention to them or do not consider them to be a barrier or reason not to do things. For example, driving to the university involves some degree of risk of being involved in a car accident; walking down the stairs involves some risk of tripping and being injured. However, so long as these risks are low many people do not factor them into their decision making (though I will discuss below some of the reasons why for some, even

[86] I include in my discussion examples both of risks of physical harm and risks of non-physical harm (including psychological and emotional harm). Whilst these harms are of course very different in nature, I assume that my account is general enough to cover risks of either kind of harm.

these low levels of risk may not be tolerable). Of course, if levels of risk increase, there reaches a point where it is no longer considered an insignificant background condition. In this section, I will show how this point – when risk becomes no longer tolerable – is determined.

There is no single, universal line at which some a level of risk becomes intolerable. Firstly, it will depend on the severity of the harm. We can tolerate much higher levels of risk of small harms. For example, I can tolerate a much higher risk of getting a paper cut than I can of breaking my arm. In addition to this, the tolerable level of risk is also relational to the individual. That is, two different individuals may be able to tolerate different levels of risk of the same harm.

There are a number of reasons for this. This include, but need not be limited to, the vulnerability of the individual to further negative effects of the harm, the ability of the individual to recover from or mitigate the harm (including the support systems they have in place to aid this), the responsibilities and obligations which the individual has to others and will be affected if they suffer the harm.

Often the ability to tolerate a harm will be connected to one's status and relative power. For example, having precarious employment may increase the knock-on effects of or decrease someone's ability to mitigate the harms of injury which makes them unable to work for a time, compared to being employed on a permanent contract. Similarly, those in positions of relative power are more likely to have greater resources and more influential support networks on which they can draw to mitigate or remedy harms – leading to the ability to tolerate higher levels of risk.

Consider the level of risk of contracting Covid-19 which different individuals can tolerate. Someone who has long-term health conditions which would be exacerbated by Covid-19 may only be able to tolerate a lower level of risk than someone who does not. Other factors which may reduce the risk of contracting Covid-19 which someone can tolerate include:

- Living with a family member who is clinically vulnerable
- Being a single parent who would greatly struggle to care for their children if they had moderately severe Covid-19 symptoms
- Being financially insecure and working in a job with insufficient sick pay and therefore struggling financially if required to take time off work due to illness

All of these factors feed into the overall impact that taking on a risk of catching Covid-19 would have on someone and their ability to mitigate or

recover from that impact. Different kinds of risks will also be related to different kinds of factors. For example, the level of risk of being ridiculed and humiliated by their peers which a student can tolerate may be reduced if they encounter this or risks of this in other aspects of their lives. If they require the classroom to be a space of respite from ridicule and humiliation, then they may be less able to tolerate risk of it in the classroom.

There are two clarifications to make. Firstly, my claim is not that the tolerable level of risk is purely subjective. The claim that the tolerable level of risk of some harm is relational to the individual should be distinguished from the claim that the tolerable level of risk of some harm for some individual is the level of risk of that harm which they are willing to or comfortable enduring. This is because people may have vastly different attitudes to risk. Some people may be extremely risk averse and be uncomfortable with even a very low level of risk.

Students' subjective attitudes towards risk does not seem to be the relevant standard to use when evaluating the safety of the classroom.[87] Requiring that the level of risk in the classroom be reduced to a level comfortable for even the most risk averse students would be too demanding. For example, imagine a student who is extremely risk averse and is uncomfortable with even a very low risk of a terrorist attack at the university. It seems excessive to say that the classroom is unsafe for them because it is impossible to completely eliminate this risk and attempting to do so would place a very large burden on the educator or institution. It also seems to misdiagnose the problem when students aren't able to participate in the classroom given their high-risk averseness. Rather than it being an issue of safety – to be addressed by making the classroom safer by reducing the level of risk – the issue arises from the individual's attitude toward risk and is best addressed by re-evaluating this attitude. That being said, the reason for the student's risk aversion and the effect which the perception of risk has on them can be relevant to the safety of the classroom for reasons I will discuss below.

Determining the tolerable level of risk of some harm for an individual will be very difficult from a third-person perspective. Often educators are not in a position to know about or judge the factors in a student's life which alter

[87] Note that this is still distinct from the subjective account of risk discussed in chapter 1, according to which safety was defined in terms of students' subjective (perceived) level of risk. The claim I am opposing here is that safety is defined in terms of the objective level of risk, but the level of objective risk at which the classroom becomes unsafe for a student is determined by their subjective attitude towards risk.

their tolerable level of risk. If they know students well, they may know about student's part-time work, their health conditions, their family situation and so on. But, unless they *really* know the student, it is unlikely that they will know the full details of factors such as the kind of support network they have available, the current state of their mental health and how it could be affected by various harms. Therefore, it is often the case that educators must rely on individual's reporting of their own tolerable level of risk. That is, although there is a fact of the matter about what level of risk a student can tolerate – independent from the level of risk they *feel comfortable* with – educators are not generally in a good epistemic position to identify this level. Students may also lack some of this knowledge or may confuse the tolerable level risk with the level of risk with which they are comfortable, but they are better placed than educators and therefore it is often required that educators take students at their word.

The second clarification to be made is the difference between (a) the objective level of risk surpassing the point of being tolerable and (b) the *perception* that the level of risk has surpassed the point of being tolerable. For example, imagine a similar case to the one mentioned above – a student who believes there to be a very high risk of a terrorist attack at the university, but this belief is not justified.[88] This relates to the distinction I made in chapter 1 between objective and subjective elements of safety – the extent to which students are actually at risk of harm in contrast with the extent to which students perceive themselves to be at risk or fear harm. As I argued in chapter 1, safety should be defined in terms of objective risk of harm.

In two types of case – those in which students are uncomfortable with even very low levels of risk and cases in which students believe the level of risk is higher than tolerable – there is a complicating factor worth noting: the psychological effects of feeling at risk (Wolff and De-Shalit, 2007, p. 68). Consider again the student who has a heightened fear of a terrorist attack on the university. Suppose that this fear is a result of the student been a survivor of the Manchester Arena bombing in 2017 and any slight increases in the risk (such as an increase in the national threat level) will cause the student to become severely distressed at the thought of an attack occurring. In this way, the heightened perception of risk may itself cause a student

[88] Students' over-estimation of the level of risk could be related to an over-sensitivity to indicators of risk. Furedi (2003) and Ecclestone (2004, 2008) both argue that cultural shifts have lead to increased sensitivity to risk. They both suggest that this has resulted in increase over time of students' aversion to risk.

harm. The impact which this feeling of fear has on the student (given the slightly elevated level of risk) increases the overall impact which the risk has on them. Therefore, it has the potential to reduce their tolerable level of risk – not because of the impact which the occurrence of the feared harm would have on them, but rather, because of the impact which their awareness of the level of risk has on them. Similarly, if a student falsely and unjustifiably believes that they are at a high risk of harm, this may still *cause* harm itself through either the distress which it causes or the impacts of the protective mechanisms which it leads the student to put in place.

5.1.2 Safety and Protection

The level of risk of harm will vary depending on the protective measures put in place. Suppose that there is a risk of harm arising from the presence of sharp tools in a lab. This risk could be significantly reduced through certain protective measures being put in place. These might include safety guards on the tools or protective equipment worn by the students. They may also include practices, skills or behaviours of the students using the tools which reduce the risk of harm, such as using proper cutting techniques and following safety guidance.

Protective measures may be implemented by a number of different actors. For example, the educational institution will be responsible for implementing certain protective measures in the designs of their buildings (for example, safety latches on windows and handrails on stairs) or the provision of materials and resources (for example, ensuring all tools have the necessary safety guards). Individual educators may be responsible for implementing and enforcing certain practices within their classroom or for supervising certain activities. Students themselves may be responsible for implementing other protective measures relating to their own behaviours and attitudes to protect either themselves or others (such as the development of PECTs).

When concerns around safety arise, they often highlight a lack of sufficient protective measures. Often, the claim that a classroom is unsafe for some student or students is a call for the educator or educational institution to implement greater protective measures. In some instances, the institution, members of the institution, or even outsiders, respond by objecting that the institution and educators are not responsible for putting in place the

requested protective measures.[89] Instead, the responsibility falls on students to protect themselves.

We might think that, if protective measures are available for a student to implement herself and these measures will reduce the level of risk of harm to a tolerable level, then this makes the classroom safe. For example, the lab is safe for a student who is easily able to use the sharp tools correctly in a way that considerably reduces the risk of injury. However, I will argue that this is not always the case by describing cases in which the classroom is not made safe by the availability of a self-protective measure.

Firstly, consider two examples where a classroom *is* made safe by the availability of a protective measure:

> COLD*:* As part of Josef's geology course he must do some fieldwork during the winter. The temperature outside is low enough that, were Josef to work outside all day wearing just a thin jumper he would be at risk of hypothermia, but when wearing a winter coat (which Josef owns) it is a comfortable temperature.

In Josef's 'classroom' – the fieldwork site – the low temperature does not make the environment unsafe despite the risk of hypothermia. This is because, Josef can easily adopt a protective measure[90] – wearing a suitable winter coat which would sufficiently reduce the risk of hypothermia below a tolerable level. In this respect (assuming there are no sources of potential harm in the environment) the environment is safe because there is a measure that Josef can easily adopt to reduce the risk of harm from hypothermia.

> ALTRUISM: In an ethics seminar, the class are discussing effective altruism. Jade expresses her support for the view and describes how she uses its principles in her own life. The majority of the rest of the group are sceptical about the benefits of effective altruism. Some of them express their disagreement with Jade and challenge, sometimes

---

[89] For example, in 2016 the dean of students at University of Chicago wrote to incoming students stating that the university did not support the creation of intellectual safe spaces, or condone the disinvitation of controversial speakers and use of trigger warnings' (Ben-Porath, 2017, p. 14).

[90] This is assuming that Josef has access to a winter coat. If he could not afford to buy one and had no other way of getting hold of one, then it would not make the classroom safe for him.

> quite strongly, some of the background assumptions she holds about our duties to others and how best to help others. Jade is surprised by these criticisms as she had not encountered them before and it makes her seriously reconsider her approach to altruism – something which is very important to her. This has the potential to be quite distressing for Jade because it brings into doubt a value which central to her daily life. She might feel quite powerless or like a lot of the things she has done in the past have been pointless. It could have a marked effect on her mental state and how she thinks about her life. She might also feel that her classmates view her as a 'bad person' for holding so strongly a view with which they disagree.

Coming into contact with ideas which challenge deeply held beliefs or assumptions is supposed to be a valuable part of higher education. Generally, students like Jade can be expected to work through these challenges, either through continuing to reflect or through discussion with others. Of course, these things can be difficult but we generally assume that there are ways for students like Jade to sit with and process their discomfort and inner turmoil so that the end result is positive – they have the opportunity to deeply reflect on their beliefs and assumptions and learn something valuable (Boler, 1999). The discomfort and turmoil themselves are not (or are very unlikely to be) significant harms (perhaps, they may not even be harms at all). The significant harms are the more serious effect that they could have or contribute to were they not to be process or worked through in a positive way. This could be a deep sense of hopelessness or depression which could come from losing one's sense of purpose. It is expected that Jade is able to either process things herself, or seek support from others so as to avoid these extreme outcomes. There are of course some instances in which an educator may have concern that a student may not be able to do this, for example because they lack adequate support systems or because of existing mental health difficulties.[91] In such a case an educator

---

[91] For example, Berman (2001, pp. 240–244) describes teaching two students who, while studying novels which depicted suicide and suicidal ideation (including Plath's The Bell Jar) found themselves 'at risk'. The experiences of the characters in the novels struck a chord with their own experiences of depression and they found the exploration of these topics had a significant impact on the mental health. Whilst many students will study literature such as this, many will be able to process difficult topics without an impact on their mental health. Some students, like the two Berman describes, are at a higher risk of

may (if able and it is appropriate) aid the student in protecting themselves by offering or directing them to support.

In both COLD and ALTRUISM, the classroom is safe for the student because they can reasonably be expected to protect themselves from the potential harms present. However, compare this to a different case:

> NUTS*:* Cerys has a severe nut allergy – even the smallest traces of nuts on a surface or in the air can cause a severe reaction. She is taking a catering class taught in a kitchen in which nuts are often used. As a form of self-protection, Cerys could wear a full hazmat suit to prevent herself from coming into contact with or breathing in traces of nuts. This form of protection significantly reduces the risk of her coming to harm as a result of the traces of nuts in the kitchen to below a tolerable level.

Suppose that it is possible for Cerys to wear a hazmat suit: she has been provided with one to wear in the classroom. However, we would not say that the classroom is safe for Cerys because she has the option to wear a hazmat suit and this would reduce the risk of harm from an allergic reaction to a tolerable level.

To make the point clear, it is true to say that 'when she wears a hazmat suit, the classroom is safe for Cerys' since it is true that wearing the suit means that there is a very low risk of harm from an allergic reaction. However, my claim instead is that we cannot say that the classroom is safe for Cerys in general *because* the option of wearing a hazmat suit is available to her (just like the fieldwork site is safe for Josef because the option of wearing a winter coat is available to him).

The difference between COLD and NUTS is the nature of the available means of protection – specifically, how *reasonable* it is to expect the student to adopt this measure. It is reasonable to expect Josef to wear a coat when it is cold since, this is a normal everyday practice. It does not get in the way of Josef going about his fieldwork tasks. On the other hand, wearing a hazmat suit is not an everyday practice, it is cumbersome, uncomfortable, and will likely make it more difficult for Cerys to participate in the practical

---

harm due to their previous mental health difficulties. Berman describes how he discussed with the students what support their had available, and when it because clear that one did not have sources of support which could protect her from harm to her mental health, he explored ways in which the course assignments could be adapted to allow her to process the difficult topics in a way that hopefully reduced the risk of harm to her mental health.

learning activities. It is not something we would normally expect Cerys to do because of the burden it places on her.

Therefore, having available options for self-protection which are effective in reducing the risk of some harm to a tolerable level is not sufficient to make the classroom safe for a student. It is only sufficient for making the classroom safe if it is reasonable to expect that the student adopts those protective measures.[92]

Below, I will describe three reasons why it may be unreasonable to expect a student to adopt a particular protective measure:

1. The protective measure is not genuinely available to them in the classroom *qua learning environment*
2. Adopting the protective measure will unfairly disadvantage the student
3. Adopting the protective measure will cause or make the student vulnerable to further harm

5.1.3 Protective measures which are not genuinely available to students

In order that we can reasonably expect a student to adopt a protective measure, that protective measure must be genuinely available to the student in two senses. Firstly, it must be within the capabilities of the student and they must have access to the resource required. For example, is it not reasonable to expect a very young student to protect themselves from harm by using sharp tools correctly if their fine motor skills are not sufficiently developed to do so.

Secondly, they must be able to adopt the protective measure within the constraints of the activities of the classroom. To illustrate this second requirement, consider two examples:

> ROCK-CLIMBING: An indoor rock-climbing venue has broken safety harnesses meaning that if a climber were to fall from any of the climbing walls, they would likely be severely injured. The broken equipment poses no risk of

[92] The requirement I propose is not just that students can be expected *to be able* to adopt but also can reasonably be expected to do so. For example, Cerys may be very clear *able* to wear a hazmat suit because one that fits her has been provided, but she still cannot actually be expected to *in fact* wear it, since doing so would be incredibly uncomfortable and would greatly hinder her ability to engage in the practical learning tasks in the classroom.

harm to people in the venue who are not climbing on the walls.

We might say that the venue is safe in the sense that, someone can enter it and by implementing the self-protective measure of not climbing on the wall, they are at a very low risk of harm. Suppose that I go to the rock-climbing venue as a place to climb. Considering the rock-climbing venue *qua place in which I rock-climb* means we take for granted that using the space involves participating in the activity of rock-climbing. That is, we presume that I will use the climbing walls. Once we take this for granted, the self-protective measure of *not* climbing on the walls is no longer available since it is inconsistent with using the venue qua place for rock-climbing.

Here is an analogous case in the classroom:

> BALLET*:* Amir is studying ballet. He is bullied and mocked by students in his class and this has severe impacts on his mental health. The bullying typically occurs when Amir dances in front of his classmates. If he did not dance during the class, then he would avoid the majority of the bullying and this would likely improve his mental health. However, this would mean that he would miss out on many of the central aspects of the class.

Withdrawing from the dancing is a self-protective measure which is open to Amir. However, doing so would impact on how he is able to engage in the learning activities. We might say that simply entering the ballet classroom is safe for Amir since he can protect himself from bullying by not dancing in the class. However, when we consider the classroom *qua place for Amir to study and practice ballet* then we presume that he will engage in the requisite activities – that is we presume that he will participate in the dancing. Therefore, the classroom *qua place for Amir to study and practice ballet* is not safe for him given that the self-protective measure of not participating is not available to him (since it is inconsistent with the classroom qua place for him to study and practice ballet).

Therefore, self-protective measures which preclude engaging in the requisite activities of the classroom qua place of learning are not genuinely available to students. And since students cannot be expected to adopt protective measures which are not available to them, they cannot be expected to adopt protective measures which preclude engaging in the requisite activities of the classroom qua place of learning.

As per the assumptions stated at the beginning of the chapter, educators have a responsibility to make the classroom and learning activities

accessible to the students to whom they purport to provide an education wherever possible. If a student is required to implement a protective measure which prevents them from participating in the central learning activities of the classroom when reasonable steps could be taken to avoid this (such as the implementation of other protective measures by the educators or institution) then this infringes on their opportunity to learn in the classroom. As a result, this violates the educators and institutions responsibility to make the classroom accessible to the student. For these reasons, protective measures which are incompatible with students' engagement in the central learning activities of the classroom cannot be reasonably expected of students.

5.1.4 Protective measures which disadvantage students

There are several ways in which the adoption of protective measures can be consistent with the activities of the classroom whilst still disadvantaging students. Firstly, some protective measures make it more difficult to engage in the requisite activities of the classroom qua learning environment even if they do not make it impossible. Consider NUTS: wearing the hazmat suit does not entirely prevent Cerys from participating in learning activities but it will make the practical tasks more difficult (suppose that the class is on cake decoration and is largely practical). She will find it difficult to move around, won't be able to see as clearly, the gloves will make it difficult to do delicate tasks and she might not be able to hear instructions clearly. Overall, wearing the hazmat suit will detract from but not completely rule out her ability to engage in the practical learning tasks in the classroom. Therefore, in order to protect herself she would need to make some sacrifices regarding her engagement in the learning activities in the classroom. If this is a significant sacrifice it will put her at a disadvantage compared to her peers who do not need to wear hazmat suits since they will be able to better engage in the learning activities.

Another way in which self-protective measures can detract from students' ability to engage in learning activities is by drawing time, energy and attention away from those activities. For example, imagine that instead of wearing the hazmat suit, Cerys had to thoroughly clean every surface and piece of equipment which she used. This could take up a lot of time and attention which she would otherwise put into engaging in the learning activities.

Cerys would be unfairly disadvantaged by the need to wear a hazmat suit or to thoroughly clean any surfaces she may come in contact with. Therefore, there is a responsibility to the educators or institution to avoid this

disadvantage wherever possible. Suppose that it is not necessary for any of the classes taught in that classroom to use products containing nuts. The classroom could be designated as 'nut free' without any significant change to the teaching. If this option is easily available to the educators, then there is an easy way in which they can avoid Cerys being at a disadvantage. This means that (i) there is a responsibility placed on the educator or institution to make this adjustment in order to avoid Cerys being at a disadvantage and (ii) it is not reasonable to expect Cerys to adopt this disadvantaging protective measure since there is an easy alternative which would avoid the need for it.

Similarly, consider an example of psychological rather than physical harm, in which one of the PECTs discussed in chapter 2 – intellectual perseverance – may be developed:

> INTELLECTUAL PERSEVERANCE: Ayesha's peers and teachers often assume that she is not very intelligent or suited to academic studies. They often don't listen to her contributions in class or disregard them because they assume they will be non-sensical or irrelevant. Over a long period of time, she might internalise these attitudes and develop intellectual servility and as a result her epistemic agency may be constrained. In order to avoid this, and to avoid the harms and disadvantages which come from being ignored and dismissed, she might persevere. She might keep attempting to participate and be taken seriously, perhaps by developing new strategies, presenting herself in certain ways, or making allies in the classroom.

Perhaps being motivated to overcome the barriers she faces in being taken seriously could protect her from the harms which may arise from them in the ways described in chapter 2. It is an epistemic character trait which changes her attitude towards the difficulties which she faces and through this attitude she can avoid the downward spiral which can result from the internalisation and acceptance of such attitudes towards her.

However, developing this attitude and persevering in attempting to be taken seriously can demand a lot of time and energy. She may have to repeat herself over and over again, say things in different ways or pay acute attention to the way she presents herself. All of this draws her attention away from actually engaging in the topic of the discussion. She might find it exhausting having to work so hard just to be heard and this might make

it harder to learn or will leave her with less energy to explore the topic.[93] Again, this puts her at a disadvantage compared to her peers who need not exercise intellectual perseverance: they have greater time and energy to spend on engaging with the learning activities.

Requiring some, but not all, students to adopt protective measures which disadvantage them in participating in the learning activities of the classroom violates the responsibility not to unfairly disadvantage students in the classroom. Therefore, given teachers responsibility not to unfairly disadvantage students, students cannot be reasonably expected to adopt such protective measures.

The reasonableness of an expectation for a student to implement a self-protective measure is dependent to some degree on the available alternatives. For example, in the case above, the ease with which adjustments could be made by the educational institution to avoid Cerys being required to wear the hazmat suit contributes to the unreasonableness of the expectation that she wear it to protect herself from harm. When no other alternative is available, the case is less clear cut. Imagine instead that it is necessary for products containing nuts to be used in the classroom because of syllabus requirements and that there are no other kitchen facilities available.

There are two things we might say in such a case. Firstly, the expectation becomes more reasonable since it is the best available means by which Cerys is protected from harm. However, such a claim has a potentially counter-intuitive result. Classrooms in which there are no easy means by which educators and institutions can protect students from harm are more easily classified as safe for students in the sense that more demanding expectations can reasonably be made of students to protect themselves from harm. But we would not want to go so far as to say that the absence of any viable means for the educator or institution to reduce the level of risk of harm to a tolerable level entails that we can reasonably expect students to protect themselves in extreme ways. Therefore, the dependence of

---

[93] Another example could be the stress and emotional labour which members of marginalised groups can be required to do in order to justify their place in the classroom. For example, responding to micro-aggressions, over-coming ignorance and stereotypes can difficult and taxing work which draws energy, time and attention away from other projects (see Berenstain, 2016). Literature on stereotype threat notes that the desire to and attempt to avoid confirmation of a negative stereotype can create a substantial cognitive load and it is in this way that awareness of a negative stereotype can impact performance (Steele, Spencer and Aronson, 2002).

reasonable expectations on the viability of other options must be trumped by other factors.

Alternatively, we can say of Cerys's case, that if wearing the hazmat suit is the only available means by which she can be protected from harm, then the classroom cannot be made safe for her. This would be consistent with Cerys still choosing to enter and participate in the class by wearing the hazmat suit, however, it will raise (perhaps unresolvable) concerns about the fairness of the situation.

5.1.5 Harmful protective measures

A third reason for which students cannot be reasonably expected to adopt protective measures is if those protective measure cause or make them vulnerable to further harms. Two such examples are the harms discussed in chapters 3 and 4. The development of vicious PECTs gives rise to harm through the damage of epistemic character. Furthermore, the development of intellectual servility as a PECT can also give rise to the additional harm of the constraint of epistemic agency.

Other protective behaviours may cause social harms, psychological or emotional harms. For example, a student may protect themselves from bullying by isolating themselves from others, but this may result in social harms such as loneliness and an inability to form friendships. Alternatively, consider a transgender woman who protects herself from discrimination and harassment from her peers by concealing her gender identity and presenting as male in the classroom. Whilst this may protect her from discrimination and harassment, it can have significant negative psychological and emotional effects as it requires her to hide and deny her identity.

It is not reasonable to expect students to subject themselves to, or make themselves vulnerable to, significant harm in the classroom. Furthermore, if the level of risk of harm resulting from such protective measures is higher than tolerable, then the classroom fails to be safe.

A classroom is not safe for a student when the only way to lower the risk of harm to a tolerable level is by implementing a protective measure that they cannot be reasonably expected to adopt. Therefore, I propose the following revision to the definition of safety offered above:

> *Safety*$_2$: A classroom C is safe for some student S if being in and participating in C does not put S at higher than tolerable level of risk of harm without S being required to protect

themself in ways that cannot be reasonably expected of them.

Now that I have offered an account of safety in the classroom, I will demonstrate the connection between safety and accessibility by arguing that a higher than tolerable level of risk of harm can make the classroom inaccessible to a student.

5.2 Access

Wherever possible, classrooms should be accessible to as many students as possible. There are many different ways in which a classroom may be inaccessible to a student or group of students – including physical and financial barriers or a lack of appropriate resources or tools. I will argue below that classrooms can also be inaccessible to a student if being in the classroom and participating in the learning activities places the student at too high a risk of harm. For this reason, there is a connection between classrooms being safe and classrooms being accessible.

I will begin by defining an accessible classroom. In doing this I will distinguish between two senses of 'accessible' which I will call *actually-accessible* and *presumed-accessible* and argue that it is the latter of these two senses with which educators and institutions should generally be concerned when ensuring that classrooms are accessible to students. Then I will argue that reducing the level of risk to a tolerable level is necessary to make the classroom presumed safe in the absence of detailed knowledge regarding the abilities of individual students to take on high levels of risk.

5.2.1 Actual- and *Presumed-Accessibility*

The most straightforward sense in which a classroom is accessible for some student S is that S is able to enter the classroom. Access is most generally discussed in relation to physical (and sometimes other forms) of disability.

Consider a paradigm example of inaccessibility:

> STEPS*:* Cara has a mobility impairment and uses a motorised wheelchair. One of her classes is scheduled in a room on the second floor without any lift access. The classroom is inaccessible to Cara because she cannot physically get to the classroom because there is no step-free access.

But access is much more than just being able to get into the room. Consider another example:

DEAF*:* Ben is deaf. In one of his classes the lecturer just speaks without slides or handouts. There is no provision for a sign language interpreter, the room does not have a hearing loop, and Ben is unable to lip-read because the lecturer wanders around the room and often has his back to him.

Ben is able to enter the classroom but is unable to participate in the learning activities and acquire any of the information being conveyed because it is not provided in a format which he can use. The lecture theatre, simply as a space, may be physically accessible to Ben because he is able to enter it, but the classroom *qua learning environment* (that is, a lecture where content is presented) is not accessible to Ben.

This is a fairly basic picture of what access involves – a classroom *qua learning environment* is accessible to students when they are able to enter and be in the space, participate in the requisite activities of the classroom and make use of the resources provided.

But there is more detail to add to this picture. Consider again STEPS – suppose that Cara's mobility impairment is such that she can walk up the stairs but doing so will most likely be incredibly painful and exhausting and there is a very high chance that it will cause her a significant injury. There is a sense in which Cara could get into the classroom: she could walk up the stairs. But clearly, the fact that this option is in principle available to her does not suffice to make the classroom accessible to her in any meaningful sense.

When entering the classroom and participating in the learning activities presents a significant level of risk of harm to a student then it is reasonable that the student will choose not to enter the classroom or participate in the learning activities. For example, imagine that Cara is presented with the options of climbing two flights of stairs or missing her class. It would not be surprising if she chose to skip the class rather than risk injury, pain, and exhaustion.

If Cara were to willingly choose to attempt to walk up the stairs because her attitude to risk is such that she is happy, and perhaps even enjoys, taking such risks, then the classroom might be said to be accessible to her. Let us call this kind of accessibility '*actual-accessibility*'. A classroom is *actually-accessible* for a student when they are able to enter and participate in the classroom, including when they choose to do so by knowingly taking on a high level of risk of harm.

There are a wide range of factors which will feed into Cara's decision about whether she is willing to walk up the stairs to attend the lecture. These

factors may include: her general level of health; how much energy she has; or whether she will need to reserve her energy for other tasks.

The decision may also depend on her attitude to risk. Even when risk is above tolerable someone might still choose to take on that risk. For example, many people are willing to take on significant risk in extreme (or even non-extreme) sports. However, different people have different attitudes to risk – some are risk averse, others enjoy risk to some degree or view the harm as a worthwhile cost of the benefits gained.

As a result, there is a grey area in which it is difficult to predict whether an individual would willingly take on a higher than tolerable level of risk. Unless the lecturer, or person in charge of timetabling, knows a lot about Cara's attitude to risk, then it is unlikely that they will be able to predict exactly which kinds and levels of risk of harm she will be willing to take on. A further difficulty arises from the fact that Cara's attitude to the level of risk arising from attempting to walk up the stairs may change over time. For example, she may be willing to take on the risk at some times and not at others depending on factors such as her overall health at the time, whether she has done any other strenuous physical activity recently, or the commitments she has over the next few days. Therefore, the classroom may be *actually-accessible* for Clara at one time but not at another.

The result of this is that educators and those making relevant decisions within the institution are unlikely to know with any certainty what measures need or need not be taken in order to make the classroom *actually-accessible* for students.[94] *Actual-accessibility* is therefore not a useful standard for educators and relevant decision makers within educational institutions to use to establish whether their classrooms are accessible to their students.

I therefore propose that a difference sense of 'accessible' is more useful in this context – what I will call *presumed-accessibility*. A classroom is *presumed-accessible* for some student when educators and relevant decision makers within the educational institution can *presume* that the student will not reasonably and justifiably choose not to enter or participate in the

[94] In many instances, students with disabilities will have support plans (or equivalent) which outline their specific support needs. Depending on how detailed it is, access to such information will change the requirements for making the classroom presumed-accessible. But it may not be specific and detailed enough to allow educators and relevant decision makers within the institution to know exactly what is required, at any particular time, for the classroom to be actually accessible.

classroom. By making the classroom *presumed-accessible*, educators and institutions are able to maximise the probability that the classroom will be *actually-accessible*. When establishing what is required to make a classroom accessible for a particular student it is important for those lacking in detailed knowledge of a students' ability and willingness to expose themselves to risk to err on the side of caution. Cara's lecturer (or those timetabling the lecture) cannot presume that Cara will be willing to risk walking up the stairs, therefore if this is the only way for her to enter the classroom, then it cannot be *presumed-accessible* for her.

What is required for the classroom to be presumed-accessible depends on the knowledge which educators and relevant decision makers within the institution possess regarding Cara's ability to take on risk of harm. The greater knowledge they possess, the more fined-grained the informed and justified assumptions they are able to make about the levels and kinds of risks she is able and willing to take on.[95]

5.2.2 Presumed-accessibility and safety

So far, I have argued that, in the absence of detailed information about a student's ability to take on high levels of risk, educators and institutions should aim for classrooms which are *presumed-accessible.* In order to ensure that a classroom is *presumed-accessible* for some student, we need to remove any reasons for which they can reasonably and justifiably choose not to enter the classroom. The more we know about the student, the more fine-grained we can make our judgements. However, it is worth noting the danger in over-assuming because of some familiarity with the student. As noted, attitudes towards risk may change over time; and even though a student chose to take on a risk previously, they might reasonably and justifiably choose not to take on that same risk at some other time.[96]

---

[95] I have discussed here presumed-accessibility for some particular students. However, many decisions made within the educational institution will be related to accessibility for hypothetical students. For example, when a new teaching building is being designed, it is not known which students will be using the building the classrooms in the future. Rather, attempts to make the classrooms accessible will be made with hypothetical students in mind. The judgements made regarding the accessibility of, say, a newly designed classroom can be understood as judgements regrading presumed-accessibility. That is they are judgements along the lines: 'given the features of this classroom, we can presume that x,y,z students will not be prevented from or reasonably and justifiably choose not to enter and participate in the classroom.'

[96] There may be further dangers in assuming too knowledge of a students' abilities and ability to take on risk. Over-cautious assumption regarding the level of risk of a student

When a classroom is safe for some student S, the level of risk of harm for S is tolerable without S having to protect herself in ways that cannot reasonably be expected of her. When the level of risk of harm is tolerable then it is below the level at which it begins to factor into our decision-making. Therefore, when the level of risk of harm is tolerable, S will not reasonably choose not to enter or participate in the classroom on the basis of risk. Therefore, when the classroom is safe for S it meets the above condition for *presumed-accessibility*[97].

Given sufficient knowledge of a students' attitude towards and ability to take on risk, then it is possible for a classroom to be *presumed-accessible* for a student whilst the level of risk is higher than tolerable. In other words, it is possible that an educator knows a student well enough (notwithstanding my earlier caution about presuming to know more than one does) to be able to presume that a student will be willing to take on some level of risk of harm *r* which is higher than tolerable. Therefore, so long as the level of risk of harm does not exceed *r* then the classroom meets this requirement for *presumed-accessibility* for the student. But since *r* is a higher than tolerable level of risk, the classroom is not safe for the student. Therefore, it is not necessary for a classroom to be safe for some student *S* in order for it to be *presumed-accessible* for *S*.

For example, imagine a student who suffers from anxiety attacks which can be triggered when speaking in front of a group. Presenting in front of the class as part of an assignment results in a higher than tolerable level of risk of an anxiety attack. The educator is aware of this, but is also aware, from extensive conversations with the student that they would like to overcome this anxiety by facing such risks head on. Therefore, they are willing to put themselves at a higher than tolerable level of risk by doing the presentation. In this case, the educator knows that the classroom and assignment are

---

can tolerate may lead to the implementation of protective measures in such a way as to undermine the agency of the student: for example, by assuming that a student will not be able to participate in a fieldwork task without giving them the option to do so. Therefore, it is important to note that whilst educators should err on the side of caution by not assuming that students will be able to take on higher than tolerable levels of risk, this should not undermine the student's agency by imposing protective measures which deny the student agency and choice regarding how they participate in the classroom and the level and kind of risks which they are willing to take on (see for example Charlton, 1998).

[97] Safety alone is not sufficient for presumed-safety since there are other conditions required for presumed-accessibility, unrelated to risk, such as access to resources in a useable format, lack of financial barriers and so on.

accessible to the student because they are willing to take on the level of risk involved.

According to my account of safety, the assignment is not safe for the student because it involves a higher than tolerable level of risk of harm. However, the classroom is *actually-accessible* for the student (because they are willing to take on that level of risk) and *presumed-accessible* from the perspective of the educator (because the educator knows that the student is willing to take on this level of risk). Therefore, the educator can satisfy her responsibility to make the classroom accessible to the student without making the classroom safe for the student (at least in respect to the harm arising from anxiety attacks).

However, educators – and even more so, those making the relevant decisions within educational institutions (regarding timetabling, for instance) – will not always know their students well enough to make this kind of presumption. Therefore, without a great deal of information about a student's attitudes to risk and their ability to take on higher than tolerable level of risk, the only way to presume that the student will not reasonably choose not to enter and participate in the classroom will be to reduce the level of risk to a tolerable level. That is, in the absence of such information, the only way to make the classroom *presumed-accessible* to the student is to make it safe for them.

5.3 Two case studies

So far, I have presented an account of safety in the classroom and argued that making a classroom safe for some student *S* is necessary for making the classroom *presumed-accessible* for *S* in the absence of detailed knowledge of *S's* attitude towards and ability to take on a higher than tolerable level of risk.

In this section, I will work through two case studies to show what the implications of my account are – in particular, what considerations it brings to the foreground regarding the responsibilities of educators.

5.3.1 Content Warnings

> RE-TRAUMATISATION: Malia had a traumatic experience in a house fire several years ago. As a result, she suffers from post-traumatic stress disorder (PTSD) which may be triggered by graphic images and videos of fires. She is studying a course on event management and part of the this covers health and safety – including fire safety. As

> part of one of the classes, the lecturer intends to show a video from an infamous fire at a music venue with inadequate fire safety measures which resulted in numerous fatalities.

Watching this video is very likely to result in re-traumatisation effects for Malia. These might include very distressing physical and psychological effects such as flashbacks and panic attacks. There are a couple of measures which could protect Malia from these effects. Firstly, receiving treatment for her PTSD may be the most effective in significantly reducing the risk of the re-traumatisation effects. However, this may be a very long process and it is likely that this is still ongoing at the time of the class. Secondly, she could avoid watching the video – perhaps by leaving the room when the video is shown – or attempt to mentally prepare herself for viewing it. We might assume that Malia was aware that fire safety would come up as part of the course and knew that there could be material or discussion that may trigger her PTSD symptoms. Given her knowledge of the course content, we might expect her to take some measures herself, such as leaving the room when potentially triggering material will be shown.

However, her ability to take these measures effectively depends on some cooperation from the lecturer. That is, she needs warning in order to be able to leave the room or to prepare herself so that she is able to watch the video. Suppose that the lecturer does not provide a detailed syllabus so that students do not know the topic of each lecture in advance. Near to the beginning of the lecture, he shows the video with no introduction explaining to the students what they are about to see. Malia has no opportunity to prepare herself or to leave the room. Therefore, she is not able to implement any protective measures to reduce the risk of harm from re-traumatisation effects.

What this draws our attention to is how the set-up of the course and the actions of the educators can affect what can reasonably be expected of students. For example, when issues of access and such are considered when designing courses and course materials, this will have an impact on what is required of students to engage with the course and materials. The use of content warnings is an example of how course material design can aid students in implementing protective measures. These may be implemented in 'formal ways' such as explicit statements on syllabi or lectures slides etc. or in a more informal ways such as briefly noting before showing the video that it contains potentially distressing content (not just for someone like Malia with a particular traumatic experience, but for other students as well.)

However, educators will not be able to predict all the ways in which students may be vulnerable to harm. For example, we would not expect a biology lecturer to predict (unless they had been informed beforehand) that one of their students has a severe phobia of bats and therefore provide advance warning that their lecture slides will contain pictures of bats. But in certain cases, the likelihood that one of their students will be affected by a particular issue or have a certain vulnerability is high enough that we can to some degree expect the educator to be sensitive to this in the preparation of their course materials. For example, the high prevalence of sexual assault on university campuses[98] makes it such that, in a lecture of 60 students, there is a significant possibility that at least one of them will have had some previous experience of sexual assault and find discussion of the topic difficult, or potentially harmful.

The lecturer should be aware of this when the content of their lecture and the related materials touch on this topic. The inclusion of certain materials can give rise to access issues. This does not entail (as I will argue below) that certain materials cannot be included, but rather, it places a responsibility on educators to consider their impact on students, and how it might be mitigated in light of their commitment to making the classroom accessible to as many students as possible.

In the following case, I will demonstrate the limitations of the effectiveness of tools such as content warnings, especially when potentially triggering content is embedded throughout a course. In some cases, it may not be possible to make a classroom safe for some students given the nature of the material.

> PHILOSOPHY OF SEX: Lucy was recently the victim of sexual assault. She is taking a module on Philosophy of Sex which will involve discussion of consent and sexual assault. Discussing these topics is likely to be very distressing for Lucy and even with prior warning she is not able to sufficiently prepare herself for these discussions in order to avoid experiencing severe distress. Skipping the sessions in which these topics will definitely

[98] A study by the National Union of Students (2014) found that 25% of respondents reported that they had experience sexual assault during their time at university. Similarly, a survey commissioned by Revolt Sexual Assault (2018) found that 42% of all students and recent graduates surveyed had experienced sexual assault and 48% of female students and recent graduate surveyed had experienced sexual assault.

> be discussed would mean that she misses a significant proportion of the teaching. It is also likely that these topics will be mentioned in other sessions because of the overlap between topics.

In Lucy's case, there is no available protective measure which would reduce the risk of severe distress without disadvantaging her in her learning (besides long-term measures which allow her to reduce the distress she feels when discussing the topic but are unlikely to have much effect by the time she takes the course). The only way to make the classroom safe for her while taking this module would be to remove the content relating to consent and sexual assault. However, this is an important topic to be covered in the course.

The outcome is that the classroom cannot be made safe for Lucy on this module. This means that the course is not presumed-accessible for her since she may reasonably and justifiably choose not to expose herself to the higher than tolerable level of risk of severe distress (though it is also possible that Lucy may choose to take the course knowing that it is unsafe for her). This need not present a problem if Lucy is not required to take the module and is offered sufficient alternatives. Not every classroom can be made safe and presumed-accessible for all students. The justification for designing the course such that the classroom will be unsafe for students like Lucy will depend in part on factors including whether the potentially harmful content is pedagogically valuable and necessary, and whether the course is optional or compulsory. While the inclusion of such materials may very well be justified, the educator's commitment to making the classroom accessible to students wherever possible, will place a burden on them to justify the inclusion of materials which make the classroom unsafe, and therefore not presumed-accessible, for students such as Lucy.

5.3.2 Dignity harms

For my second case study, I turn to a type of case which has often been at the centre of media attention towards 'safety' in the classroom or on campus.[99] This type of case typically involves some form of speech which

[99] Sigal Ben-Porath (2017) discusses a number of incidents which gained media attention in the U.S. including the demands of students at Middlebury College to disinvite the controversial speaker Charles Murray in 2017 on the grounds that his previous work endorsed allegedly racist views (ibid., pp.50-53) and a letter written by the dean of students at University of Chicago to incoming students in 2016 (see fn.29). In a recent case in the U.K. a group of students at the University of Sussex called for the firing of philosopher

students have cited as undermining their safety (or the safety of some other group of students) in the classroom. Typically, the speech in question is deemed to cause harm, especially to minority, marginalised or otherwise vulnerable groups. In such cases students often call on educational institutions to respond.

Consider the following case:[100]

> AFFIRMATIVE ACTION: A class contains students with a mixture of backgrounds, including some who come from areas, or are members of marginalised groups, which typically have a low rate of progression to higher education but whose applications to the university were encouraged and supported by widening access or 'affirmative action' schemes. A conversation about Liberal Education strays slightly off topic and a student makes a comment that 'affirmative action has lowered the quality of education because it lets in a lot of people who shouldn't really be here'. The educator does not intervene and a debate ensues about the truth of this claim. Maya, another student in the class who participated in one of the schemes in question, approaches the educator and complains that the speech was harmful and she should have intervened.

I will not argue here in defence of any particular form of response on the part of the educator. Instead, my focus is on whether or not the institution or educator does in fact have a responsibility to respond to such speech when it occurs in the classroom.

One response to Maya's complaint is the criticism that such intervention on the part of the educator would 'coddle' students (Lukianoff and Haidt, 2019; see also Campbell and Manning, 2018). The thought being that students are asking for protection from something from which they can and should

---

Kathleen Stock on the grounds that her 'gender critical' views caused harm to trans students and other members of the LGBTQIA+ community (Parsons, 2021).

Whilst these cases often extended beyond the classroom itself, a central theme is the claim that harm can arise from the expression, discussion or giving a platform to certain views and the demand from students that the university respond to prevent this harm. As can be seen in the media responses to the Stock case (see for example Adams, 2021; Davies, 2021), many argued that such responses from the university were unnecessary (or worse, violated academic freedom) and the students demanding them were oversensitive.

[100] This case is adapted from an example used by Easton (2018, p. 170).

protect themselves. That is, they should have available to them, a variety of psychological resources (such as traits of character, skills, and patterns of thinking) with which they can reasonably be expected to protect themselves from the harms caused by the speech in question.[101] Furthermore any intervention will deprive students of opportunities to develop such self-protective mechanisms which are required for surviving the world beyond the classroom. Education is supposed to, in part, equip students for the world outside the classroom. The problem with this response is that, as I will show below, these psychological protective measures do not actually protect students against some of the harms which can be caused by such speech.

Speech can cause a range of different harms. Firstly, it may cause harm through offense. At least in principle, the risk of this kind of harm can be reduced by students themselves adopting 'psychological' protective measures, such as, learning to ignore or brush off such speech or developing PECTs. The extent to which such protective measures are available and whether students can reasonably be expected to adopt them cannot be settled in general terms but is context dependent. For this reason, I will not attempt to settle it here but leave open the possibility that those at risk of being harmed can reasonably be expected to protect themselves by adapting their own responses to the speech or by ignoring or brushing it off.

However, it is a mistake to focus solely on offense in such cases (Lawrence III, 1990, p. 461; Applebaum, 2003, p. 155). Another way in which speech can cause harm is by constituting an 'assault on dignity' (Waldron, 2012), where assault on dignity is understood as distinct from offense and distress.[102]

Speech has the potential to cause harm by undermining important goods which contribute to an individual's good standing in her community. These goods are: the entitlements afforded by her good standing in the community, including rights and protections; and the assurance of continued enjoyment of these entitlements - which in part comes from being treated as though she has these entitlements in her everyday interactions with others. Together, these entitlements and assurances amount to what Waldron (2012) calls 'dignity'.

---

[101] See for example Haidt and Lukianoff (2019).

[102] Ben-Porath (2016, 2017) and Callan (2016) both argue that classrooms should be safe from assaults on dignity but need not (and in some cases, should not) be 'intellectually safe' – safe from challenges to fundamentally held beliefs.

A student has good standing in her community when she is afforded equal entitlements, such as the right to take up a proportionate amount of the discussion time and teacher's attention, be taken seriously and treated with respect. She has assurance of these entitlements when her interactions with others in the classroom reflect a respect of these commitments and from this she can assume that she will continue to enjoy these entitlements. For example, if she is regularly ignored, interrupted or talked over by her peers such that she cannot contribute to the discussion, then she lacks assurance that her peers respect her entitlement to take up proportionate discussion time and be taken seriously. She lacks this assurance regardless of whether her peers have any real power to change her entitlements – that is, even if the teacher who has ultimate control over who speaks when, still regards her as having and upholds that entitlement.

Speech can undermine an individual's possession of each of these two goods. It can undermine assurances when it embodies a lack of respect for, or recognition of, an individual's entitlements as a member of the community in good standing. This lack of assurance that one is treated as being a member of the community in good standing and having equal entitlements can have a range of impacts. In their everyday lives, people often take for granted the kinds of treatment they can expect from others and the assumptions they make will influence how they approach interactions with others. If they lack assurance that they will be treated with respect by others, they are likely to approach such interactions with less trust and more reserve. In the classroom, students who lack assurances that they will be treated as members of the community in good standing may be less willing to take intellectual risks or allow themselves to become vulnerable. For example, Elizabeth Ellsworth (1989, pp. 315–16) describes how, without high levels of trust in their peers, her students held back from taking risks and speaking honestly and openly, for fear of being misunderstood or becoming too vulnerable. Because they did not have enough confidence in how their peers would react to such openness, they stayed silent.

A student can lack assurance either because speech in fact embodies a lack of respect of their entitlements (for example, the statement from her peer that 'people who benefitted from affirmative action schemes do not really deserve to be here' communicates a lack of respect for Maya's entitlement to be in the classroom even if the speaker themselves has no power to undermine that entitlement) or because the student perceives a lack of

respect for their entitlements (regardless of the speakers intention or awareness of the impact of their speech).[103]

Lacking assurance has two types of negative effects. One is the insecurity of one's entitlements and the goods they confer. The other is the impact which it has on individual's ability to be able to plan and go about their daily life (Waldron, 2012, pp. 83–4).[104] To some extent, students may be able to reduce the impact of such a lack of assurance by changing the way they respond to it. For example, avoiding 'catastrophising' (Lukianoff and Haidt, 2019, pp. 7–8) and over-generalising or over-inflating the likelihood they will not be treated as equal members of the community based on single, isolated incidences may reduce the impact which a lack of assurance will have on students. However, there are limits to such protective measures, especially when the speech casts legitimate doubt on how they will be more broadly treated within the community and whether their entitlements will be respected.

Secondly, certain forms of speech have the ability to erode individuals' entitlements and therefore undermine the first aspect of dignity. It can do this by warranting certain derogatory or discriminatory treatment[105] or by bringing into question the individual's equal status. For example, the comment that 'some people are only here because of affirmative action' has the potential to cast its targets' entitlement to be in the classroom and status as an intellectual equal into question. Depending on how it is responded to, such a comment can act as to initiate a debate as to whether certain students are really entitled to be in the classroom and whether they have equal status in the learning community of that classroom or educational institution. If such a debate can be had, then those students' status is not secure or established.

This second harm – the questioning of a student's equal status and entitlement to be in the classroom – occurs independently of the student's attitude towards or interpretation of the speech (or even regardless of whether they are aware of the speech). This has important implications for the kind of protective measures which can be implemented to reduce the

---

[103] Micro-aggressions (Sue, 2010) may be an example of this latter kind.

[104] See also John (2011, pp. 70–71) on the 'subjective' element of security.

[105] Applebaum (2003) argues, echoing Butler, that certain kinds of speech in the classroom can support systems of oppression. Tirrell (2017) has also argued that 'toxic' speech can illegitimately rank individuals as inferior and warranting certain inferences about their appropriate treatment.

risk of this kind of harm. 'Psychological' measures, such as changing one's attitude towards the speech, avoiding 'catastrophising' or overgeneralising, do not offer any protection against this second form of harm because they do not prevent entitlements from being undermined.

So long as the speech gives rise to this second form of harm, students cannot be reasonably expected to protect themselves by means of psychological protective measures. There are other means by which the risk of harm from speech can be reduced. For example, Christina Easton (2018) suggests a range of ways in which some 'words that wound' can be addressed through discussion without the need for silencing. These include: showing that certain things are not open questions or that there is a broad consensus on certain empirical questions and as a result, certain views cannot reasonably be defended; engineering discussions towards more sympathetic views and demonstrating that problematic views aired are not widely shared; and bringing to light false assumptions or inferences in problematic views being expressed.

Sometimes students will be able to do these things themselves. But sometimes, given existing power structures or due to the nature of the entitlements being questions this will not be effective. For example, in AFFIRMATIVE ACTION, Maya could attempt to speak up in defence of her entitlement to be in the classroom, but because the speaker already doubts her intellectual abilities, her arguments will not be given any weight. Therefore, such approaches may need to be implemented by another student, whose entitlement to be in the classroom is not under question, or the educator in order to be effective.

5.4 Summary

In this chapter, I have argued that for a classroom to be safe, it need not be completely free from risk of harm. Rather, for a classroom to be safe for some student, the level of risk of harm must be at a level which is tolerable for that student. Further, a classroom is not safe for a student when the risk of harm is reduced to a tolerable level only by the student protecting themselves in ways that cannot be reasonably expected of them. Students cannot be reasonably expected to protect themselves in way which are not genuinely available to them in the classroom qua learning environment, will unfairly disadvantage them, or which cause, or make them vulnerable to further harm (though this list need not be exhaustive).

I then showed that educators very often have a responsibility to ensure that the classroom is safe for their students arising from their commitment to

make the classroom accessible to students wherever possible. To do this, I argued that a classroom cannot be presumed to be accessible to a student when the student could reasonably and justifiably choose not to enter and participate in the classroom on the grounds that doing so would expose them to a higher level of risk than they are willing to take on. Given that educators and relevant decision makers within the educational institution often do not have detailed knowledge of the level of risk of harm to which a student is willing to expose themselves, making the classroom presumed-accessible for that student requires reducing the level of risk of harm to a tolerable level. That is, making the classroom presumed-accessible for the student requires making it safe for them.

Conclusion

This thesis had two central aims. Firstly, to develop an account of safety in the classroom which illuminates the relationship between safety and protection. Secondly, to explore the notion of self-protection within the framework of critical character epistemology.

In chapter 1 I defended an approach according to which safety is defined in terms of objective risk of harm. I argued that this account is able to adequately explain the value of certain conditions which intuitively can make the classroom safe: a lack of objective risk of harm; a lack of subjective risk of harm; and a lack of negative emotional responses. I argued that the value of these things, insofar as it is properly described as 'safety', can be explained in terms of a lack or low level of objective risk of harm. Furthermore, defining safety in terms of objective risk of harm avoids the worry that safety interferes with pedagogical and liberatory goals because it requires the exclusion of too many uncomfortable experiences from the classroom. According to the objective account, uncomfortable experiences only need to be excluded to make the classroom safe if they have the potential to cause harm. This, I argue, is consistent with pedagogical and liberatory aims.

Having defended the objective approach to defining safety, I presented my own account in chapter 5. According to this account a classroom C is safe for some student S if being in and participating in C does not put S at higher than tolerable level of risk of harm without students being required to protect themselves in ways that cannot be reasonably expected of them. I also argued in chapter 5 that, unless educators have detailed knowledge of a student's ability to take on risk, then safety is required to make the classroom presumed-accessible to that student. This is because, without detailed knowledge of students' ability to take on risk, educators cannot assume that students will be able to take on a higher than tolerable level of risk.

As an application of this account of safety, I demonstrated in chapter 2 that a range of protective epistemic character traits (PECTs) can be developed to aid students in protecting themselves against potential harms arising in

a hostile classroom environment. These include both traits which are conventionally recognised as virtues and those which are conventionally recognised as vices. Which PECTs students develop in response to the risk of harm may be influenced by the costs associated with various PECTs and the impact of their social location on the developmental trajectory of their epistemic character.

In chapters 3 and 4, I identified particular harms which can arise from the development of PECTs conventionally identified as epistemic vices. These harms are damage to epistemic character (chapter 3) – character development which takes the individual further from the ideal of flourishing – and the constraint of epistemic agency (chapter 4). The incurrence of these harms is one of the possible costs of developing PECTs which are conventionally identified as vices. As argued in chapter 5, a means of self-protection cannot be reasonably expected of students if it causes or makes them vulnerable to significant harm. Therefore, students cannot be reasonably expected to protect themselves through the development of PECTs which cause them significant harm by damaging their epistemic character and/or constraining their epistemic agency. Further, classrooms in which students must protect themselves via the development of such PECTs in order to reduce the risk of harm to a tolerable level are not safe for those students.

There are two implications of my account of safety in the classroom worth noting. Firstly, it will not always be possible to make a classroom safe for all students. Some students may have particular vulnerabilities to harm which mean that reducing the level of risk of harm to below tolerable level is not possible or would require protective measure which are not practicable in the classroom. Furthermore, there may be conflicts between the requirements of making the classroom safe for different students. For example, imagine a scenario in which one student has a medical assistance dog who alerts them if they are about to have a seizure and another student who has a severe allergy to dogs. Without the dog, the risk of harm from seizures for the first student is higher than tolerable. But if the dog is present, then there is a higher than tolerable level of risk of harm to the second student due to her allergy.

Given these complexities, it would be too demanding to require that educators and educational institutions make all classrooms safe for all students. Rather, given the link between safety and accessibility, the commitment to making education accessible wherever possible should create a burden on educators and educational institutions to justify why a classroom cannot be made safe for a student.

Secondly, the account I have offered is demanding in the sense that, making a classroom safe for a student can require a great deal of knowledge on the part of the educator or relevant decision makers within the educational institution. As I have highlighted, what counts as safe for a particular student is dependent on many individual and contextual factors. These include: their particular vulnerabilities to harm; the extent to which their character development has been shaped by relationships of power and oppression and the trajectory in which their character is able to develop as a result of this; the ways in which their peers will respond to them when they exercise various traits; the support networks they have available to them; and the extent to which they can tolerate, recover from or mitigate harms. It is unlikely that educators will have detailed knowledge of these things. They might not anticipate all of the things which cause harm to students. As discussed in §5.3.1, students can have a range of phobias or be vulnerable to retraumatisation by a range of triggers of which educators may not be aware. Anticipating all of the potential sources of harm will be an impossible task for educators. It may also be an unhelpful exercise – pre-empting a maximally broad range of potential sources of harm for students, just on the basis that they could be potentially harmful to some student would be excessively restrictive.

Furthermore, as safety is dependent on students not being required to protect themselves in ways that cannot be reasonably expected of them, knowing whether a classroom is safe for some student depends on knowing what can reasonably be expected of them. And this will require knowledge of many of the factors listed above.

As a result, it may be seen as a limitation of my account that it is very difficult, or perhaps impossible, to put into practice given the demands made of educators and educational institutions. However, the aim of this thesis is not to provide an account of safety that is immediately practicable for educators and educational institutions. Rather, I have presented a framework into which relevant information can be inputted to answer specific questions about whether a particular classroom is safe for a particular student. Further work must be done to provide practical guidance to educators about what safety means in any particular classroom.

But whilst the context dependence of safety makes it difficult to produce any general guidance on making classrooms safe, the account I have defended does provide a framework for interpreting individual cases in which the safety of the classroom is called into question by explaining the relationships between the various factors to be considered when evaluating the safety of the classroom.

Bibliography

Adams, R. (2021) *Kathleen Stock says she quit university post over 'medieval' ostracism*, *The Guardian*. Available at: https://www.theguardian.com/education/2021/nov/03/kathleen-stock-says-she-quit-university-post-over-medieval-ostracism (Accessed: February 5, 2022).

Alfano, M. (2013) "Identifying and defending the hard core of virtue ethics," *Journal of Philosophical Research*, 38, pp. 233–260.

Alfano, M. and Robinson, B. (2017) "Gossip as a Burdened Virtue," *Ethical Theory and Moral Practice*, 20(3), pp. 473–487.

Allen, P. (1973) "Free Space," in Koedt, A., Levine, E., and Rapone, A. (eds) *Radical Feminism*. Chicago: Quadrangle Books, pp. 271–279.

Anderson, D. (2021) "An Epistemological Conception of Safe Spaces," *Social Epistemology*, 35(3), pp. 285–311.

Applebaum, B. (2003) "Social justice, democratic education and the silencing of words that wound," *Journal of Moral Education,* 32:2, pp. 151-162

Arao, B. and Clemens, K. (2013) "From Safe Spaces to Brave Spaces: A New Way to Frame Dialogue Around Diversity and Social Justice," in Landreman, L.M. (ed.) *The Art of Effective Facilitation: Reflections from Social Justice Educators*. Sterling: Stylus Publishing, pp. 135–150.

Aristotle (2002) *Nicomachean Ethics*. Translated by S. Broadie and C. Rowe. Oxford: Oxford University Press.

Arthur, J., Kristjánsson, K. and Thoma, S. (2016) *Is Grit the Magic Elixir of Good Character?* Jubilee Centre for Character and Virtues. Available at: https://www.jubileecentre.ac.uk/userfiles/jubileecentre/pdf/insight-series/Is_Grit_the_Magic_Elixir_of_Good_Character_InsightSeries2016.pdf (Accessed: November 14, 2020)

Astell, M. (1999) *A Serious Proposal to the Ladies: Parts I and II*. Edited by P. Springborg. Ontario: Broadview Literary Texts.

Austin, J.L. (1975) *How to Do Things with Words*. Cambridge: Harvard University Press.

Baehr, J. (2011) *The Inquiring Mind: On Intellectual Virtues and Virtue Epistemology*. Oxford: Oxford University Press.

Baehr, J. (ed.) (2016) *Intellectual Virtues and Education: Essays in Applied Virtue Epistemology*. New York: Routledge.

Baehr, J.S. (2006) "Character in epistemology," *Philosophical Studies*, 128(3), pp. 479–514.

Baril, A. (2019) "What makes the epistemic virtues valuable?," in Battaly, H. (ed.) *The Routledge Handbook of Virtue Epistemology*. London: Routledge, pp. 69–80.

Barnes, E. (2020) "Gender and Gender Terms," *Nous*, 54(3), pp. 704–730.

Bartky, S. (1990) *Femininity and Domination: Studies in the Phenomenology of Oppression*. New York: Routledge.

Battaly, H. (2013) "Detecting epistemic vice in higher education policy: Epistemic insensibility in the seven solutions and the REF," *Journal of Philosophy of Education*, 47(2), pp. 263–280.

Battaly, H. (2014) "Varieties of Epistemic Vice," in Matheson, J. and Vitz, R. (eds) *The Ethics of Belief*. Oxford: Oxford University Press.

Battaly, H. (2017) "Intellectual Perseverance," *Journal of Moral Philosophy*, 14, pp. 669–697.

Battaly, H. (2018) "Can Closed-mindedness be an Intellectual Virtue?," *Royal Institute of Philosophy Supplement*, 84, pp. 23–45.

Battaly, H. (2019) "A Third Kind of Intellectual Virtue: Personalism," in Battaly, H. (ed.) *The Routledge Handbook of Virtue Epistemology*. London: Routledge, pp. 115–126.

Battaly, H. (2020) "Closed-Mindedness As an Intellectual Vice," in Kelp, C. and Greco, J. (eds) *Virtue Theoretic Epistemology: New Methods and Approaches*. Cambridge: Cambridge University Press, pp. 15–41.

Battaly, H. (2021) "Countering Servility through Pride and Humility," *Midwest Studies in Philosophy*, 45, pp. 333–370.

Begum, H. (2015) *Safe and self-defining spaces: Solidarity with Goldsmith SU BME Network, University College London Union*. Available at: uclu.org/articles/safe-and-self-defining-spaces-solidarity-with-goldsmith-su-bme-network (Accessed: May 3, 2018).

Bennett, J. (1990) "Why is belief involuntary," *Analysis*, 50, pp. 87–107.

Ben-Porath, S. (2016) "Safety, Dignity and the Quest for a Democratic Campus Culture," *Philosophical Inquiry in Education*, 24(1), pp. 79–85.

Ben-Porath, S. (2017) *Free Speech on Campus*. Philadephia: University of Pennsylvania Press.

Benson, P. (1994) "Free Agency and Self-Worth," *The Journal of Philosophy*, 91(12), pp. 650–668.

Berenstain, N. (2016) "Epistemic Exploitation," *Ergo*, 3(22), pp. 569–590.

Bergamini, M. (2018) *Are "Safe Spaces" In Universities Breeding Intolerance?*, *The Huffington Post*. Available at: https://www.huffingtonpost.co.uk/entry/are-safe-spaces-in-universities-breeding-intolerance_uk_5aa6871be4b06b6a749ce3c4 (Accessed: November 8, 2020).

Berman, J. (2001) *Risky Writing: Self-Disclousure and Self-Transformation in the Classroom*. Amherst: Univesity of Massachusetts Press.

Bishop, J. (1989) *Natural Agency: An Essay on the Causal Theory of Action*. Cambridge: Cambridge University Press.

du Bois, W.E.B. (2015) *The Souls of Black Folk*, *Oxford University Press*. New Haven: Yale University Press.

Boler, M. (1999) *Feeling Power: Emotions and Education*. London: Taylor & Francis.

Boostrom, R. (1998) "'Safe spaces': Reflections on an educational metaphor," *Journal of Curriculum Studies*, 30(4), pp. 397–408.

Bouattia, M. (2018) *The Narrative On University Safe Spaces And No-Platform Policies Couldn't Be Further From The Truth*, *The Huffington Post*. Available at: https://www.huffingtonpost.co.uk/malia-bouattia/free-speech-campus-safe-spaces_b_14678920.html (Accessed: November 8, 2020).

Brar, C. (2015) *We asked what you think about Milo Yiannapolous coming to Bristol*, *The Tab Bristol*. Available at: https://thetab.com/uk/bristol/2015/10/23/asked-think-milo-yiannopoulos-coming-bristol-19977 (Accessed: March 24, 2022).

Bristol Students' Union (2014) *Safe Space Policy*. Available at: https://www.bristolsu.org.uk/resources/safe-space-policy-f299 (Accessed: March 24, 2022)

Brogaard, B. (2012) "Intellectual Flourishing as the Fundamental Epistemic Norm," in Turri, J. and Littlejohn, C. (eds) *Epistemic Norms*. Oxford: Oxford University Press, pp. 11–33.

Callan, E. (2016) "Education in Safe and Unsafe Spaces," *Philosophical Inquiry in Education*, 24(1), pp. 64–78.

Campbell, B. and Manning, J. (2018) *The Rise of Victimhood Culture: Microaggressions, Safe Spaces, and the New Culture Wars*. London: Palgrave Macmillan.

Card, C. (1996) *The Unnatural Lottery: Character and Moral Luck*. Philadephia: Temple University Press.

Carter, A.M. (2015) "Teaching with Trauma: Trigger Warnings, Feminism, and Disability Pedagogy," *Disability Studies Quarterly*, 35(2).

Cassam, Q. (2019) *Vices of the Mind*. Oxford: Oxford University Press.

Césaire, A. (1972) *Discourse on Colonialism*. Translated by J. Pinkman. New York: Monthly Review Press.

Charlton, J. (1998) *Nothing About Us Without Us: Disability Oppression and Empowerment*. Oakland: University of California Press.

Code, L. (1984) "Towards a 'Responsibilist' Epistemology," *Philosophy and Phenomenological Research*, 45(1), pp. 29–50.

Cooper, J. (2008) *Cognitive Dissonance: Fifty Years of a Classic Theory*. London: SAGE.

Crerar, C. (2018) "Motivational Approaches to Intellectual Vice," *Australasian Journal of Philosophy*, 96(4), pp. 753-766.

Crerar, C. (2021) "Must We Love Epistemic Goods?," *Philosophical Quarterly*, 71(4), pp. 886–907.

Cunliffe, R. (2021) *If Rod Liddle has the right to lecture Durham students about sex work, where's my invitation to talk about cats?*, *The Newstatesman*. Available at: https://www.newstatesman.com/society/2021/12/if-rod-liddle-has-the-right-to-lecture-durham-students-about-sex-work-wheres-my-invitation-to-talk-about-cats (Accessed: February 22, 2022).

Curren, R. (2019) "Virtue Epistemology and Eduacation," in Battaly, H. (ed.) *The Routledge Handbook of Virtue Epistemology*. New York: Routledge, pp. 470–482.

Daukas, N. (2019) "Feminist Virtue Epistemology," in Battaly, H. (ed.) *The Routledge Handbook of Virtue Epistemology*. London: Routledge, pp. 379–391.

Davies, J. (2021) *Academics warn universities should stop pandering to students who must be told to "accept academic freedoms or LEAVE" after professor was "hounded out" of her role in trans row*, *Mail Online*. Available at: https://www.dailymail.co.uk/news/article-10147039/Academics-warn-universities-stop-pandering-students.html (Accessed: March 28, 2022)

Dillon, R.S. (2012) "Critical Character Theory," in Crasnow, S.L. and Superson, A.M. (eds) *Out from the Shadows: Analytical Feminist Contributions to Traditional Philosophy*. Oxford: Oxford University Press, pp. 83–114.

Dotson, K. (2011) "Tracking Epistemic Violence, Tracking Practices of Silencing," *Hypatia*, 26(2), pp. 236–257.

Dotson, K. (2014) "Conceptualizing Epistemic Oppression," *Social Epistemology*, 28(2), pp. 115–138.

Dweck, C.S. (2006) *Mindset: The new psychology of success*. New York: Random House.

Earl, S. and Arthur, J. (no date) *Character and Academic Attainment: Does Character Education Matter?* Jubilee Centre for Character and Virtues. Available at: https://www.jubileecentre.ac.uk/userfiles/jubileecentre/pdf/insight-series/Insight_Briefing_Paper_JA_SE.pdf (Accessed November 14, 2022)

Easton, C. (2018) "'Words that Wound' in the classroom: Should they be silenced or discussed?," in Downs, D.A. and Surprenant, C.W. (eds) *The Value and Limits of Academic Speech: Philosophical, Political, and Legal Perspectives*. New York: Routledge.

Ecclestone, K. (2004) "Learning or Therapy? The Demoralisation of Education," *British Journal of Educational Studies*, 52(2), pp. 112–137.

Ecclestone, K. (2008) *The Dangerous Rise of Therapeutic Education*. London: Routledge.

Echols, A. (1989) *Daring to Be Bad: Radical Feminism in America 1967-1975*. Minneapolis: University of Minnesota Press.

Eddo-Lodge, R. (2017) *Why I'm No Longer Talking to White People About Race*. London: Bloomsbury Circus.

Edinburgh University Students' Union (no date) *Safe Space Policy*. Available at: https://www.eusa.ed.ac.uk/about/policy/internal/safespaces (Accessed: March 24, 2022)

Ellison, J. (2020) *Letter to the class of 2020*. Available at: https://news.uchicago.edu/sites/default/files/attachments/Dear_Class_of_2020_Students.pdf (Accessed: March 28, 2022)

Ellsworth, E. (1989) "Why doesn't this feel empowering? Working through the repressive myths of critical pedagogy," *Harvard Educational Review*, 59(3), pp. 297–324.

Emler, N. (2001) *Self-esteem: The costs and causes of low self-worth*. York: York Publishing Services.

Equality Act 2010 [online]. Available at: https://www.legislation.gov.uk/ukpga/2010/15/contents (Accessed: March 25, 2022).

Equality and Human Rights Commission (2014) *Equality Act 2010 Technical Guidance on Further and Higher Education*. Available at: https://www.equalityhumanrights.com/sites/default/files/equalityact2010-technicalguidance-feandhe-2015.pdf (Accessed: March 28, 2022)

Fisher, B.M. (2000) *No Angel in the Classroom: Teaching through Feminist Discourse*. Lanham: Rowman & Littlefield Publishers.

Freres, D.R. *et al.* (2002) "Preventing depressive symptoms in middle school students: The Penn Resiliency Program," *International Journal of Emergency Mental Health*, 4(1), pp. 31–40.

Fricker, M. (2007) *Epistemic Injustice: The Ethics and Power of Knowing*. Oxford: Oxford University Press.

Furedi, F. (2003) *Therapy Culture: Cultivating Vulnerability in an Uncertain Age*. London: Routledge.

Garran, A.M. and Rasmussen, B.M. (2014) "Safety in the Classroom: Reconsidered," *Journal of Teaching in Social Work*, 34(4), pp. 401–412.

Gillham, J.E. *et al.* (2012) "Resilience Education," in David, S., Boniwell, I., and Ayers, A.C. (eds) *Oxford Handbook of Happiness*. Oxford: Oxford University Press, pp. 657–680.

Goldsmiths Students' Union (2012) *Goldsmiths Students' Union Policy: Safe Space Policy*. Available at: https://www.goldsmithssu.org/pageassets/yourunion/governance/policies/Safe-Space-Policy.pdf (Accessed: March 24, 2022)

Hantzis, D.M. and Looser, D. (1995) "Of Safe(r) Spaces and 'Right' Speech: Feminist Histories, Loyalties, Theories, and the Dangers of Critique," in Williams, J. (ed.) *PC Wars: Politics and Theory in the Academy*. New York: Routledge, pp. 222–252.

Haslanger, S. (2012) *Resisting Reality: Social Construction and Social Critique*. Oxford: Oxford University Press.

Hieronymi, P. (2008) "Responsibility for believing," *Synthese*, 161(3), pp. 357–373.

Holley, L.C. and Steiner, S. (2005) "Safe Space: Student Perspectives on Classroom Environment," *Journal of Social Work Education*, 41(1), pp. 49-64.

hooks, bell (1994) *Teaching to Trangress: Education and the Practice of Freedom*. New York: Routledge.

Hooks, B. (1990) *Yearning: Race, Gender, and Cultural Politics*, *Yearning: Race, Gender, and Cultural Politics*. Boston: South End Press.

Hookway, C. (2010) "Some Varieties of Epistemic Injustice," *Episteme*, 7(02), pp. 151–178.

Hornsby, J. and Langton, R. (1998) "Free Speech and Illocution," *Legal Theory*, 4(01), pp. 21-37.

Jenkins, K. (2016) "Amelioration and Inclusion: Gender Identity and the Concept of Woman*," *Ethics*, 126(2), pp. 394–421.

John, S. (2011) "Security, Knowledge and Well-being," *Journal of Moral Philosophy*, 8, pp. 68–91.

Johnson, C. (2021) "Teaching to the Test: How Schools Discourage Phronesis," in Kidd, I.J., Battaly, H., and Cassam, Q. (eds) *Vice Epistemology*. Abingdon: Routledge, pp. 225–239.

Jones, K. (2012) "The Politics of Intellectual Self-trust," *Social Epistemology*, 26(2), pp. 237–251.

Jones, W.E. (2012) "Higher Education, Academic Communities, and the Intellectual Virtues," *Educational Theory*, 62(6), pp. 695–711.

Jubilee Centre for Character and Virtues (2012) *A Framework for character education in schools*. Available at: https://www.jubileecentre.ac.uk/527/character-education/framework (Accessed: November 14, 2020).

Kenney, M.R. (2001) *Mapping Gay L.A.: The intersection of Place and Politics*. Philadelphia: Temple University Press.

Kessler, M.J. (2018) "The Difference Between Being Offended and Taking Offence," in Downs, D.A. and Surprenant, C.W. (eds) *The Value and Limits of Academic Speech: Philosophical, Political, and Legal Perspectives*. New York: Routledge, pp. 136–154.

Kidd, I.J. (2016) "Intellectual Humility, Confidence, and Argumentation," *Topoi*, 35(2), pp. 395–402.

Kidd, I.J. (2019a) "Epistemic Corruption and Education," *Episteme*, 16(2), pp. 220–235.

Kidd, I.J. (2019b) "Epistemic Courage and the Harms of Epistemic Life," in Battaly, H. (ed.) *The Routledge Handbook of Epistemic Injustice*. London: Routledge, pp. 244–255.

Kidd, I.J. (2020) "Epistemic Corruption and Social Oppression," in Kidd, I.J., Cassam, Q., and Battaly, H. (eds) *Vice Epistemology: Theory and Practice*. London: Routledge, pp. 69–85.

Kidd, I.J. (2021) "Character, Corruption, and 'Cultures of Speed' in Higher Education," in Mahon, Á. (ed.) *The Promise of the University: Reclaiming Humanity, Humility and Hope*. New York: Springer, pp. 17–28.

Kidd, I.J. (forthcoming) "From vice epistemology to critical character epistemology," in Alfano, M., Klein, C., and de Ridder, J. (eds) *Social Virtue Epistemology*. New York: Routledge.

Kings College London Students' Union (no date) *KCLSU Safe Space Policy*. Available at: https://www.kclsu.org/pageassets/safespace/KCLSU-Safe-Space-Policy.pdf.

Kotzee, B. (2017) "Education and Epistemic Injustice," in Kidd, I.J., Medina, J., and Pohlhaus, G. (eds) *The Routledge Handbook of Epistemic Injustice*. London: Routledge, pp. 324–335.

Lackey, J. (2021) "Eyewitness Testimony and Epistemic Agency," *Noûs*. [Pre-print] DOI: https://doi.org/10.1111/nous.12380

Lawford-Smith, H. (manuscript) *Women-only spaces and the right to exclude*. Available at: https://philpapers.org/rec/LAWWSA-4. (Accessed: March 12, 2022).

Lawrence III, C.R. (1990) "If the Hollers Let Him Go: Regulating Racist Speech on Campus," *Duke Law Journal*, 1990, pp. 431-483.

Leeds Revolutionary Feminist Group (1981) *Love your enemy?: the debate between heterosexual feminism and political lesbianism*. London: Onlywomen Press.

Little, L. (2016) *Women of colour need their own space - so I've set up a feminist magazine*, *The Guardian*. Available at: https://www.theguardian.com/commentisfree/2016/mar/23/feminists-created-equal-black-women (Accessed: November 8, 2020).

Ludlow, J. (2004) "From Safe Space to Contested Space in the Feminist Classroom," *Transformations: The Journal of Inclusive Scholarship and Pedagogy*, 15(1), pp. 40–56.

Lukianoff, G. and Haidt, J. (2019) *The Coddling of the American Mind: How Good Intentions and Bad Ideas are Setting Up a Generation for Failure*. London: Penguin Books.

Lynch, M.P. (2018) "Arrogance, truth and public discourse," *Episteme*, 15(3), pp. 283–296.

Mackenzie, C. (2008) "Relational autonomy, normative authority and perfectionism," *Journal of Social Philosophy*, 39(4), pp. 512–533.

Mackenzie, C. (2014) "Responding to the agency dilemma autonomy, adaptive preferences, and internalized oppression," in Marina A.L. Oshana (ed.) *Personal Autonomy and Social Oppression: Philosophical Perspectives*. New York: Routledge, pp. 48–67.

Manners, P. (2019) 'Trans Inclusion in Women Only Spaces,' *The Journal of Contemporary Community Education Practice Theory,* 10 (1)

Manson, N.C. (2012) "Epistemic restraint and the vice of curiosity," *Philosophy*, 87(2), pp. 239–259.

McHugh, C. (2013) "Epistemic Responsibility and Doxastic Agency," *Philosophical Issues*, 23, pp. 132–157.

Medina, J. (2013) *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and Resistant Imaginations*. Oxford: Oxford University Press.

Mele, A. (ed.) (1997) *The Philosophy of Action*. Oxford: Oxford University Press.

Meyer, I.H. (2003) "Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: Conceptual issues and research evidence," *Psychological Bulletin*, 129(5), pp. 674–697.

Meyer, I.H. (2007) "Prejudice and discrimination as social stressors," in Meyer, I.H. and Northridge, M.E. (eds) *The Health of Sexual Minorities*. Washington, D.C.: APA, pp. 242–267.

Mills, C.W. (2007) "White ignorance," in S. Sullivan and N. Tuana (eds) *Race and Epistemologies of Ignorance*. Albany: SUNY Press, pp. 13–38.

Montmarquet, J.A. (1993) *Epistemic Virtue and Doxastic Responsiblity*. Milton Keynes: Rowman & Littlefield.

Monypenny, A. (2021) "Between vulnerability and resilience: A contextualist picture of protective epistemic character traits," *Journal of Philosophy of Education*, 55(2), pp. 358–370.

National Union of Students (2014) *Hidden Marks: A study of women students' experiences of harassment, stalking, violence and sexual assault*. Available at: https://www.nusconnect.org.uk/resources/hidden-marks-a-study-of-women-students-experiences-of-harassment-stalking-violence-and-sexual-assault (Accessed: February 5, 2022).

Noble, T. and McGrath, H. (2012) "Well-being and resilience in education," in David, S., Boniwell, I., and Ayers, A.C. (eds) *The Oxford Handbook of Happiness*. Oxford: Oxford University Press, pp. 609–626.

Nussbaum, M.C. (1986) *The Fragility of Moral Goodness: Luck and Ethics in Greek Tragedy and Philosophy*. Cambridge: Cambridge University Press.

Olson, D. (2015) "A Case for Epistemic Agency," *Logos & Episteme*, 6(4), pp. 449–474.

Orozco, J. (2013) "On the Limits of Virtue Epistemology," *American Philosophical Quarterly*, 50(2), pp. 103–120.

Palfrey, J. (2017) *Safe Spaces, Brave Spaces: Diversity and Free Expression in Education*. Cambridge: The MIT Press.

Parsons, V. (2021) *LGBT+ students claim 'monumental victory' after anti-trans professor Kathleen Stock quits*, *Pink News*. Available at: https://www.pinknews.co.uk/2021/10/29/kathleen-stock-trans-sussex/ (Accessed: February 5, 2022).

Patton, L.D. (2006) "The voice of reason: A qualitative examination of black student perceptions of black culture centers," *Journal of College Student Development*, 47(6), pp. 628–646.

Pells, R. (2017) *University "safe space" policies levase academics in fear of losing their jobs, claims professor*, *The Independent*. Available at: https://www.independent.co.uk/news/education/education-news/university-safe-spaces-academics-professors-fear-lose-jobs-students-free-speech-political-correct-pc-dennis-hayes-a7815991.html (Accessed: November 8, 2020).

Pohlhaus Jr, G. (2014) "Discerning the Primary Epistemic Harm in Cases of Testimonial Injustice," *Social Epistemology*, 28(2), pp. 99–114.

Radicalesbians (1970) *Woman Identified Woman*. Available at: https://repository.duke.edu/dc/wlmpc/wlmms01011 (Accessed: March 28, 2022)

Ramsay, P. (2017) "Is Prevent a safe space?," *Education, Citizenship and Social Justice*, 12(2), pp. 143-158.

Redmond, M. (2010) "Safe Space Oddity: Revisiting Critical Pedagogy," *Journal of Teaching in Social Work*, 30(1), pp. 1–14.

Revolt Sexual Assault (2018) *Students' experiences of sexual violence*. Available at: https://revoltsexualassault.com/wp-content/uploads/2018/03/Report-Sexual-Violence-at-University-Revolt-Sexual-Assault-The-Student-Room-March-2018.pdf (Accessed: March 28, 2022)

Roberts, R.C. and Wood, W.J. (2007) *Intellectual Virtues*. Oxford: Oxford University Press.

Roth, M.S. (2019) *Don't Dismiss "Safe Spaces," The New York Times*. Available at: https://www.nytimes.com/2019/08/29/opinion/safe-spaces-campus.html (Accessed: November 8, 2020).

Rothschild, E. (1995) "What is Security?," *Daedalus*, 124(3), pp. 53–98.

Russell, J.S. (2015) "Resilience," *Journal of the Philosophy of Sport*, 42(2), pp. 159–183.

Sanders, K.N. (2014) "Black Culture Centers: A Review of Pertinent Literature Kimberly N. Sanders," *Urban Education Research and Policy Annuals*, 4(1), pp. 30–38.

Sarachild, K. (1978) "Consciousness-Raising: A Radical Weapon," in Redstockings (ed.) *Feminist Revolution*. New York: Random House.

Shahvisi, A. (2018) "From Academic Freedom to Academic Responsibility: Privileges and Responsibilities Regarding Speech on Campus," in Downs, D.A. and Suprenant, C.W. (eds) *The Value and Limits of Academic Speech: Philosophical, Political, and Legal Perspectives*. New York: Routledge, pp. 266–284.

Sosa, E. (1985) “Knowledge and Intellectual Virtue,” *The Monist*, 68(2), pp. 226–245.

Sosa, E. (2013) “Epistemic agency,” *Journal of Philosophy*, 110(11), pp. 585–605.

Steele, C., Spencer, S. and Aronson, J. (2002) “Contending with images of one’s group: the psychology of stereotype and social identity threat,” in Zanna, M.P. (ed.) *Advances in Experimenal Social Psychology vol. 34*. San Diego: Academic Press, pp. 379–440.

Steele, C.M. and Aronson, J. (1995) “Stereotype Threat and the Intellectual Test Performance of African Americans,” *Journal of Personality and Social Psychology*, 69(5), pp. 797–811.

Stengel, B.S. (2010) “The Complex Case of Fear and Safe Space,” *Studies in Philosophy and Education*, 29, pp. 523–540.

Stokas, A.G. (2015) “A Genealogy of Grit: Education in the New Gilded Age,” *Educational Theory*, 65(5), pp. 513–528.

Students’ Union University of Manchester (no date) *Safe Space Policy*. Available at: https://s3-eu-west-1.amazonaws.com/nusdigital/document/documents/28180/a4e86ca9c4fe8db8c609304706e32e89/Safe_Space_Policy_AMENDED_FOLLOWING_SENATE.pdf (Accessed: March 24, 2022)

Sue, D.W. (2010) *Microaggressions in Everyday Life: Race, Gender, and Sexual Orientation*. Hoboken: Wiley.

Sussex Students’ Union (no date) *Safer space & good conduct policy*. Available at: https://sussexstudent.com/about-us/policies/safer-space (Accessed: March 24, 2022)

Tanesini, A. (2016) “‘Calm down, dear’: Intellectual arrogance, silencing and ignorance,” *Aristotelian Society Supplementary Volume*, 90(1), pp. 71–92.

Tanesini, A. (2018) “Intellectual Servility and Timidity,” *Journal of Philosophical Research*, 43, pp. 21–42.

Tanesini, A. (2020) “Ignorance, arrogance, and privilege: Vice epistemology and the epistemology of ignorance,” in *Vice Epistemology*, pp. 53–68.

Tanesini, A. (2021) *The Mismeasure of the Self: A Study in Vice Epistemology*. Oxford: Oxford University Press.

Tessman, L. (2001) “Critical Virtue Ethics: Understanding Opppression as Morally Damaging,” in DesAutels, P. and Waugh, J. (eds) *Feminists doing ethics*. Rowman & Littlefield.

Tessman, L. (2005) *Burdened Virtues: Virtue Ethics for Liberatory Struggles*. Oxford: Oxford University Press.

The Mancunion (2015) *"Transphobe" Julie Bindel banned from Free Speech debate*. Available at: https://mancunion.com/2015/10/07/update-yiannopoulos-also-banned-from-censorship-event/ (Accessed: March 21, 2022).

The Roestone Collective (2014) "Safe space: Towards a reconceptualization," *Antipode*, 46(5), pp. 1346–1365.

Tirrell, L. (2017) 'Toxic Speech: Toward an Epidemiology of Discursive Harm,' *Philosophical Topics*, 45(2), pp. 139-161.

Tuana, N. (2006) "The Speculum of Ignorance: The Women's Health Movement and Epistemologies of Ignorance," *Hypatia: A Journal of Feminist Philosophy*, 21(3), pp. 1–19.

Valk, A.M. (2002) "Living a Feminist Lifestyle: The Intersection of Theory and Action in a Lesbian Feminist Collective," *Feminist Studies*, 28(2), pp. 303–332.

Wadham, J. *et al.* (2017) *Blackstone's Guide to the Equality Act 2010*. Oxford: Oxford University Press.

Waldron, J. (2012) *The Harm of Hate Speech*. Cambridge: Harvard University Press.

Walker, D. (no date) *How Can Resilience be Developed in UK Schools?* Available at: https://www.jubileecentre.ac.uk/userfiles/jubileecentre/pdf/insight-series/WalkerD.pdf (Accessed: November 11, 2019).

Watson, L. (2021) "Vices of Questioning in Public Discourse," in Kidd, I.J., Battaly, H., and Cassam, Q. (eds) *Vice Epistemology*. Abingdon: Routledge, pp. 239–258.

Webster, A. (2021) "Socially Embedded Agency Lessons from Marginalized Identities," in *Oxford Studies in Agency and Responsibility: Volume 7*. Oxford: Oxford University Press, pp. 104–129.

Whitten, S. (2018) *Why "safe spaces" at universities are a threat to free speech*, *The Conversation*. Available at: https://theconversation.com/why-safe-spaces-at-universities-are-a-threat-to-free-speech-94547 (Accessed: November 8, 2020).

Williams, B. (1973) *Problems of the Self*. Cambridge: Cambridge University Press.

Wilson, J.W. (2018) *I've never had a student ask for a safe space. Here's what they ask for.*, *Vox*. Available at: https://www.vox.com/first-person/2018/12/12/18131186/college-campus-safe-spaces-trigger-warnings (Accessed: November 8, 2020).

Winans, A.E. (2005) "Local pedagogies and race: Interrogating white safety in the rural college classroom," *College English*, 67(3), pp. 253–273.

Wise, E. (2015) *FemSoc oppose journalism event featuring controversial speaker Milo Yiannopoulos as it 'violates safe space policy,' The Bristol Tab*. Available at: https://thetab.com/uk/bristol/2015/10/19/femsoc-members-consider-journalism-society-boycott-over-event-featuring-controversial-journalist-milo-yiannopoulos-19833 (Accessed: March 24, 2022).

Wolfe-Robinson, M. (2021) *Durham head steps back after calling students 'pathetic' at Rod Liddle event*, *The Guardian*. Available at: https://www.theguardian.com/education/2021/dec/09/durham-head-steps-back-after-calling-students-pathetic-at-rod-liddle-event (Accessed: March 28, 2022)

Wolff, J. and De-Shalit, A. (2007) *Disadvantage*. Oxford: Oxford University Press.

Wollstonecraft, M. (2004) *A Vindication of the Rights of Woman*. Edited by M. Brody. London: Penguin Books.

Yiannopoulos, M. (2015) *I'm so done with the trans outrage brigade: Why I'm supporting "drop the T," Breitbart News*. Available at: https://www.breitbart.com/politics/2015/11/10/i-am-so-done-with-the-trans-outrage-brigade-why-im-supporting-drop-the-t/ (Accessed: March 24, 2022).

Zagzebski, L. (1996) *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge.* Cambridge: Cambridge University Press.

Zagzebski, L. (2001) "Must Knowers be agents?," in Fairweather, A. and Zagzebski, L. (eds) *Virtue Epistemology: Essay in epistemic virtue and responsibility*. Oxford: Oxford University Press, pp. 142–157.