

"/--\", Jone Bengoa, 2015





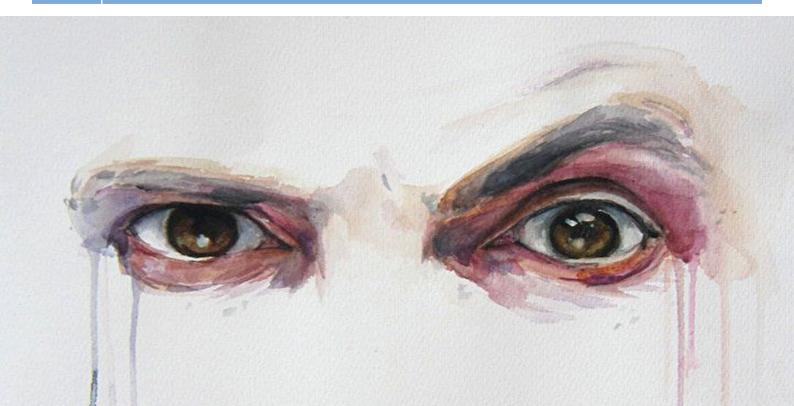




# Wednesday, 19 June

Time	Sibson	Sibson
	Lecture Theatre 2	Seminar Room 5

9.30	WELCOME & COFFEE IN THE SIBSON ATRIUM		
10.00	Plenary Session 1: Lecture Theatre 2  In defence of vengeful anger  Krista Thomason  (Swarthmore College)		
11.15	Mercedes Corredor (University of Michigan) Policing unintentional slights: bad sex and moral disapproval	Costanza Porro (King's College London) State blame for criminal wrongdoing: a skeptical view	
12.15	LUNCH IN THE	SIBSON ATRIUM	
13.15	Plenary Session 2: Lecture Theatre 2  A hope standard for punishment  Kimberley Brownlee  (University of Warwick)		
14.30	Daniel Telech & Leora Dahan Katz (Van Leer Jerusalem Institute, Polonsky Academy) Punitive disappointment	Jake Wojtowicz (King's College London) Agent-regret and punishment	
15.30	COFFEE IN THE	SIBSON ATRIUM	
16.00	Manuela Irarrázabal (University College London) Punishment, pleasure, and pain in Greek tragedy	Andreas Carlsson (University of Oslo) Guilt, blameworthiness, and the point of blame	
17.00	Plenary Session 3: Lecture Theatre 2  Suffering and punishment  Michael Brady  (University of Glasgow)		
18.15 WINE RECEPTION IN THE SIBSON ATRIU		THE SIBSON ATRIUM	
19.30	CONFERENCE DINNER		



Time	Sibson Lecture Theatre 2	Sibson Seminar Room 5		
10.00	Craig Agule (Rutgers University-Camden) Being sympathetic to bad-history wrongdoers	Aleksandra Świderska (University of Warsaw) The benefit of looking harmed: Injured robots may appear more human-like and less uncanny		
11.00	COFFEE IN THE SIBSON ATRIUM			
11.30	Plenary Session 4: Lecture Theatre 2 Punishment and sympathy in times of 'bubbles'  Sabine Roeser  (TU Delft)			
12.45	5 LUNCH IN THE SIBSON ATRIUM			
13.45	Lilith Newton (University of Edinburgh) Blameworthiness and punishment for the right reason	Kirstine la Cour (University College London) Threats and Protests		
Plenary Session 5: Lecture Theatre 2  Moving between frustration and anger: Punishing, empowering,  Mary Carman (Wits University)		Punishing, empowering, but not trusting		
16.00	COFFEE IN THE	SIBSON ATRIUM		
16.30	SYMPOSIUM			
17.30	END OF CONFEREN	NCE & FAREWELLS		



## **Organising Team**

Alexandra Couto (University of Kent) A.Couto@kent.ac.uk

Alexandra Trofimov (University of Kent) A.Trofimov-692@kent.ac.uk

Lubomira Radoilska (University of Kent) L.V.Radoilska@kent.ac.uk

Lauren Ware (University of Kent) L.Ware@kent.ac.uk

### **Local Volunteers**

Alyx Robinson, Daniel Auker-Howlett, Joel Yalland, Joe Jones, James Rothwell, William Levack-Payne, Sam Holden, Evgenia Kim, Hugh Robertson-Ritchie, and Stefan Abram

Far Left: "??", Jone Bengoa, 2015. Left: Méthode pour apprendre à dessiner les passions, Charles Le Brun, 1698.

#### Plenary Session 1: In defence of vengeful anger Krista Thomason (Swarthmore College)

Vengeful anger has a particularly bad reputation. Adages and proverbs warn that desiring revenge "digs two graves" and that an "eye for an eye" leaves everyone blind. We characterize the vengeful person as too quick to anger, as taking slights too seriously, and as holding grudge: she should let go. Additionally, vengeful anger supposedly betrays morally problematic views o other people. To seek retaliation is to view the offender as a villain or enemy to be destroyed and to enjoy her suffering at the moment of revenge. Yet as often as vengeful anger is derided and lamented, we seem loath to give it up. Revenge tales in books and films are popular and satisfying Further, philosophers and legal scholars have long noted the close connection vengeful ange: bears to justice. The idea that a wrongdoer ought to pay for his crimes or "get what's coming to him" is at the heart of both the desire to see justice done and the desire for revenge. In this paper, I offer a defense of vengeful anger. In particular, I will argue that the wish for paybacl that is central to vengeful anger is neither irrational nor vicious. At the core vengeful anger is a wish or desire for what I will call forced empathy. Contrary to the claim that the payback wish is an irrational desire to turn back time or a simple desire to hurt the wrongdoer, I arguthat payback is a wish that the wrongdoer be forced to appreciate how her actions affected the victim. I then examine the appropriateness of the role of vengeful anger in punishment

Policing Unintentional Slights: Bad Sex and Moral Disapprova
Mercedes Corredor (University of Michigan)

The goal of this paper is to address how we ought to go about morally disapproving of those who unintentionally commit harms and who do so as a consequence of their existence in social contexts that set them up for moral ignorance and consequent wrongdoing. Often this debate is addressed from the perspective of those who have received a *poor* moral upbringing and concerns the extent to which these individuals might be excused on the basis of their suboptimal social conditioning. My concern here is not so much with those who have had difficult upbringings, riddled with strife and hardship, but rather with those who are beneficiaries of privilege, and who have turned out morally worse because of it. Importantly, my concern is not with privileged wrongdoers of times past, but with existing wrongdoers who are shaped by current and pervasive ideologies that are responsible for providing the shape and frame of all of our moral outlooks.

I hone in on questions of responsibility for those for whom acting morally is made more difficult by the fact that what it is to act morally is still in the process of being worked out, including by those on the frontlines who are most committed to the pursuit of achieving moral knowledge. I call these *transitional contexts* and I seek to motivate that our current practices of blaming, and of calling-out and piling-on shame in particular, might both be morally and prudentially uncalled for in such contexts. Transitional contexts, on my view, are ones in which a subset of the moral community develops knowledge about how all members of the moral community are likely to be morally fallible given that our habits have been shaped by oppressive structures which we jointly perpetuate. This knowledge needn't have an exculpatory dimension; indeed, part of what the subgroup comes to learn is that even these questions—questions of blame and responsibility—are, at least for the time being, genuinely open questions. To elucidate this, I look at cases of 'bad sex behavior', such as that described in Kristen Roupenian's 'Cat Person' case, and suggest that with this new analysis in mind, a turn towards calling-in shame, which has reconciliation as an aim, is a fitting social strategy.

## State Blame for Criminal Wrongdoing: a Skeptical View Costanza Porro (King's College London)

It is generally believed that censure is a distinctive aspect of criminal law and the many who hold expressivist and communicative theories of punishment argue that punishment is *protanto* justified precisely because it expresses censure for criminal action and blame towards criminal offenders. In this paper, I challenge the view that blame should play a central role in the criminal domain. Both those who advocate for the centrality of criminal blame and the few who question it tend to adopt a similar line of reasoning to that adopted when reflecting on moral blame, assuming that whether blame on the part of the state is justified in the criminal case largely depends on whether and under which conditions blame is the appropriate response to wrongdoing in interpersonal cases. For this reason, one criticism of criminal blame is that

## State Blame for Criminal Wrongdoing: a Skeptical View (Cont.)

#### Costanza Porro (King's College London)

moral blame, on which it supposedly depends, is exclusionary and stigmatising. I argue that the problem with criminal blame does not lie in the nature of moral blame, but in the often overlooked disanalogy between criminal and moral blame. My aim is to reflect on this disanalogy by looking at the different nature of the relationship between the state and its citizens and that between the blaming and the blamed party in interpersonal contexts.

David Shoemaker suggests that the distinction between moral and criminal blame is bes explained by appealing to the different ways in which blame operate in equal and unequal authority relationships. Instead, I argue that it is in virtue of the unique nature of th relationship between the state and its citizens that state blame is objectionable. Whether on can enjoy equal status in the community crucially depends on the action of the state; this is core difference between state blame and interpersonal blame, because individuals and the attitudes, including blame, taken singularly do not have an impact on others' status. This als differentiates other authorities from the state: no authority other than the state can have suc a global effect on the status of those whom it exercises authority over. Thus, while interperson blame is a way to treat others as moral equals and responsible agents, state blame undermine offenders' status as social equals. It does so in three ways. Firstly, state blame, in both idea and non-ideal circumstances, is an exclusionary attitude which is incompatible with the state obligation to treat offenders with equal respect and undermine their status. Secondly, it create and reinforces objectionable inequalities of esteem which undermine the status of the addresses of disesteem because they are regarded as inferior by others. Finally, if we understand blame a a moral dialogue with offenders, the traditional settings of the criminal justice system are no the right place for a fruitful exchange because they are ill-equipped for managing emotions an attitudes such as blame, shame and guilt. Restorative justice settings appear more conducive a fruitful moral dialogue and the restoration of moral relationships.

# Plenary Session 2: A hope standard for punishment Kimberley Brownlee (University of Warwick)

Each year, approximately 100 people, including dozens of young people, take their own lift while they are in a UK prison. Thousands of people either engage in self-harm or abuse other prisoners. Such behaviours often correlate with mental health struggles, but also, often, are pleas for help and hope. This paper considers what role hope should play within our practice of lawful punishment. The paper argues that one standard that a given punishment must meet in order to be legitimate and morally justifiable is that it be compatible with a reasonable personate retaining hope for her future after punishment.

#### Punitive Disappointment

#### Daniel Telech & Leora Dahan Katz (The Van Leer Jerusalem Institute, Polonsky Academy)

When blame is understood as emotion-involving, it is typically understood as involving anger, or 'angry emotions', like resentment and indignation (Strawson 1962; Wallace 1994 McKenna 2012; Watson 1987; 1996; Wolf 2011; Owens 2012; Bennett; 2002). While we do no contest that angry blame is a genuine phenomenon, we think it is a mistake to think the insofar as blame is affective, or emotion-involving, it must be angry. We argue that a type of disappointment, moral disappointment, is an emotion of blame, despite being an attitude of sadness rather than anger. Disappointment in another, on the basis of the blamee's failure t live up to a normative ideal, involves liability to sanction-like responses, such that it would b unfair to have this attitude toward agents who could not understand ideals as reason-giving Although disappointment is not a retaliatory emotion, it nonetheless tends to affect the blamer concern for the blamee's avoidance of suffering. It does so through the diminution of blamer propensity not to direct harmful responses toward the blamee. When, for example, Jill cease to confide in a colleague, Alex, who has let her down painfully, Jill may deprive him Alex of certain good, thereby setting back the interests of—and so, harming—Alex (Feinberg 1970). A a sad attitude of blame rather than an angry attitude of blame, disappointment might also b characterized by its propensity to leave targets open to suffering in indirect ways: to the exter that the blamee's disappointing action weakens the trust I invest in him, I may merely car less about the avoidance of his suffering. In either case, because of the connection betwee

n	
У	
t	
t	
l	
е	
e	
a r	
0	
1	
ıl	
S	
ıl -	
S	
S	
S	
S	
t	
d	
С	
_	
е	
r	
e	
S	
et n	
. 1	
g	
<b>ŀ</b> ;	
t	
t	
f	
f	
0	
е	
5.	
S	
S	
S	
а	
S	
е	
t	
е	
n	

#### Punitive Disappointment (Cont.)

#### Daniel Telech and Leora Dahan Katz (The Van Leer Jerusalem Institute, Polonsky Academy)

moral disappointment and liability to sanctions, conduct that is punitive can sometimes be reasonably motivated by, and expressive of, not resentment (as per Sidgwick (1874: III.V) and Feinberg (1970: 68)), but disappointment. Disappointment-based punitive responses need not be pursued for any anticipated rehabilitative, deterrent, or otherwise consequentialist effects. Disappointment-based responses may rather be retributive, as will be developed.

In Section I, we demonstrate that in the contemporary literature blame and punishment (to the extent they are associated with emotions) are associated almost exclusively with angry emotions. Section II introduces the emotion of moral disappointment, offering an analysis of its nature and fittingness conditions. Section III demonstrates that disappointment can play a role analogous to anger with respect to punitive responses. Both emotions are negatively valanced attitudes that characteristically: i) express disapproval of an agent's moral failure; ii) call for the blamee's recognition of his moral failure (and readiness to repair); and iii) involve liability to suffering that is not merely a side-effect of the interaction. Section IV then considers whether angry and disappointed responses to moral failures must be mutually exclusive, and whether they may be responsive to common features of moral failure. Section V extends the discussion by demonstrating the relevance of disappointment-expressive blame to the explanation and justification of institutional forms of punishment.

#### Agent-Regret and Punishment

#### Jake Wojtowicz (King's College London)

Sometimes we do harmful things without being at fault, like Williams's famous lorry-driver who accidentally kills a child. Such actions call for an emotional response: not guilt, for this attaches to faulty behaviour, but agent-regret, which involves lamenting what I have done whilst recognising that I was not at fault.

If guilt is an emotion of self-punishment, agent-regret is close to being one: it involves suffering, sometimes intensely, on the basis of what I have done. In this talk, I explore how agent-regret can call for various responses that have affinities with punishment, but resist the idea that we should ever punish someone who rightly feels agent-regret.

First, I explore the idea that we should punish these characters. A victim-centred approach focusses on how the victim needs to direct their resentment at someone, and the agent naturally plays that role (see MacKenzie 2017, Sussman 2018). But although the victim may appropriately feel some sort of resentment, it should not be the resentment of blame or punishment. Blame and punishment go with fault, and the agent was not at fault. We should discourage any such feelings.

The agent-centred idea arises from how several real-life incidents of agent-regret present u with agents who either relish their suffering or seem to want to suffer. Consider the woma who accidentally killed a child and is exasperated by friends who dismiss it as an accider (Gregory 2017); she did it and feels responsible for it, and her friends' kind-hearted attempt downplay this. I argue that the Good Will Hunting "It's not your fault" approach is wrongheaded Telling an agent that it was not their fault runs close to telling them they were not responsibl when, in a very real sense, they were. But for us to punish her, or for her agent-regret to tur punitive, would be just as problematic. Instead, I argue that we need to vindicate the fact she wa responsible without punishing her. We can do this by supporting her efforts to make amend encouraging her to make amends, or even by punishing her for failing to make amends; she wa not at fault in what she did, but she might be at fault in not making amends. I relate this to Joh Gardner's continuity thesis, as a particularly plausible way of understanding what she migl do to make amends. What she has done also affects how she sees herself, and how her victir sees her. One way of responding to agent-regret is to pay heed to how the woman's action sit in her identity. We must work out how both she and her victim can regard her in light of the action, without resenting her in a blaming way. I explore several non-punitive ways that w can move beyond the identities given to us by our harmful acts. I end by reflecting on how proper understanding of agent-regret, and the non-punitive responses to it, can help us achieve a realistic response to the accidental harms we foist on the world.

,	

#### Punishment, pleasure, and pain in Greek tragedy

#### Manuela Irarrázabal (University College London)

This paper explores the tension between pleasure and pain in the dramatic representation of anger and punishment in Greek tragedy from the perspective of embodied cognitive sciences. The discussion focuses on Aeschylus' Agamemnon and Euripides' Hippolytus as two instances in which the representation of punishment is pervaded with erotic vocabulary and leads to the destruction, not only of the desired object but also of the whole close social environment. The main argument I pursue here is that these tragedies show a conceptualisation of punishment that is imbedded in bodily experience that places it in a liminal position between pleasure and pain. Furthermore, as I will argue, this liminality is strongly connected to notions of anger in Greek literature.

The conceptualisations of anger punishment in terms of pleasure and pain can be found across genres in Ancient Greek literature. Plato's famous argument that it is better to suffer injustice than to commit is importantly based in the idea that punishment is not a matter of self-gratification but rather of an experience 'healing', both for the perpetrator of injustice and for the community affected by it (*Gorgias* 468b-479e). In *Rhetoric II*, Aristotle defines emotions in terms of changes of opinion in regard to judgments, accompanied by pleasure and pain (1378a). In the case of anger, the element of pleasure is linked to the perception that punishment is attainable placing the link between punishment and pleasure at the centre of the understanding of anger Greek tragedy, following ideas present in epic poetry, develops the notion of punishment as connected to pleasure and pain in ways that not only support and provide a background for Plato and Aristotle, but also provide thought- provoking details about how these concepts are connected.

The analysis of the tragedies starts out from embodied cognitive theories. By 'embodied' I mean that cognition depends upon the experience of having a body with all its sensorimotor capacities and that these sensorimotor capacities are embedded in a biological, psychological and cultural context. Our conceptual system is grounded in our bodily experience. Abstract and ontological conceptions of anger and punishment are never independent from metaphoric models, images schematic models, and symbolic models informed by bodily experience. This paper explores how the body, considered both as a lived, experiential structure and as the context for cognitive mechanisms is present in the literary representation of punishment. Desires, as well as appraisals, are subject to the norms and constraints of society, and therefore are an enriching source of information about a culture and its various literary representations. This paper aims to provide a background for other considerations of punishment, either in philosophy or drama in which pleasure and pain are involved and affect the social environment in which it takes place.

#### Guilt, Blameworthiness, and the Point of Blame

#### Andreas Brekke Carlsson (University of Oslo)

Several philosophers have recently noted that blame aims at a specific kind of emotional uptake When blaming other people we want them to painfully recognize what they have done; we want them to feel guilt or remorse. Which implications do this observation about the nature of blame have for our conception of blameworthiness?

First, I will discuss two influential accounts of blame, according to which blame is essentially communicative. Miranda Fricker (2016) argues that paradigmatic cases of blame should be seen as speech acts, whose illocutionary goal is to bring the wrongdoer to feel a pained recognition of what he has done. Blame, Fricker argues, may function as a proleptic mechanism: If the wrongdoer did not recognize that he had reasons to act differently, one might nevertheless treat him as if he did, thereby bringing him to recognize reasons he did not recognize at the time of action. I will argue that Fricker's account of blame leads to an implausible account of blameworthiness. If the agent did not recognize that he should have acted differently at the time of action he is not blameworthy, although Fricker's account entails that he is.

Coleen Macnamara (2015) takes the communicative function of blame to support a mora competence requirement on blameworthiness. Since blame is communicative and is aimed at the emotional uptake of guilt, blame would only be felicitous when directed at agents who have the capacity respond to blame with guilt. Hence, only agents with this capacity can be blameworthy. I will present a counterexample. Suppose an agent knowingly took steps to

## Wednesday, 19 June & Thursday 20 June

#### Guilt, Blameworthiness, and the Point of Blame (Cont.)

#### Andreas Brekke Carlsson (University of Oslo)

eliminate her capacity for feeling guilt in order to commit a wrong action more efficiently and with less pain afterwards. Macnamara seems committed to the view that this agent is not blameworthy for his action, because blame would no longer be communicatively apt. But this seems to be wrong result. Even though blame would no longer be felicitous in this case, the agent still seems to be *blameworthy*.

I will then propose an account of blameworthiness that avoids these problems, but nevertheless makes sense of the observation that blame aims at guilt. Drawing on earlier work, I will argue that to be blameworthy is to deserve to feel guilty. On this account, we avoid the proleptic mechanism since one does not deserve the pain of guilt unless one recognizes that one has reason to act otherwise. We also avoid the implication that people who deliberately make themselves incapable of feeling guilty are not blameworthy, since they would still deserve to feel guilt in virtue of their choice, even though blame is no longer communicatively apt. Finally we can make sense of the idea that blame aims at guilt. Blame aims at guilt because feeling guilty is what blameworthy agents deserve. Expressing blame thus serves both the communicative function of telling people that they are blameworthy, and the retributive function of causing them the specific pain they deserve: the pain of recognizing their wrongdoing.

# Plenary Session 3: Suffering and punishment Michael Brady (University of Glasgow)

In this presentation I offer a defence of the Communicative Theory of Punishment against recent criticisms due to Matt Matravers. According to the Communicative Theory, the intentional imposition of suffering by the judiciary is justified because it is intrinsic to the condemnation and censure that an offender deserves as a result of wrongdoing. Matravers raises a number of worries about this idea – grounded in his thought that suffering isn't necessary for censure, and as a consequence sometimes the imposition of suffering can be unjust. I respond by arguing that the imposition of suffering is an essential part of a suite of emotional responses that wrongdoing merits. Words are not enough for censure and condemnation.

## Being Sympathetic to Bad-History Wrongdoers

**Craig Agule** (Rutgers University-Camden)
Bad-history wrongdoers pose a philosopl

Bad-history wrongdoers pose a philosophical puzzle for theorizing about blame. When we confront their wrongdoings, bad-history wrongdoers often seem deserving of blame. However, when we confront their bad-history, the impulse to blame weakens or disappears this completely. For many philosophers, the bad-history cases are grist for theorizing about responsibility, as we wrestle with whether the true account of responsibility has a history component. In this paper, I present an alternative account of the bad-history cases. Both blame and sympathy are fitting in the bad-history cases, but these two responses conflict in the practical aims and in their demands upon our attention. Those conflicts compromise both responses, which explains our reactions to the cases without needing to complicate our accounts of moral responsibility.

In the first section of the paper, I argue that the extant analyses fail to satisfyingly explain our experience of the bad-history cases. Some philosophers argue that, because our contemporary psychologies are the causal upshots of our histories, we should not be surprised if bad-history agents' moral competence is compromised. Other philosophers think the bad-history cases reveal a contingent role history can play, such that some histories directly undermine responsibility. Finally, some incompatibilists think that what is perspicuously true of the bad-history agents is in fact true of all of us, that our behaviors are the upshots of our histories beyond our control. All of these responses to the bad-history cases urge us to see bad history as yielding compromised moral responsibility.

But the responsibility explanations are too shallow to explain our puzzled and conflicted responses to the bad-history cases. At best, those explanations show that our first impressions of the cases are mistaken, but then we should see the bad-history cases as cases of compromised responsibility, sad in the way that all cases of wrongdoing are, but not essentially complicated.

In the remainder of my paper I offer reason to think we should be uncomfortable with the badhistory cases. The bad-history cases are overlap cases, cases where more than one response is fitting. It is fitting to blame the bad-history wrongdoers, because of their wrongdoing, and

Not	

#### Being Sympathetic to Bad-History Wrongdoers (Cont.)

#### Craig Agule (Rutgers University-Camden)

it is fitting to sympathize with the bad-history wrongdoers, because of their bad history. Accordingly, we have reason to blame, and we have reason to sympathize. However, we cannot fully and simultaneously do both. Blame's characteristic motivations are in tension with sympathy's characteristic motivations, and the attention partially constitutive of blame is in tension with that partially constitutive of sympathy.

Because blame and sympathy conflict, we cannot do both. Some of our reason to act will go unsatisfied. Moreover, our reason to blame gives us reason not to fully sympathize, and our reason to sympathize gives us reason not to fully. Thus, no matter how we resolve the conflict we will act against some serious *pro tanto* reason that we have. This explains why the badhistory cases are distinctively and persistently discomfitting, and it does so without committing us to any particular account of moral responsibility.

# The benefit of looking harmed: Injured robots may appear more human-like and less uncanny

#### Aleksandra Świderska (University of Warsaw)

Will increasingly autonomous machines become regarded as moral subjects in the future? If so, might human-like robots be granted moral rights? The present work aims to empirically investigate how apparent infliction of harm may change human perception of robotic minds, and how we relate to visibly injured artificial entities. In moral psychology, a human-like mind is theorized to comprise two distinct dimensions, experience (capacities related to sensations and feelings) and agency (abilities linked to cognition). Humans tend to spontaneously attribute both types of mental states to non-human and even non-living entities. Being perceived as capable of experience and agency is closely tied to the ascription of moral status. Appearing mindless reduces moral standing and equates an entity with a mere object deprived of all emotional experiences. Further, perception of a human-like mind elicits the ascription of moral rights and responsibilities. A moral patient is generally seen as capable of experience, especially of suffering, and deserves to be protected from harm. A moral agent, in turn, is perceived to possess the means to inflict harm. Together, moral patients and agents constitute a dyad, in which moral interactions unfold. This dyad is often understood as a foundation for the agent to be entitled to punishment for malevolent behavior.

As recently discussed, an examination of moral patiency may provide a revolutionar alternative to the moral philosophy of machine ethics, by re-examining how we think abou the moral status of other entities from the perspective of a relational turn (e.g., Gunkel, 2018 Thus, when we ask who can suffer, and who can be blamed, we should consider such processe of basic human cognition before proceeding to examine their role in practices of punishment More specifically, the question arises in earnest, how humans perceive harm to robots, and how observing injury and (apparent) pain may shape what/who we perceive as having a mino as human-like, or alien and uncanny. Currently, we conducted a study to investigate how facial wounds shaped attributions of mind to male and female human-like and robotic avatar (2 x 2 x 2 between-subjects factorial design). The results revealed that harmed entities wer perceived as more capable of experience than their unharmed counterparts. This effect, terme in prior research the harm-made mind, may emerge from the entity's assumed participation in a moral interaction, whereby it is automatically cast in the role of a moral patient and imbue with corresponding mental capacities. Surprisingly, harmed avatars were evaluated as mor affiliative and less uncanny, which may indicate that visible harm rendered the entities mor friendly and socially-oriented. In contrast to previous findings, this suggests that increases in perceived experience may be associated with reduced uncanniness of anthropomorphic robots at least if the type of wound is similar to the way a human would be injured. We will discuss th implications of our results from a relational standpoint, and outline how future work could she more light on how humans construe the minds and suffering of human-like machines.

## **Plenary Session 4:** Punishment and sympathy in times of 'bubbles' **Sabine Roeser** (TU Delft)

Public debates about controversial technological and scientific developments such as climate change, vaccination and genetic modification are frequently heated and end up in stalemates. This is due to the scientific and moral complexities of these risks, which lead to strong emotional responses by people. The way emotions are typically treated in debates increase estrangement and polarization. People from different 'bubbles' blame each other for seeing the world in an irrational and lopsided way. In my presentation, I will argue that this is due to a misunderstanding of emotions. Rather than seeing emotions as irrational states, I argue that emotions can contribute to emotional-moral reflection and public deliberation on technological risks. I will argue that emotions are crucial to debates about technological risks, because emotions can point out what morally matters. Conventional, quantitative approaches leave out important ethical considerations such as justice, fairness, autonomy and legitimacy. Addressing emotions in a different way can help to overcome stalemates: by focusing on what unites us rather than what divides us, emotions can contribute to sympathy and understanding of shared values, which can in turn contribute to finding commonly shared solutions. However, these insights are challenged by developments on social media, where polarization and venting of anger, blame and extreme viewpoints seem to dominate, as compared to understanding, forgiveness and finding common ground, so different platforms for deliberation are needed.

## Blameworthiness and Punishment for the Right Reason Lilith Newton (University of Edinburgh)

In a recent paper, Clayton Littlejohn has argued that it is permissible to convict, so sanction the punishment of, defendants in criminal trials only if the fact-finder—the jury in a trial by jury, or the judge in a bench trial—knows that the defendant is guilty. This is because punishment is permissible only if it serves a particular expressive function: to "properly express blame ... for some specific deed" (2018: 17). Punishment can serve this function only if it is imposed for the right reason: "the fact that the defendant is guilty of the offense" (23). But the fact-finder can impose punishment for this reason only if she knows it to be the case: facts "can be our reasons for thinking things, feeling things, and doing things ... if [beliefs in the facts] constitute knowledge" (28).

I argue that Littlejohn's argument for his knowledge constraint on the permissibility of criminal conviction and punishment should be rejected, for conviction and punishment are intuitively permissible in cases where expressing blame would be inappropriate. Following Fischer and Ravizza (1998), I endorse two necessary conditions for blameworthiness. Firstly, for an agent to be blameworthy for  $\phi$ -ing,  $\phi$ -ing must be morally wrong. Secondly, it must be the case that the agent could have done other than  $\phi$ . I argue that neither condition need be met in order for conviction and punishment to be permissible. For many crimes *mala prohibita*, for example unlawful possession of a controlled drug, a fact-finder may be right to convict a defendant who has not acted morally wrongly (in using drugs with no intent to sell), and may not even have been able to do otherwise, in some relevant sense (for example, if she is addicted to drugs). Littlejohn's appeal to blameworthiness thus cannot ground his knowledge constraint on permissible conviction and punishment.

However, I argue that Littlejohn's notion of convicting and punishing 'for the right reason' is something that the law should care about. If we conceive of the 'right reason' for conviction and ensuing punishment to be that the criminal standard of proof has been met—the defendant's guilt has been proved beyond reasonable doubt—then the notion of convicting and punishing 'for the right reason' comes apart from questions of blameworthiness. The fact-finder can convict, so sanction the punishment of, a defendant who cannot be appropriately blamed for any bad deed. Then something valuable can be salvaged from Littlejohn's account of permissible conviction and punishment. I will argue, however, that what we are left with is insufficient to ground a knowledge constraint on permissible conviction and punishment. The strongest constraint Littlejohn is entitled to is one on which conviction and punishment are permissible only if the fact-finder knows that the criminal standard of proof has been satisfied; she need not know that the fact-finder is guilty.

#### Threats and Protests

#### Kirstine la Cour (University College London)

Consider the following familiar sequence of phenomena from moral life: 1) Someone wrongs me; 2) I blame them; 3) They apologise; 4) I forgive. In "Articulating an Uncompromising Forgiveness" (2001), Hieronymi proposes that we conceive of these phenomena and the relations between them as an interplay between threat and protest. In short, by wronging me, you make a threatening claim against me, and by blaming you, I protest and reject that claim. By apologising, you diffuse your threat, thereby removing my need to continue protesting, rendering me instead ready to forgive.

Hieronymi's proposal is interesting for several reasons. While all the phenomena mentioned above have received considerable philosophical consideration, less attention has been paid to articulating the relationships between them. By offering an account of the whole chain, the threat-and-protest proposal enables us to attend to and explain a number of significant features of our accountability practices, which, I hold, remain under-explored or insufficiently well-understood in the philosophical literature. Amongst these, I focus on the following two: the distinctively interpersonal, other-directed character of these practices, and their temporal duration. Both features appear productively illuminated by likening the string of phenomena to the moves in a conversational exchange. The account seems both to explain how blame can be a form of "moral address", itself inviting a response from its target, and to make sense of the possibility and propriety of continued blame and its eventual cessation.

My objective in the paper is to argue that this appearance is misleading, and that the proposal in fact accounts satisfactorily for neither of these. Specifically, I show that if, as Hieronymi proposes, blame is understood as a protest against and repudiation of a wrongdoer's threatening claim, apology cannot have the function of diffusing the threat, thereby paving the way for forgiveness, for what the protest signifies is precisely that the threat cannot be sustained. Unless the victim's protest is somehow undermined or otherwise unsuccessful—something I argue we should not accept—Hieronymi's own construal of these terms should not allow for the possibility of the coexistence of threat and protest. The upshot of this is that the account fails to explain both the continuation of blame and its cessation. Moreover, while the account instructively captures certain important aspects of the communicative, other-directed character of our practices, it is insufficiently discriminating about who gets to do the talking and how. I end by suggesting some routes for amelioration.

# **Plenary Session 5:** Moving between frustration and anger: Punishing, empowering, but not trusting

#### Mary Carman (Wits University)

Anger has been widely defended as an emotion that can have political and moral value. While it may be a punishing emotion - in our own experiences and for those we are angry with - there are good arguments to think that it can nevertheless be an empowering one. Yet, many cases of political anger arise out of frustration, particularly frustration in response to a lack of change. That frustration can move into anger after some catalysing event, then revert to frustration. Take, for instance, the recurring student protests in South Africa since 2015 against the lack of access to tertiary education in the country. This can be seen as frustration at a perceived lack of political will to bring about meaningful change in pursuit of justice, but which flares into anger after a catalysing event, such as the financial exclusion of returning students from impoverished families in 2019. After emotion-fuelled protests and still no substantial change, the anger simmers down but frustration remains. Frustration, however, is not open to the same kinds of philosophical defences that are available for anger: where anger can be punishing but empowering, frustration is plausibly only punishing. Can we learn anything by highlighting the role of frustration in many cases of political anger? In this paper, I offer a philosophical analysis of frustration as a political emotion and argue that the movement between frustration and anger (and back) draws out the centrality of a break-down of trusting that can get lost if one only focuses on anger. Reframing the frustration-anger movement within a context of trusting potentially has implications for how we respond to—indeed, if we punish—the associated anger and expressions of that anger.

## Map of Campus & Venues

# FINDING YOUR WAY AROUND THE CAMPUS Colleges Other buildings Other buildings Wain road Minor road Minor road Wain entrance, suitable for disabled access Main entrance Main entrance Main entrance Cycle path National cycle route 1 Cycle lock up Cycle lock up, sheltered Cycle lock up, sheltered





J14 Locke Building

J15 Eliot College Extension

Jennison Building

Ingram Building

E5 Darwin College

Cornwallis East



to Canterbury

**Sibson Building,** University of Kent, Park Wood Rd, Canterbury CT2 7PE