

## Punishment and Transformation

Jennifer Lackey  
Northwestern University

In this paper, I argue that strict, long-term punishments are epistemically irrational. By appealing to the radical changes in mental states that can be brought about through transformations or transformative experiences, and showing that punishment needs to be sensitive to such mental states, I argue that strict, long-term punishments screen off the possibility of being sensitive to epistemic information that is highly relevant. I conclude that rationality demands that justice be an ongoing process, open to revision in light of changes in both those being punished and those doing the punishing.

### 1. Preliminary Remarks

It will be helpful to begin with some clarifying remarks about the terminology I will be using, as well as the general framework for the issues to be discussed. By “strict,” I mean punishments that are not open to revision. Prison sentences without the possibility of parole are classic examples of strict punishments as I am understanding them here. Moreover, I am distinguishing between a punishment being *revised* and one being *overturned*. A prison sentence without the possibility of parole cannot be reduced or otherwise modified, but it can be overturned if, for instance, exculpatory evidence is discovered. “Long-term” will obviously be a somewhat loose notion, and may even depend in part on the nature of the action being punished. Losing the car for a year might be a long-term punishment for a teenager violating curfew, for instance, while this same amount of time might seem short-term for a violent offense. My purpose in this paper is not to wade into questions of this sort. Instead, I will focus on clear, paradigmatic cases of long-term punishments and argue that they are epistemically irrational. To the extent that the premises on which my arguments rely are true of other punishments, they can be applied accordingly.

I will often use natural life sentences as a central case of a strict, long-term punishment. A sentence of “natural life” means that the remainder of one’s natural life will be spent behind bars. There are no parole hearings, no credit for time served, no possibility of release. Short of a successful appeal or an executive pardon, such a sentence means that the convicted will, in no uncertain terms, die behind bars.

While natural life sentences will be the central case of strict, long-term punishments, they are by no means the only one. *De facto* life sentences, which are sentences that exceed the life expectancy of the convicted, are also clear examples. A 25 year-old man who is given a 60-year sentence, for instance, will most likely end up with the same outcome as a 25-year-old given a natural life sentence: death while in prison. Indeed, this is even clearer when we factor in the toll that incarceration takes on mortality rates. While the average life expectancy of an American is 78.8 years,<sup>1</sup> it is 64 years for those who are incarcerated.<sup>2</sup> Thus, even many significantly shorter sentences turn out to be *de facto* life ones once this is taken into account.

Moreover, truth-in-sentencing policies, which “require those convicted and sentenced to prison to serve at least 85 percent of their court-imposed sentence,”<sup>3</sup> turn many lengthy prison sentences into strict ones. For instance, the state of Illinois adopted its version of truth-in-sentencing, which requires, among other things, that those convicted of murder serve 100 percent of their sentences.<sup>4</sup> Given this, every sentence for murder in Illinois is a strict one, with no parole hearings and thus no possibility of revision.

The scope of this paper will be on the question of the rationality of given punishments, even if they have been appropriately applied at the outset. There are a variety of issues that might make us rethink our punishment policies. For example, in recent years there have been many cases of people being exonerated after it has been discovered they were wrongfully convicted. There has also been an increasing appreciation for the ways in which structural racism and economic disadvantages affect

how punishments are distributed. These are very serious concerns, but I will set them aside in this paper to focus instead on another issue: the ongoing rationality, and therewith justice, of strict, long-term punishments, regardless of their origin.

Finally, all of my arguments will be epistemic in nature, with the conclusion that strict, long-term punishments are *epistemically irrational*. By “epistemic,” I mean of, or related to, knowledge and knowledge-related goals. For instance, it might be rational from a moral point of view to believe in your son’s innocence, despite the overwhelming evidence on behalf of his guilt, because of the moral duties parents have to their children. It might also be practically rational to believe that you’re going to survive your terminal diagnosis in the face of massive amounts of evidence to the contrary because this belief will significantly improve the quality of your remaining days. But believing against the evidence is clearly at odds with achieving knowledge-related goals—such as truth, justification, understanding, and so on—and thus in neither case would the agent involved be epistemically rational.

## 2. The Transformation Argument

With these points in mind, let’s turn to my first argument on behalf of the conclusion that strict, long-term punishments, such as natural life sentences, are epistemically irrational. I will call this the *Transformation Argument* (TA).

The first premise of the TA is:

1. Punishment at a given time ought to be sensitive to the relevant evidence available at that time.

One way of supporting (1) is by appealing to a widely accepted view in epistemology, according to which action is epistemically appropriate only when it is grounded in a sufficiently good epistemic position. A central candidate for this position is knowledge. For instance, Timothy Williamson

maintains that the “epistemic standard of appropriateness” for practical reasoning can be stated as follows: “One knows  $q$  iff  $q$  is an appropriate premise for one’s practical reasoning” (Williamson 2005, p. 231).<sup>5</sup> Similarly, according to John Hawthorne and Jason Stanley, “Where one’s choice is  $p$ -dependent, it is appropriate to treat the proposition that  $p$  as a reason for acting iff you know that  $p$ ” (Hawthorne and Stanley 2008, p. 578).<sup>6</sup> We can call the thesis found in these passages the *Knowledge Norm of Practical Reasoning*, or the KNPR, and formulate it as follows:

KNPR: It is epistemically appropriate for one to use the proposition that  $p$  in practical reasoning if and only if one knows that  $p$ .

As stated, there are two dimensions to the KNPR; one is a necessity claim and the other is a sufficiency claim. More precisely:

KNPR-N: It is epistemically appropriate for one to use the proposition that  $p$  in practical reasoning only if one knows that  $p$ .

KNPR-S: It is epistemically appropriate for one to use the proposition that  $p$  in practical reasoning if one knows that  $p$ .

While both versions of the Knowledge Norm of Practical Reasoning are of interest, my focus here will be on only the necessity claim.

Actions are typically the result of practical reasoning, and those actions that are epistemically proper must be the product of practical reasoning that itself meets a sufficient epistemic standard.<sup>7</sup> If, for instance, the KNPR-N is correct and this standard is knowledge, then acting on  $p$  requires that one know that  $p$ .

In place of KNPR generally and KNPR-N more specifically, others have argued for weaker constraints on practical rationality and action. Ram Neta, for instance proposes the following:

JBK-Reasons Principle: Where S's choice is  $p$ -dependent, it is rationally permissible for S to treat the proposition that  $p$  as a reason for acting if and only if S justifiably believes that she knows that  $p$ . (Neta 2009, p. 686)

On Neta's view, then, epistemically proper action requires, not an appropriate connection with knowledge, but, rather, with justified belief that one knows. In particular, in order for it to be epistemically proper to act on  $p$ , one must justifiably believe that one knows that  $p$ . Others have defended even weaker views, especially with respect to particular kinds of action,<sup>8</sup> such as a *justified belief* norm,<sup>9</sup> a *rational credibility* norm,<sup>10</sup> a *reasonable-to-believe* norm,<sup>11</sup> and a *supportive reasons* norm.<sup>12</sup>

It is not my aim here to defend one of these views over another. Instead, my point is to make clear that it is widely accepted that there *is* an epistemic standard governing action, regardless of the strength of the norm. Support for this more general point comes from our practices of praise and blame. If, for instance, you fail to pick up my daughter from school today at 3:35 PM, despite the fact that you do so every weekday, an appropriate defense would be, "I knew, or I had good reason to believe, that all students needed to stay until 4:00 PM today because this is what the school's website indicated." Your action here would be justified by appealing to the epistemic status of the beliefs guiding it and, assuming that they meet the correct standard, you wouldn't be subject to criticism for failing to pick up my daughter.

This brings us back to premise (1). Punishment is an action, and so in order for it to be epistemically proper, it must meet an epistemic standard. Regardless of whether this standard is knowledge or something weaker, none permit ignoring relevant evidence. Indeed, even if knowledge is understood in a highly externalist way, such that it requires something along the lines of reliability or truth-tracking, evidence cannot be disregarded. Counterevidence, for instance, can always function as a defeater, either a rebutting or an undercutting one.<sup>13</sup> Even if, say, I reliably form the belief that there is a fox in my backyard on the basis of perception, evidence that my belief is false

(rebutting), or unreliably formed or sustained (undercutting), might still defeat my justification. If you tell me that foxes have never been seen in your area, and that your neighbor has a Shiba Inu who frequently escapes into your yard, then I have evidence against the truth of my fox belief. Similarly, if my optometrist tells me that I'm wearing glasses with a wildly incorrect prescription, then I have evidence against the reliability of the basis of my fox belief. Even if we are not all evidentialists, according to which evidence is the central source of epistemic goods, it is clear that there is no way of circumventing the crucial role that evidence plays in our belief-forming practices. As David Hume famously said, "a wise man proportions his belief to the evidence."

Moreover, relevant evidence includes not only what is possessed at a given time, but also what a subject *should have* at that time. Suppose, for instance, that the total body of evidence that a racist possesses justifies his white supremacist beliefs, but only because he deliberately avoids the acquisition of any evidence to the contrary. Perhaps he carefully chooses to be surrounded by only other racists, and he reads only news sources that defend his beliefs. Even though the evidence in his possession supports his racist beliefs, it doesn't follow from this that they are justified. Why not? Because there is evidence that he ought to possess, or ought to consider. Such evidence is often characterized in terms of *normative defeat*, in contrast to *doxastic defeat*, which involves counterevidence that is already possessed.<sup>14</sup> But the central point is that "relevant evidence" in (1) is not limited to only the evidence that a subject possesses. This is important, especially in matters of criminal justice. It clearly won't do for a police officer to fail to follow through with highly relevant and credible leads, only to fall back on the view that his belief in the defendant's guilt is justified relative to the wildly incomplete evidence he in fact possesses. Similarly, it won't do for the State to ignore who prisoners have become after serving decades in prison, only to claim that the very lengthy sentences are justified relative to the evidence that was had at the time of their convictions.

Support for (1) is thus found in the combination of the following three claims: (i) action is governed by an epistemic norm; (ii) any version of this norm will include room for relevant evidence; and (iii) punishment is an action. Otherwise put, punishment, being an action, should be grounded in epistemically proper beliefs of the one doing the punishing, such as the State, and epistemically proper beliefs cannot be insensitive to relevant evidence. Thus, punishment at a given time ought to be sensitive to the relevant evidence available at that time.

Such a view is certainly borne out in many of our practices. Perhaps the most striking is found in exonerations, many of which are made possible by evidence uncovered even decades after the original convictions. Since 1989, for instance, there have been 344 people in the United States exonerated by DNA testing.<sup>15</sup> In many of these cases, it is only because of the technological evolution of forensic DNA profiling that evidence became available to prove the innocence of the defendants. Were future evidence screened off so that punishment was based only on the evidence available at the time of the original convictions, the innocence of these hundreds of men and women would never have been established. Similar considerations apply to other sources of exoneration—a key eyewitness recants, testimony becomes available that the defendant was tortured and issued a false confession as a result, theories once accepted by the scientific community become widely challenged, such as evidence of arson or Shaken Baby Syndrome,<sup>16</sup> and so on. In each of these cases, our practices in the criminal justice system clearly support (1). Indeed, were this not to be the case—were our criminal justice practices to be at odds with (1)—epistemic and moral errors of massive proportions would be rampant.

The second premise of the Transformation Argument is:

2. Punishment ought to be sensitive to the mental states of the one being punished, where this includes his or her mental states after the time of the punishable act.<sup>17</sup>

By “mental states” here, I’m including beliefs, desires, emotions, intentions, and dispositions. Mental

states that are paradigmatically relevant are, for instance, whether the punishee appreciates the wrongness of his or her actions, feels remorse, intends to avoid wrongdoing in the future, and so on. Moreover, notice how weak this premise is. It does not say that punishment ought to be sensitive to *only* the mental states of the punishee. Thus, many other factors might bear on punishment, such as retribution, deterrence, and the mental states of the victims, if there are any. It also does not require that punishment be *highly* sensitive to the punishee's mental states. In this way, punishment may even be largely determined by factors besides the beliefs, desires, emotions, intentions, and dispositions of the one being punished. Perhaps the primary aim of punishment is to serve as a deterrent, and the mental states of the punishee are relevant only insofar as they bear on or supplement this. But what this premise does make clear is that the mental states of the person being punished cannot be screened off when considering the legitimacy of a given punishment.

Premises (1) and (2) are related, yet they are nonetheless importantly different. (1) makes the general point that punishment needs to be sensitive to the relevant evidence available at the time in question, while (2) maintains that the mental states of the punishee, even those available after the time of the punishable act, are included in the body of relevant evidence. Thus, (1) could be true while (2) is false if one accepts that punishment needs to be sensitive to relevant evidence, but denies that the mental states at issue are relevant in this sense. Moreover, I include both premises in the argument since the justification for each is different, thereby providing a more comprehensive picture of the epistemic dimensions of punishment.

Support for premise (2) can be found in many of our practices. On behalf of the general thesis that mental states are relevant to punishment, the most obvious example is that the mental states of a person at the time a punishable act is committed are often taken to be of great significance. A clear example here is the role that *mens rea*, or a "guilty mind," plays in the criminal justice system. For instance, a person who engages in illegal activity because he or she honestly



misperceives reality—that is, he or she makes a “mistake of fact”—lacks *mens rea* and is said to be such that he or she ought not be charged with, or convicted of, a crime. If, say, I reasonably but mistakenly believe that your MacBook Air is mine and I leave our department meeting with it, the absence of a guilty mind means that I shouldn’t be charged with, or convicted of, theft. Similar considerations apply to the difference between first-degree murder and lesser charges, such as second-degree murder or manslaughter: only the former requires premeditation which, in turn, involves mental states, such as the intention to kill, beliefs about how to bring this about, and so on. Indeed, some convictions, such as those involving conspiracy charges, are grounded entirely in the mental states of the accused. Clearly, then, the very nature of the criminal charges brought against people depends heavily on their mental states, thereby directly affecting the punishment, or lack thereof, handed out to them.

But the crucial part of premise (2) for my purposes is the role played by mental states after a punishable act, and here it should be noted that we often regard the mental states of a punishee long after an offense has been committed to be relevant to the punishment deserved. In explaining given sentences, for instance, judges often appeal to the mental states of the convicted at the time of the trial or the sentencing, even if this is long after the crime in question was committed. This is clear in the recent highly controversial case of Brock Turner, the student from Stanford who was convicted of three felony counts of sexual assault and sentenced to six months in county jail, three years of probation, and a requirement that he register as a sex offender. In defending what many regarded as an egregiously light sentence, the judge “said that he believed the defendant felt genuine remorse.”<sup>18</sup> Now it is clear that the judge is not speaking about Turner’s mental states at the time of the sexual assault, nor even immediately afterward, but, rather, at the time of the trial and sentencing, which occurred well after a year of the crimes themselves. Moreover, even though there was a public outcry about the punishment Turner received, many saying that it was disproportionate to the

seriousness of the crimes, the criticism did not focus on whether Turner's remorse, or lack thereof, ought to have made a difference. Indeed, when Turner himself seemed to blame what occurred on a "party culture" of "drinking," the fact that he seemed to fail to appreciate his own agency in the sexual assault that he perpetrated fueled calls for a more stringent sentence.<sup>19</sup> Thus, even though the judge and the public are at opposite ends of the spectrum on the sentence involved in this particular case, they are crucially united in appealing to Turner's mental states at the time of the trial and sentencing to at least partly justify their competing positions.

Of course, there is no magic number after which an act is committed that we would say the mental states of a punishee fail to matter to the punishment deserved. Some trials take place decades after a crime is committed, and it is not unusual for judges, prosecutors, and defense attorneys to point to who the defendant currently is to defend a particular sentence. In 2014, for instance, a psychology professor, Norma Patricia Esparza, was offered a 6-year prison sentence in exchange for pleading guilty to voluntary manslaughter for her purported role in the 1995 murder of a man who she said raped her. This deal was made nearly 20 years after the crime was committed, and the district attorney's office made clear that Esparza's mental states were crucial to the offering of the plea, saying that it reflected Esparza's "acknowledgement and acceptance of her role in the victim's murder."<sup>20</sup> Once again, acknowledgement and acceptance crucially involve mental states, such as beliefs, and the district attorney is talking about Esparza's states now, not 20 years earlier.

This point is not supported merely anecdotally, as "[m]any state courts have found remorse to be an appropriate mitigating factor to consider when assigning criminal punishment" and "have found the absence of remorse to be an appropriate aggravating factor when calculating an appropriate criminal punishment," (Ward 2016, p. 131).<sup>21</sup> Moreover, it is not only judges who regard remorse as a relevant factor for sentencing, but jurors as well, and it is often the behavior and demeanor of the defendant at the trial that most influence jurors' assessments of remorse.<sup>22</sup> Given

that trials often take place years after the criminal behavior in question purportedly took place, it follows straightforwardly that current practices in our criminal justice system strongly support premise (2).

Perhaps the clearest and most powerful example on behalf of premise (2) is the evidence taken to be relevant at parole hearings or resentencing hearings. Some of the key conditions for release or reduced sentences include conceding guilt,<sup>23</sup> good conduct in prison, completing classes, having been sufficiently rehabilitated, and not posing a danger to society. All of these involve mental states to varying degrees. It is, for instance, most natural to understand someone's being rehabilitated and not posing risks to society in terms of changes in their beliefs, desires, intentions, emotions, and dispositions. For instance, gang members once thought to be friends might now seem to be exploitative, bouts of anger might be channeled into education and advocacy rather than revenge, and the intention to be a role model for one's children might replace trying to impress one's peers.

We also see the relevance of current mental states in other punitive contexts. Compare two students known to have cheated: the first fully acknowledges that looking at her notes during an exam was wrong, is clearly contrite, and promises to never do so again, while the second flagrantly and steadfastly lies about it and shows no evidence that he won't cheat again. It is fairly standard for the second student's punishment to be harsher than the first's.

The third premise of the Transformation Argument is:

3. People can change, often in profoundly transformative ways, which can involve radical changes in their corresponding mental states.

Such transformations can be clearly seen by considering the two ends of the spectrum of life. On the early side, it is now widely known that the prefrontal cortex of the brains of adolescents and emerging adults is still developing, leading to their being more likely than adults to act on impulse,

engage in dangerous or risky behavior, and misread social cues and emotions.<sup>24</sup> Indeed, “the frontal lobes, home to key components of the neural circuitry underlying ‘executive functions’ such as planning, working memory, and impulse control, are among the last areas of the brain to mature; they may not be fully developed until halfway through the third decade of life” (Johnson, Blum, and Giedd 2009, p. 216). This fact by itself raises a host of questions about the level of responsibility that adolescents and emerging adults bear for their actions and the appropriate punishments that should be handed out to them. If, for instance, the underdeveloped brains of adolescents at least partly explain criminal behavior—behavior that wouldn’t have occurred had they been adults—then holding them fully responsible for their actions, and punishing them as adults, seems wildly off the mark. But the point that I wish to emphasize here is that normal brain development often results in profound changes between adolescence and adulthood, ones that make the beliefs, decisions, and actions of the other seem foreign and perplexing. This is perhaps why adults often look back on some of the actions of their younger selves with embarrassment and even horror, and why teenagers frequently feel disconnected from, and poorly understood by, the adults around them.

On the later side, attention has been drawn to the fact that people tend to age out of crime.

John F. Pfaff makes this point when he writes:

Let’s start with the flaws in how we incapacitate violent offenders. Even that sentence, which seems so banal, is actually quite misleading. “Violent offenders.” We use this term a lot, but we shouldn’t... For almost all people who commit violent crimes... violence is not a defining trait but a transitory state that they age out of. They are not violent people; they are simply going through a violent phase. Locking them up and throwing away the key ignores the fact that someone who acts violently when he’s eighteen years old may very well be substantially calmer by the time he’s thirty-five. (Pfaff 2017, pp. 190-1.)

Moreover, only 1% of serious crime is committed by people over the age of 60. According to

Jonathan Turley, “Everyone agrees on what is the most reliable predictor of recidivism: age. As people get older, they statistically become less dangerous.”<sup>25</sup> Turley refers to this period as “criminal menopause,” a phenomenon where people lack the desire, and often the means, to engage in criminal activity. This raises serious questions about the rationale for punishment involving the elderly. If, for instance, there is abundant evidence that aging prisoners have been rehabilitated, and a negligible chance that they would pose a safety risk upon being released, then the skyrocketing numbers of elderly prisoners in the United States seem to cry out for explanation.<sup>26</sup> Again, though, the central point I wish to make here is the way in which this fact supports premise (3). For if even people who repeatedly engaged in criminal activity as young adults completely eschew this as they age, then there is a clear sense in which they have changed in transformative ways.

At the early end of the spectrum of life, then, there is the possibility that people *might change*; at the later end, there is the reality that they *have changed*. Both facts support premise (3).<sup>27</sup>

Moreover, there is a further kind of transformation that is importantly relevant here. In her recent book, L. A. Paul focuses on a phenomenon that she calls *transformative experience*, which involves experiences that are both *epistemically* and *personally* transformative. To be clear, I am distinguishing ordinary *transformations*, such as the typical changes that happen between adolescence and adulthood, from *transformative experiences*, which pick out the phenomena specifically described by Paul. An *epistemically* transformative experience is one that provides new information that could not have been learned without having that kind of experience. A paradigmatic example here is that of Mary from Frank Jackson’s well-known knowledge argument,<sup>28</sup> who is a scientist specializing in the neurophysiology of vision. She acquires all of the physical information there is about color and related matters, but she herself has spent her entire life in a black and white room. Now imagine what happens when Mary leaves her room for the first time and see a red apple. It seems plausible that, despite possessing knowledge of all of the physical information there is about color, she still

learns something new; namely, what it is like to see color, and red in particular. Jackson uses this to conclude that physicalism is false, but what is relevant for our purposes is that Mary has an epistemically transformative experience. She learns what it is like to see red, which is something she couldn't have learned without an experience of this sort. While this is an extraordinary case, there are also plenty of ordinary examples of epistemically transformative experiences, such as what I would come to know were I to taste a durian fruit for the first time.<sup>29</sup>

An experience is *personally* transformative if it “changes you enough to substantially change your point of view, thus substantially revising your core preferences or revising how you experience being yourself” (Paul 2014, p. 16). In an important sense, personally transformative experiences change who you are by radically altering your point of view. “Such experiences may include experiencing a horrific physical attack, gaining a new sensory ability, having a traumatic accident, undergoing major surgery, winning an Olympic gold medal, participating in a revolution, having a religious conversion, having a child, experiencing the death of a parent, making a major scientific discovery, or experiencing the death of a child” (Paul 2014, p. 16). While personally transformative experiences are different than those that are merely epistemically transformative, this does not mean that there is not a central epistemic dimension to these. As Paul says, “...if a personally transformative experience is a radically new experience for you, it means that important features of your future self, the self that results from the personal transformation, are epistemically inaccessible to your current, inexperienced self” (Paul 2014, p. 17).

Transformative experiences simpliciter, then, are those that are both epistemically and personally transformative in these senses. In addition to the examples Paul provides above, one experience that is frequently transformative for people and is particularly relevant for our purposes here is incarceration. For instance, in a very recent CNN article entitled “Ex-con transforms to entrepreneur behind bars,” the author, Coss Marte, writes, “My personal transformation happened

behind bars. I was sent to ‘the box’ after an altercation with a prison officer. After I was beaten, I was shoved into the cell and forced to do nothing but think. I asked myself ‘Why?’ How did I end up here?’<sup>30</sup> Marte goes on to describe how his beliefs, desires, and actions radically changed after this moment. He began exercising, losing 70 pounds in 6 months, he turned to God, he realized that selling drugs was wrong, and he “began to believe that [his] purpose was to give back instead of destroying individuals around [him].” This is not uncommon. Indeed, a recent article on “prison conversions” begins as follows: “The jail cell conversion from ‘sinner to saint’ or from nonbeliever to true believer is a well-known, indeed almost clichéd character arch [sic] in feel-good fiction, history, and media accounts” (Maruna, Wilson, and Curran 2006, p. 161). The authors explain these conversions in terms of what psychologists William R. Miller and Janet C’deBaca call “quantum changes,”<sup>31</sup> which are sudden identity transformations that are “qualitatively different from the more common, incremental changes in human development” (Maruna, Wilson, and Curran 2006, p. 161). Such quantum changes clearly involve radically new beliefs, such as those described by Marte above.

We are now in a position to see the Transformation Argument unfold as follows:

1. Punishment at a given time ought to be sensitive to the relevant evidence available at that time.
2. Punishment ought to be sensitive to the mental states of the one being punished, where this includes his or her mental states after the time of the punishable act.
3. People can change, often in profoundly transformative ways, which can involve radical changes in their corresponding mental states.
4. Strict, long-term punishments screen off any relevant future evidence, including the radically different mental states of people who have changed in transformative ways.
5. Therefore, strict, long-term punishments are epistemically irrational.

I have already offered defenses of (1)-(3), so let me say a few words about (4) and (5). Recall that

strict punishments are those that are closed to revision, and thus it clearly follows that they screen off any non-exculpatory evidence beyond what is available at the time the punishment is handed out. Focusing on our paradigmatic case, natural life sentences say to all involved that there is no possible piece of information that could be learned between sentencing and death that could bear in any way on the punishment the convicted is said to deserve, short of what might ground an appeal. Nothing. So no matter how much a juvenile is transformed behind bars, and no matter how unrecognizable an elderly prisoner is from his earlier self, this is utterly irrelevant to whether they should be incarcerated. Our absence of knowledge about the future, our ignorance of what is to come, our lack of a crystal ball, is in no way a barrier to determining now what someone's life ought to be like decades from now.

While natural life sentences are the clearest example here, the same considerations hold for all strict punishments, such as any versions of truth-in-sentencing that require those convicted of crimes to serve 100 percent of their sentences.

However, screening off the possibility of even considering future, non-exculpatory evidence in relation to punishment is irrational. For as we have seen, future evidence, such as the radically different mental states of the punishee, can be relevant to the punishment that is deserved. This is especially clear when we consider all of the support on behalf of (3) showing that transformations while incarcerated are not only possible but likely. Indeed, consider this: if we take two defendants with different mental states regarding their crimes as deserving of different punishments at the time of sentencing, why would we not regard two stages of the same person — one at 19 and another at 49 — with radically different attitudes toward his crime, as deserving of different punishments?<sup>32</sup> Current selves and future selves can vary from one another no less than two altogether distinct people do.<sup>33</sup>

There is a related worry here that is worth mentioning briefly that appeals to the plausible



moral principle, due to Aristotle, that *likes ought to be treated alike*.<sup>34</sup> If, for instance, two students produce work identical in quality, justice and fairness seems to require that I assess them comparably, just as I should sanction them similarly if these same two students are then found to have cheated under the same conditions. To fail to do so would be to violate this principle by not treating like cases alike, a violation at the heart of much discrimination that is deeply problematic. But now notice: there is no reason why moral principles of this sort should apply only once rather than in an ongoing way. Otherwise put, *we ought to treat likes alike across time*. Suppose, for instance, that A received a 40-year sentence while B received a 20-year sentence for the same crime. It may be perfectly appropriate to punish A more harshly than B at T1 because, though they committed the same crime, A's mental states crucially differ from B's at T1. Perhaps A takes delight in the memory of the crime or vows to commit more such crimes in the future, while B does not. At T2, however, suppose that A transforms or undergoes a transformative experience that renders his current mental states relevantly similar to the way B's were at T1. If A and B continue to be punished differently by serving sentences that vary significantly in length, then likes are not being treated alike.

But the problem that I want to raise here is an epistemic one: strict, long-term punishments *screen off the possibility of treating likes alike*. In particular, given premise (2) of the Transformation Argument, treating likes alike in relation to punishment requires sensitivity to the punishee's current mental states. However, strict, long-term punishments are incapable of taking this information into account. Thus, even if there are moral questions about the scope and detail of treating likes alike, this problem does not depend on settling them. For what is at issue here is an epistemic barrier *to even aiming* to treat likes alike, regardless of what this aiming fully amounts to. In other words, so long as we accept premise (2), strict long-term punishments close the door to treating likes alike under any interpretation, and hence close the door to satisfying a deep and powerful principle of morality.

### 3. The Transformative Choice Argument

As we saw, Paul develops the notion of a transformative experience, which involves an experience that is both epistemically and personally transformative. Such experiences, Paul argues further, raise problems for making rational choices involving them. This can be seen by considering a decision-theoretic framework. When calculating expected utilities within such a framework, two factors are taken into account: our subjective probabilities and our preferences. Suppose, for instance, that I am deciding whether to run a marathon. If I am calculating expected utilities, I should consider both the subjective probabilities that I would succeed, or fail, in running the marathon, and my preferences regarding doing so or not. I should decide to run the marathon, on this framework, if my calculations yield that this is the thing to do because it has the highest expected utility.

But Paul argues that there is a problem when decisions involve transformative experiences. Suppose, for instance, that I'm deciding whether to have a child for the first time. According to Paul:

“...[such an example brings] out two problems with our ordinary way of making important personal decisions from our subjective perspective when the decisions involve transformative choices.

First, transformative choices involve epistemically transformative experiences, compromising our ability to rationally assign subjective values to radically new outcomes. The subjective value of the lived experience that is the outcome of choosing to undergo the new experience is epistemically inaccessible to you, and this results in a type of ignorance that standard decision-theoretic models are ill-equipped to handle.

Second, because of the personally transformative nature of the experience, your preferences concerning the acts that lead to the new outcomes can also change. In particular, having the new experience may change your post-experience preferences, or change how

your post-experience self values outcomes. Transformative choices, then, ask you to make a decision where you must manage different selves at different times, with different sets of preferences. Which set of preferences should you be most concerned with? Your preferences now, or your preferences after the experience?” (Paul 2014, pp. 47-8)

The Problem of Transformative Choice, then, involves two key dimensions: first, since there is epistemic impoverishment with respect to the nature of a transformative experience prior to having it, there is an *ignorance of truths* that compromises our ability to rationally assign *subjective values* to the outcomes in question. For instance, prior to tasting a durian fruit or having a child, I am ignorant of the nature of this gustatory experience and of what it is like to be a mother, and so my ability to calculate expected utilities within a decision-theoretic framework is severely limited. Second, transformative experiences can bring about a *radical change in our preferences*. These preferences might be unknown to me, which is another layer of ignorance that compromises our ability to calculate expected utilities. But even where I might know that my preferences will radically change, there is still the other issue that Paul raises: whose preferences do I take into account, mine now, or my future self’s?<sup>35</sup>

I now want to consider the Problem of Transformative Choice in relation to the decision of whether to impose on someone a strict, long-term punishment, such as a natural life sentence. Suppose that I’m a judge and I’m deciding what the “appropriate punishment” is for a defendant by calculating the relevant expected utility of imposing a natural life sentence or not. Let’s remain as neutral as possible on what precisely an “appropriate punishment” amounts to, but one feature that I defended in the previous section is that it has to be sensitive to the current mental states of the person being punished. Given that there is ample evidence that people often radically change while they’re incarcerated for lengthy periods, either through typical transformations that occur between adolescence and being elderly or because of a transformative experience, there is a crucial ignorance

of truths on my part as the judge: who will this person be in 10, 20, or 30 years? In particular, in calculating the expected utility of imposing a natural life sentence, I need to consider the subjective probability that I will succeed, or fail, in appropriately punishing the defendant. But if the future mental states of the punishee are partially determinative of whether the punishment is appropriate, and relevant mental states can change dramatically over time via transformations and transformative experiences, then I am in a state of critical ignorance when making my decision. Hence, we have a parallel of the first dimension of the Problem of Transformative Choice.

It is interesting to emphasize the first-person and third-person aspects of this parallel. In Paul's original Problem of Transformative Choice, the decision is a first-person one: the fact that *I* might undergo a transformative experience compromises my ability to calculate the expected utility of my choice. Here, however, the decision is a third-person one: the fact that *someone else* might undergo a transformative experience compromises my ability to calculate the expected utility of my choice about that person's life.<sup>36</sup> While both versions crucially involve epistemic impoverishment that compromises my ability to assign the relevant subjective values to the outcomes, the object of the ignorance varies: my own experiences versus another's.<sup>37</sup>

The other dimension of this problem involves the radical change in preferences that can come about via either transformations or transformative experiences, and we again find a parallel with decisions about imposing strict, long-term punishments. Preferences regarding appropriate punishments can, and do significantly change.<sup>38</sup> Some are the result of what I have been calling mere transformations, such as slow and steady changes in social attitudes, while others might be more properly regarded as transformative experiences, but both are relevant here. Let's begin with an example of the former: social attitudes regarding incarceration have radically changed in the past 20 years or so, a fact that might be powerfully illustrated by the transformation of the views of both Bill and Hillary Clinton on the matter. In her *The New Jim Crow: Mass Incarceration in the Age of*

*Colorblindness*, Michelle Alexander writes:

[I]n 1992, presidential candidate Bill Clinton vowed that he would never permit any Republican to be perceived as tougher on crime than he. True to his word, just weeks before the critical New Hampshire primary, Clinton chose to fly home to Arkansas to oversee the execution of Ricky Ray Rector, a mentally impaired black man who had so little conception of what was about to happen to him that he asked for the dessert from his last meal to be saved for him until the morning. After the execution, Clinton remarked, “I can be nicked a lot, but no one can say I’m soft on crime.” (Alexander 2012, p. 56).

After being elected, Clinton signed into law the well-known Violent Crime Control and Law Enforcement Act of 1994,<sup>39</sup> which included, among various other things, \$9.7 billion in funding for prisons, a significantly expanded federal death penalty, and mandated life sentences for criminals convicted of a violent felony after two or more prior convictions, including drug crimes (this was known as the “three-strikes” provision). In 1996, Hillary Clinton gave a speech in New Hampshire in support of this Act, where she made her now infamous comment about “superpredators”:

...[we] have to have an organized effort against gangs...just as in a previous generation we had an organized effort against the mob. We need to take these people on. They are often connected to big drug cartels, they are not just gangs of kids anymore. They are often the kinds of kids that are called superpredators — no conscience, no empathy. We can talk about why they ended up that way, but first, we have to bring them to heel.<sup>40</sup>

These sorts of attitudes led to sentences becoming harsher and longer, resulting in skyrocketing incarceration rates in the United States. The decisions were based in part on preferences regarding appropriate punishment by lawmakers, judges, prosecutors, and the broader society at large.

Recent years, however, have seen a dramatic shift in attitudes regarding criminal justice, with it now being widely agreed that earlier views about sentencing were simply wrong. Last year, Bill

Clinton acknowledged that the 1994 crime bill was problematic in various ways: "...I want to say a few words about [criminal justice reform]. Because I signed a bill that made the problem worse and I want to admit it."<sup>41</sup> And, of course, Clinton isn't alone; a recent headline on NPR reads, "20 Years Later, Parts Of Major Crime Bill Viewed As Terrible Mistake."<sup>42</sup> Importantly, this acknowledgement is not simply the result of seeing that the crime bill contributed directly to the problem of mass incarceration, but it is also due to a fundamental transformation of attitude:<sup>43</sup> "Criminal justice policy was very much driven by public sentiment and a political instinct to appeal to the more negative punitive elements of public sentiment rather than to be driven by the facts."<sup>44</sup> Hillary Clinton, too, has now called for an end to mass incarceration, distanced herself from much of the 1994 crime bill, and has expressed regret about her "superpredators" comment, saying, "Looking back, I shouldn't have used those words, and I wouldn't use them today."<sup>45</sup>

Judges and victims' families, too, experience transformations that lead to radical changes in attitudes and preferences. In a recent case, a judge testified at a hearing that he wrongly convicted a defendant of murder because he was "prejudiced during the trial":

"As I read [a transcript of the original trial], I couldn't believe my eyes," the former judge said in an interview. "It was so obvious I had made a mistake. I got sick. Physically sick."

Mr. Barbaro's change of heart led to a highly unusual spectacle this week in a Brooklyn courtroom: He took the witness stand in State Supreme Court to testify at a hearing that his own verdict should be set aside.<sup>46</sup>

In another case, Jeanne Bishop, the sister of a woman who was murdered, along with her husband and unborn child, wrote a book about forgiving the man who committed the murders, David Biro. Bishop writes:

Many people—there is no shortage of them—are willing to write off the David Biros of the world. I was one of those people. Here was our argument: Look at what he did! It's so evil,

so depraved, that only a malignant heart could have concocted it. He is without feeling. Any remorse he might express later is only a sham. He will never change.

It is not true. I know this from my own transformation. God changed my heart. Why not the heart of David Biro? Why not the hearts of the thousands of people languishing in prison who have committed crimes for which we are willing to lock them up forever, without a second thought?" (Bishop 2015, p. 152).

These passages represent radical changes in attitudes and preferences at some of the most crucial levels of the decision-making process about punishment within the criminal justice system—lawmakers, judges/juries, victims' families, and society more broadly. It is difficult to say precisely what caused these radical changes, or whether they involved gradual transformations, transformative experiences, or some combination thereof. But regardless of the causal origin, they raise a pressing, epistemic problem for making decisions about strict, long-term punishments, one that parallels the second dimension of Paul's Problem of Transformative Choice; namely, that we might be deeply epistemically impoverished with respect to our future preferences regarding appropriate punishment and, even if we're not, there is the crucial question: whose preferences do we take into account, ours now, or those of our future selves?

Notice, however, that the two dimensions of the Problem of Transformative Choice do not thereby lead to the conclusion that any decisions made in the face of the corresponding epistemic impoverishment are thereby epistemically irrational. It may not, for instance, be irrational to taste a durian fruit, even in a state of ignorance about both the "what it is like" of such an experience and one's own future preferences regarding it. But in the case of imposing strict, long-term punishments on people, the case is different. While I will not attempt to generate general epistemic principles here for when, and only when, epistemic impoverishment leads to irrationality, I will highlight several features that distinguish tasting durian fruit from our topic at hand.

First, when handing out strict, long-term punishments, especially lengthy prison sentences, the stakes are high, and this can bear on whether the action in question is rational. There is no need to grant a thesis as controversial as pragmatic encroachment, according to which the standards for knowledge can vary, depending on the stakes, in order to maintain that stakes can bear on questions of epistemic rationality.<sup>47</sup> For instance, it may be irrational for me to not exercise greater epistemic caution around peanuts than you do when my child has a life-threatening allergy to them and yours does not. This doesn't necessarily mean that given the same evidence, you know, say, that there aren't peanuts in this slice of cake while I do not. Rather, it might be the case that we both know that there aren't peanuts in the cake but, given the incredibly high stakes, I need a grade of knowledge closer to certainty to act on this knowledge while you do not.<sup>48</sup> Similarly, there might be a greater degree of epistemic irrationality when the door is closed to relevant evidence in high-stakes situations than in low-stakes ones. Imposing a strict 1-day punishment of no dessert, for instance, is far less irrational than a natural life sentence is, even though neither is open to revision, and at least one reason for this is that the stakes are much higher in the latter than they are in the former. The same is true in the case of durian fruit: the stakes are so low that it might be rational for me to taste it, even while I'm in a state of epistemic impoverishment about such an experience and my future preferences regarding it. But clearly the same isn't true when talking about depriving a fellow human being of freedom for decades.

Relatedly, the decision in question crucially involves the lives of others, not just of myself. This feature can be built into the high stakes nature of imposing strict, long-term punishments, but it is worth highlighting in its own right. It may, for instance, be adequate for me to quickly check that my life jacket is on properly before sailing, but it might be necessary for me to double or triple check that this is so when I'm responsible for my child or yours.

Finally, there is a simple practical matter: some decisions are final, and so there is no



possibility of leaving the door open to revision in light of new evidence. For instance, once one has a child, there is no turning back, even if one later discovers that it is not what one hoped it would be. In contrast, there is a very easy solution to the epistemic impoverishment we face when deliberating about punishments: don't make them strict. Recognize that profound and relevant changes can happen at every level of the process, and that our ignorance of the future should lead us to avoid binding out present selves when we don't have to. Indeed, even choices that we expect to significantly constrain our future selves, such as marriage, can be revisited on the basis of new evidence. This is precisely why divorce is legal.

We are now in a position to see that there is a slightly modified version of the Transformation Argument here, which we may call the *Transformative Choice Argument*. While transformations and transformative experiences are relevant in both arguments, this title emphasizes that the structure of this argument models Paul's involving transformative choice. Like the earlier argument, however, the Transformative Choice Argument brings to light the epistemic irrationality of strict, long-term punishments. There are two different and possibly even interacting ways in which transformations and transformative experiences enter the epistemic picture here. On the one hand, the fact that the *punished* might radically change leads to an ignorance of truths on the part of the punisher, thereby compromising the punisher's ability to rationally assign subjective values to the outcomes in question. On the other hand, the fact that the *punisher* might radically change leads to an ignorance about which preferences ought to be taken into account when calculating expected utilities within a decision-theoretic framework. Moreover, we can certainly imagine these experiences interacting. Imagine, for instance, David Biro's experience upon hearing that the sister of his victims not only forgave him for the murders, but actually chose to cultivate a relationship with him. In such a case, Jeanne Bishop's transformative experience might be a catalyst for David Biro's own transformative experience. Given these two areas in which ignorance might compromise the

rationality of decision-making, it is seriously epistemically problematic to make high-stakes decisions that are closed to the possibility of future revision. Strict, long-term punishments are, therefore, irrational.<sup>49</sup>

It is crucial to note that this argument depends on premises already defended in the previous section, such as (1) and (3). The main difference, and what has been the primary focus of the arguments in this section, is that premise (2) needs to be broadened to include not only the mental states of the punishee after the time of the punishable act, but also those of the other relevant parties, such as the lawmakers, judges/juries, victims' families, and society more broadly. Rather than being understood as an entirely different argument, then, it is best to think of the Transformative Choice Argument as a modified or extended version of the Transformation Argument, according to which:

1. Punishment at a given time ought to be sensitive to the relevant evidence available at that time.
- 2\*. Punishment ought to be sensitive to the mental states of the punishees and punishers,<sup>50</sup> where this includes their mental states after the time of the punishable act.
3. People can change, often in profoundly transformative ways, which can involve radical changes in their corresponding mental states.
4. Strict, long-term punishments screen off any relevant future evidence, including the radically different mental states of people who have changed in transformative ways.
5. Therefore, strict, long-term punishments are epistemically irrational.

While only the second premise explicitly differs between the Transformation Argument and the Transformative Choice Argument, it should be clear that my development of Paul's framework also provides further support for (3) and (4). By focusing on the possibility of transformative experiences at many levels of the punitive process, and by providing examples of how the mental states of

people at these levels bear on the appropriateness of the punishments in question, we have a deeper and more expansive understanding of how people can change over time, and how strict-long-term punishments screen off this relevant evidence in problematic ways.

#### 4. Objections and Replies

I will now consider several objections to the arguments offered in this paper and provide some responses to them.

One theme that runs through many of the arguments is that we should not close epistemic doors to future evidence that might bear on the appropriateness of a given punishment, particularly the mental states of the punishee. This is perhaps most evident in my defense of premises (1) and (2) of the Transformation Argument. However, we have a number of practices, both legal and social, that close epistemic doors to varying degrees in ways that might seem to be at odds with these premises. Consider statutes of limitations: in many states, for instance, the statute of limitation on defamation is one year. Given this, even if new evidence is uncovered regarding a purportedly defamatory claim two years after it was first made public, it cannot be used for a lawsuit because of the statute of limitation. But then doesn't this seem to fly in the face of premises (1) and (2), according to which punishment at a given time ought to be sensitive to the evidence available at that time, including the mental states of the punishee? In particular, one who has defamed another cannot be punished for doing so in many states beyond a year, even if there is new evidence that bears on it. Doesn't this, then, screen off the possibility of punishment tracking the available evidence in such cases?

The short answer to this question is yes, it does, but this need not undermine premises (1) and (2). The reason for this is that the justification for such statutes of limitation might be compatible with them being problematic in other ways. To see this, notice that statutes of limitation

are longer the more serious the injury or crime, with some crimes having no statute of limitation at all, such as murder. At least one of the explanations for this is that a lack of closure places a significant burden on those who might be open to claims or charges, and this needs to be weighed against the seriousness of the original matter. For instance, the stakes in at least many defamation cases are fairly low with minimal damages when compared with the burden of people living under a threat of a defamation lawsuit for the rest of their lives, but clearly such a burden is outweighed when talking about murder. In an effort to achieve the most just system overall, then, the law must function at a level of generality that isn't always optimal in other ways. Thus, there may be times when our practices close epistemic doors—in violation of (1) and (2)—because this is what is regarded as all-things-considered best. However, this doesn't mean that screening off available evidence isn't problematic in other ways, such as epistemically and morally.

A second objection that might be raised to the view defended here is that I seem to be endorsing a problematically weak conception of punishment. For instance, suppose that someone commits a particularly heinous crime and has a transformative experience within minutes of being arrested. Or suppose that there is a transformative experience pill that brings about a radical change in preferences upon taking it, and someone takes the pill immediately after committing an act of extreme violence. Am I saying that in both of these sorts of cases, the person in question shouldn't be punished because he or she had a radical change in relevant mental states?

By way of response to this objection, notice, first, that none of my arguments in this paper depends on the strong thesis that *only* the mental states of the person being punished matter for whether a punishment is warranted. Rather, I am committed to the weaker thesis that mental states are *a* factor in determining appropriate punishment. Hence, even if a transformative experience pill, or less extraordinary means, brings about a radical shift in, say, the punishee's attitude toward his or her actions, there may still be other reasons why punishment is called for, such as retribution, justice

for the victim or the victim's family, functioning as a deterrent to others, and so on. Moreover, and this is meant to be merely suggestive, the history of a transformation or a transformative experience might matter, and hence the fact that radical changes are directly brought about through manipulation might diminish the bearing it has on whether a given punishment is warranted. There are instructive parallels to draw on here, as it is often noted that historical factors matter to autonomy,<sup>51</sup> responsible agency,<sup>52</sup> and moral accountability and moral character.<sup>53</sup> Suppose, for instance, that I don't want to do the hard work involved in cultivating virtues so I instead take a "courage pill." Even if this leads to my having dispositions to behave in ways typical of those who are courageous, it may be doubted that I in fact have courage or that I have the right kind of courage to ground moral responsibility. This is because the origin of my character trait matters, either to whether I have it in the first place, or to the way in which it reflects normative facts about me. A similar line might be run here: a transformative experience pill might not be adequate for bearing on the punishment a punishee deserves because it might have the wrong kind of history. This might be made particularly vivid by imagining someone who is otherwise remorseless setting out to commit a cruelly violent act against another, counting on popping a transformative experience pill after the fact to reduce his or her prison sentence.<sup>54</sup>

A third objection is why I focus on strict, long-term punishments rather than merely strict ones. In particular, if the central objection to such punishments is that they screen off potentially valuable epistemic information, why wouldn't the problem arise with respect to all punishments that are not open to revision, regardless of whether they are long-term?

There is a sense in which this objection is correct that strictness is the feature of punishments that shoulders much of the epistemic significance here. In particular, the epistemic irrationality of the punishments at issue lies primarily in their screening off potentially relevant evidence by being closed to revision. But this doesn't mean that being long-term has no bearing on

the rationality at all. First of all, the longer the punishment, the more time there is for transformations or transformative experiences to occur in all involved. Consider, for instance, the fact mentioned earlier that only 1% of serious crime is committed by people over the age of 60. Given this, a strict 40-year sentence on a 20-year-old is far more likely to screen off epistemically relevant evidence than is one imposed for two years on a 20-year-old. The long-term nature of the first punishment, then, makes it more epistemically objectionable than the second, even though they are both strict. Moreover, as noted earlier, stakes can bear on questions of epistemic rationality insofar as there might be a greater degree of epistemic irrationality when the door is closed to relevant evidence in high-stakes situations than in low-stakes ones, even without endorsing pragmatic encroachment. Again, as mentioned above, imposing a strict 1-day punishment of no dessert is far less irrational than a natural life sentence is, even though neither is open to revision, and at least one explanation here is the difference in stakes.

Fourth, it might be objected that if my view is that epistemic rationality requires that we leave the epistemic door open to decreasing the length of punishments, wouldn't the very same reasoning lead to the possibility of increasing them? Suppose, for instance, that a defendant feels remorse at the time of his sentencing and this is partly responsible for him receiving a 20-year-sentence, but at the end of the 20 years he has transformed and feels no remorse at all. Should years be added to his sentence because of the change in his mental states?

The short answer to this question is that if mental states of a punishee are relevant evidence for the punishment deserved, and if mental states can radically change over time—either positively or negatively—then I am committed to saying that from an epistemic point of view, the door should be left open to revising our beliefs and corresponding actions in light of such changes, *regardless of the direction*. But notice that this is in fact the way that even some very successful penal systems work. In Norway, for instance, there are no natural life sentences. The maximum sentence—even for

someone like Anders Behring Breivik, who was convicted of murdering 77 people<sup>55</sup>—is 21 years, though such a sentence can, through “preventive detention,” be extended indefinitely for five years at a time if the person in question is regarded as a continuing threat to society.<sup>56</sup> Thus, despite receiving a 21-year-sentence, Breivik could have many years added to this punishment, depending on his mental states down the road.

Moreover, notice that my conclusion here takes into account *only epistemic considerations*. There are many other kinds of factors that bear on the appropriateness of punishment, such as moral, practical, psychological, and so on, that can impose limits on punishment from both directions. On the one hand, and as mentioned above, even if a punishee radically transforms in 24 hours after committing a crime, justice might still call for a *minimum* punishment, such as a set number of years in prison, regardless of the relevant mental states. This sentence would be driven by factors independent of the evidential value of the punishee’s mental states. On the other hand, in the interest of providing closure for those involved, or to avoid endless work for the courts, or to err on the side of punishing too lightly rather than too severely, there might be a cap set on the *maximum* punishment, such that no matter what the mental states of the punishee are, it cannot be extended on this basis alone. So, for instance, even if a bank robber changes from feeling regret to being remorseless while being incarcerated, justice might require that the sentence not extend beyond, say, 10 years. The upshot of these considerations, then, is that even if my view commits me to saying that epistemic rationality requires leaving the door open to the length of punishments decreasing or increasing, based on the relevant mental states, epistemic factors might ultimately be swamped by other ones in practice.

Finally, it might be argued that some strict punishments are epistemically acceptable because we are *certain* that an action warrants a particular minimal punishment, and thus the conclusion that *all* strict long-term punishments are epistemically irrational is too strong. For instance, it might be

absolutely clear that at least a 20-year prison sentence is appropriate for first-degree murder, and so there might be no reason to leave the epistemic door open when handing out such a sentence.

Given this, only the weaker conclusion that *some* strict long-term punishments are irrational can be drawn.

By way of response, notice that there is, and has been, widespread variation in what is regarded as appropriate punishment, both across the globe and within our own history. For instance, punishment has involved what we now would regard as barbaric torture, such as being hanged, drawn, and quartered, or being flayed. Some countries, including the United States, continue to execute as punishment, while most Western countries have abolished the death penalty on the grounds that it is cruel and inhuman. Some nations have fines or light prison sentences for robbery, while others punish with caning or cutting off the offender's hand. Given these examples, it might be tempting to think that public attitudes have changed in a linear fashion toward greater leniency, but the United States stands out as a clear counterexample to this. In the past twenty years alone, the length of incarceration has increased dramatically in this country, with natural life sentences striking many Americans as the norm for a murder conviction. Moreover, what is regarded as the maximum penalty in a nation can, in turn, shape what its citizens regard as justice being served. When Breivik was sentenced to 21 years in prison in Oslo in 2012, *The New York Times* reported, "some parents who lost children in the attack appeared to be satisfied with the verdict, seeing it as fair punishment that would allow the country, perhaps, to move past its trauma. 'Now we won't hear about him for quite a while; now we can have peace and quiet,' Per Balch Soerensen, whose daughter was among the dead, told TV2, according to The Associated Press. He felt no personal rancor toward Mr. Breivik, he was quoted as saying."<sup>57</sup> If such a sentence were given for the same crime in the United States, it is not difficult to imagine a dramatically different response. Given this, there is overwhelming evidence of our own fallibility where punishment is concerned, with history showing



us time and time again that views can change as quickly as it takes for a new political office to be filled. At the very least, then, this recognition should give rise to at least some doubt that precludes certainty regarding the appropriateness of a particular strict long-term punishment. In other words, history provides us with excellent evidence for humility where punishment is concerned.

## 5. Conclusion

Many types of arguments have been leveled against long-term punishments, especially natural life sentences. Economic ones focus on the ballooning costs of mass incarceration and the toll this takes on government budgets, especially as the age and medical expenses of prisoners rapidly increase. Legal ones ask whether such sentences are cruel and unusual and therefore violate the Eighth Amendment, particularly for juveniles. Social arguments ask whether natural life sentences discourage reform by providing no incentive for rehabilitation. Moral concerns are grounded in the dignity and rights of agents, while psychological objections call attention to the myriad causes of deviant behavior and their responsiveness to appropriate treatment. But what has been absent from these conversations is an epistemic argument that has to do with us—with our inability to say now what someone will be like years from now, especially when transformations and transformative experiences are a possibility.

In this paper, I have attempted to fill this gap in the discussions. I have argued that strict, long-term punishments are epistemically irrational, with natural and *de facto* life sentences being the paradigmatic examples. Of course, this doesn't mean that, as a matter of fact, no one should ever serve lengthy prison sentences. Instead, the point is that no matter what the offense is, the door ought to be left open for revisiting the punishment in light of new evidence, particularly that brought about through radical changes in mental states.<sup>58</sup> In the criminal justice system, this means that it is irrational for the possibility of parole to be taken off the table at the outset of any sentence.

This will promote practices that are not only epistemically proper, but also ones that treat those being punished with respect, understanding, and mercy.<sup>59</sup>

### References

- Alexander, Michelle. 2012. *The New Jim Crow: Mass Incarceration in the Age of Colorblindness*. New York: The New Press.
- Audi, Robert. 1997. "The Place of Testimony in the Fabric of Knowledge and Justification." *American Philosophical Quarterly* 34: 405-22.
- , 1998. *Epistemology: A Contemporary Introduction to the Theory of Knowledge*. London: Routledge.
- Bergmann, Michael. 1997. "Internalism, Externalism and the No-Defeater Condition." *Synthese* 110: 399-417.
- Bishop, Jeanne. 2015. *Change of Heart: Justice, Mercy, and Making Peace with My Sister's Killer*. Louisville, KY: Westminster John Know Press.
- BonJour, Laurence. 1980. "Externalist Theories of Epistemic Justification." *Midwest Studies in Philosophy* 5: 53-73.
- , 1985. *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press.
- BonJour, Laurence and Ernest Sosa. 2003. *Epistemic Justification: Internalism vs. Externalism, Foundations vs. Virtues*. Oxford: Blackwell Publishing.
- Burge, Tyler. 1993. "Content Preservation." *The Philosophical Review* 102: 457-488.
- , 1997. "Interlocution, Perception, and Memory." *Philosophical Studies* 86: 21-47.
- Chisholm, Roderick M. 1989. *Theory of Knowledge*, 3rd edn. Englewood Cliffs, N.J.: Prentice-Hall.
- Corwin, Emily P., Robert J. Cramer, Desiree A. Griffin, and Stanley L. Brodsky. 2012. "Defendant

- Remorse, Need for Affect, and Juror Sentencing Decisions.” *The Journal of the American Academy of Psychiatry and the Law* 40: 41-9.
- Dolinko, David. 1991. “Some Thoughts About Retributivism.” *Ethics* 101: 537-59.
- Douven, Igor. 2006. “Assertion, Knowledge, and Rational Credibility.” *The Philosophical Review* 115: 449-85.
- Eisenberg, Theodore, Stephen P. Garvey, and Martin T. Wells. 1997-1998. “But Was He Sorry? The Role of Remorse in Capital Sentencing.” *Cornell Law Review* 83: 1599-1637.
- Fantl, Jeremy and Matthew McGrath. 2009. *Knowledge in an Uncertain World*. Oxford: Oxford University Press.
- Fischer, John and Mark Ravizza. 1998. *Responsibility and Control*. Cambridge: Cambridge University Press.
- Fricker, Elizabeth. 1987. “The Epistemology of Testimony.” *Proceedings of the Aristotelian Society*, supp. vol. 61: 57-83.
- . 1994. “Against Gullibility,” in Bimal Krishna Matilal and Arindam Chakrabarti (eds.), *Knowing from Words*. Dordrecht: Kluwer Academic Publishers: 125-61.
- Goldberg, Sanford. Forthcoming. “Should Have Known.” *Synthese*.
- Goldman, Alvin. I. 1986. *Epistemology and Cognition*. Cambridge, MA: Harvard University Press.
- Hawthorne, John. 2004. *Knowledge and Lotteries*. Oxford: Oxford University Press.
- Hawthorne, John and Jason Stanley. 2008. “Knowledge and Action.” *The Journal of Philosophy* 105: 571-90.
- Jackson, Frank. 1982. “Epiphenomenal Qualia.” *The Philosophical Quarterly* 32: 127-36.
- Johnson, Sara B., Robert W. Blum, and Jay N. Giedd. 2009. “Adolescent Maturity and the Brain: The Promise and Pitfalls of Neuroscience Research in Adolescent Health Policy.” *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine* 45: 216-221.

- Kapitan, Tomis. 2000. "Autonomy and Manipulated Freedom." *Philosophical Perspectives* 14: 81-103.
- Kvanvig, Jonathan L. 2009. "Assertion, Knowledge, and Lotteries" in Patrick Greenough and Duncan Pritchard (eds.), *Williamson on Knowledge*. Oxford: Oxford University Press.
- Lackey, Jennifer. 2007. "Norms of Assertion." *Noûs* 41: 594-626.
- , 2008. *Learning from Words: Testimony as a Source of Knowledge*. Oxford: Oxford University Press.
- Levy, Ken. 2014. "Why Retributivism Needs Consequentialism: The Rightful Place of Revenge in the Criminal Justice System." *Rutgers Law Review* 66: 629-84.
- Mabbott, J.D. 1939. "Punishment." *Mind* 48: 152-67.
- McDowell, John. 1994. "Knowledge by Hearsay," in Bimal Krishna Matilal and Arindam Chakrabarti (eds.), *Knowing from Words*. Dordrecht: Kluwer Academic Publishers: 195-224.
- McKinnon, Rachel. 2013. "The Supportive Reasons Norm of Assertion." *American Philosophical Quarterly* 50: 121-35.
- Maruna, Shadd, Louise Wilson, and Kathryn Curran. 2006. "Why God Is Often Found Behind Bars: Prison Conversions and the Crisis of Self-Narrative." *Research in Human Development* 3: 161-84.
- Mele, Alfred. 1995. *Autonomous Agents*. Oxford: Oxford University Press.
- Miller, William R. and Janet C'deBaca. 1994. "Quantum Change: Toward a Psychology of Transformation," in Todd F. Heatherton and Joel Lee Weinberger (eds.), *Can Personality Change?* Washington D.C.: American Psychological Association Books: 253-80.
- Neta, Ram. 2009. "Treating Something as a Reason for Action." *Noûs* 43: 684-99.
- Nozick, Robert. 1981. *Philosophical Explanations*. Cambridge, MA: The Belknap Press.
- Paul, L. A. 2014. *Transformative Experience*. Oxford: Oxford University Press.
- Pfaff, John F. 2017. *Locked In: The True Causes of Mass Incarceration and How to Achieve Real Reform*. New

- York: Basic Books.
- Plantinga, Alvin. 1993. *Warrant and Proper Function*. Oxford: Oxford University Press.
- Pollock, John. 1986. *Contemporary Theories of Knowledge*. Totowa, N.J.: Rowman and Littlefield.
- Pullman, Philip. 2001. *The Amber Spyglass*. London: Scholastic Press.
- Reed, Baron. 2006. "Epistemic Circularity Squared? Skepticism about Common Sense." *Philosophy and Phenomenological Research* 73: 186-97.
- , 2010. "A Defense of Stable Invariantism." *Noûs* 44: 224-44.
- Stanley, Jason. 2005. *Knowledge and Practical Interests*. Oxford: Oxford University Press.
- Tuerkheimer, Deborah. 2014. *Flawed Convictions: "Shaken Baby Syndrome" and the Inertia of Injustice*. Oxford: Oxford University Press.
- Ward, Bryan H. 2006. "Sentencing without Remorse." *Loyola University Chicago Law Journal* 38: 131-67.
- Williams, Michael. 1999. *Groundless Belief: An Essay on the Possibility of Epistemology*, 2<sup>nd</sup> edn. Princeton: Princeton University Press.
- Williamson, Timothy. 2000. *Knowledge and its Limits*. Oxford: Oxford University Press.
- , 2005. "Contextualism, Subject-Sensitive Invariantism and Knowledge of Knowledge." *The Philosophical Quarterly* 55: 213-35.
- Winston, Kenneth I. 1974. "On Treating Like Cases Alike." *California Law Review* 62: 1-39.

---

<sup>1</sup> Source: <http://www.cdc.gov/nchs/fastats/life-expectancy.htm>, accessed 15 October 2016.

<sup>2</sup> Source: <http://fairsentencingofyouth.org/wp-content/uploads/2010/02/Michigan-Life-Expectancy-Data-Youth-Serving-Life.pdf>, accessed 15 October 2016.

<sup>3</sup> Source: <http://www.icjia.org/cjreform2015/pdf/Olson%20-%20Impact%20of%20TIS%20Report.pdf>, accessed 15 October 2016.

---

<sup>4</sup> Source: <http://www.icjia.org/cjreform2015/pdf/Olson%20-%20Impact%20of%20TIS%20Report.pdf>, accessed 15 October 2016.

<sup>5</sup> This is the insensitive invariantist version of the KNPR. The contextualist version is: “A first-person present-tense ascription of ‘know’ with respect to a proposition is true in a context iff that proposition is an appropriate premise for practical reasoning in that context” (Williamson 2005, p. 227).

<sup>6</sup> Hawthorne and Stanley restrict their conditions to “*p*-dependent choices” since *p* may simply be irrelevant to a given action. Since the cases discussed in this paper all involve *p*-dependent choices, I will ignore this complication in what follows.

<sup>7</sup> Even if an action is not actually the product of practical reasoning, it might still be the case that it needs to be capable of being endorsed by practical reasoning, if one were to engage in it.

<sup>8</sup> This is prevalent in the literature on the norm governing epistemically proper assertion.

<sup>9</sup> See, for instance, Kvanvig (2009). It should be noted that this norm is distinct from Neta’s *justified-belief-that-one-knows* norm.

<sup>10</sup> See Douven (2006).

<sup>11</sup> See Lackey (2007).

<sup>12</sup> See McKinnon (2013).

<sup>13</sup> For various views of defeaters, see BonJour (1980 and 1985), Nozick (1981), Goldman (1986), Pollock (1986), Fricker (1987 and 1994), Chisholm (1989), Burge (1993 and 1997), Plantinga (1993), McDowell (1994), Audi (1997 and 1998), Bergmann (1997), Williams (1999), BonJour and Sosa (2003), Hawthorne (2004), Reed (2006), and Lackey (2008).

<sup>14</sup> See Lackey (2008). See also Goldberg (forthcoming) for a very illuminating account of when a subject “should have known” a given proposition.

<sup>15</sup> Source: <http://www.innocenceproject.org/dna-exonerations-in-the-united-states/>, accessed 15 October 2016.

<sup>16</sup> See, for instance, Tuerkheimer (2014).

---

<sup>17</sup> I should say at the outset that my arguments will most likely be largely unpersuasive to those who espouse an uncompromising retributivism about punishment, according to which a given punishment is inflicted *solely* because a criminal deserves it. “This desert can take two forms: either a ‘moral debt’ that only criminal punishment can discharge or a *right* to criminal punishment that the criminal ‘earned’ through her commission of a crime” (Levy 2014, p. 645). However, I should also emphasize that nothing I say in this paper rules out punishment having a retributivist *component* or *dimension*.

It is interesting to note that attitudes about retributivism have changed rather dramatically over the years. In 1939, J.D. Mabbott wrote, “The retributive view is the only moral theory except perhaps psychological hedonism which has been definitely destroyed by criticism” Mabbott (1939, p. 152). In the years since then, however, retributivism has grown in popularity. See, for instance, Dolinko (1991) for a discussion of the development of retributivism. For a recent defense of a version of retributivism, see Levy (2014), where he argues that “retributivism...is fueled by vengeance” (2014, p. 684) and that “vengeance plays a *legitimate* and *central* role in the criminal justice system” (2014, p. 633).

<sup>18</sup> Source: <https://www.washingtonpost.com/news/grade-point/wp/2016/06/14/i-take-him-at-his-word-judge-in-stanford-sexual-assault-case-explained-controversial-sentence/>, accessed 15 October 2016.

<sup>19</sup> Source: <https://www.theguardian.com/us-news/2016/jun/07/brock-turner-statement-stanford-rape-case-campus-culture>, accessed 15 October 2016.

<sup>20</sup> Source: <http://www.cbsnews.com/news/professor-admits-role-in-killing-of-her-alleged-rapist/>, accessed 15 October 2016.

<sup>21</sup> I should mention that some legal scholars, including Ward himself, push back against the relevance of remorse in criminal sentencing because of the “difficulties inherent in both defining remorse and in identifying it in the criminal defendant” (Ward 2016, p. 133). Despite this, however, there can be no doubt that remorse in fact plays a very substantive role in our criminal justice system.

<sup>22</sup> See, for instance, Eisenberg, Garvey, and Wells (1997-1998) and Corwin, Cramer, Griffin, and Brodsky (2012).

---

<sup>23</sup> Source: <http://www.npr.org/2016/02/18/467057603/from-a-life-term-to-life-on-the-outside-when-aging-felons-are>

[freed?utm\\_source=npr\\_newsletter&utm\\_medium=email&utm\\_content=20160219&utm\\_campaign=npr\\_email\\_a\\_friend&utm\\_term=storyshare](http://www.npr.org/2016/02/18/467057603/from-a-life-term-to-life-on-the-outside-when-aging-felons-are), accessed 15 October 2016. See also

<https://www.youtube.com/watch?v=JgbEnjnLtgc>, accessed 15 October 2016.

<sup>24</sup> Source: [https://www.aacap.org/AACAP/Families\\_and\\_Youth/Facts\\_for\\_Families/FFF-Guide/The-Teen-Brain-Behavior-Problem-Solving-and-Decision-Making-095.aspx](https://www.aacap.org/AACAP/Families_and_Youth/Facts_for_Families/FFF-Guide/The-Teen-Brain-Behavior-Problem-Solving-and-Decision-Making-095.aspx), accessed 15 October 2016.

<sup>25</sup> Source: [http://articles.courant.com/1997-02-18/news/9702180121\\_1\\_prison-population-three-strikes-average-inmate/2](http://articles.courant.com/1997-02-18/news/9702180121_1_prison-population-three-strikes-average-inmate/2), accessed 31 October 2016.

<sup>26</sup> Experts project that the elderly prison population in the U.S. will be over 400,000 in 2030, compared with 8,853 in 1981. [https://www.aclu.org/files/assets/elderlyprisonreport\\_20120613\\_1.pdf](https://www.aclu.org/files/assets/elderlyprisonreport_20120613_1.pdf), accessed 15 October 2016.

<sup>27</sup> A recent article in *Quartz* has the headline: “You’re a Completely Different Person at 14 and 77, the Longest-Running Personality Study Ever Has Found”: <https://qz.com/914002/youre-a-completely-different-person-at-14-and-77-the-longest-running-personality-study-ever-has-found/>. The article discusses work recently published in *Psychology and Aging* that looks at six personality traits of subjects at 14 and then again at 77—self-confidence, perseverance, stability of moods, conscientiousness, originality, and desire to learn—and concludes that there is “no significant stability of any of the 6 characteristics” over the 63-year interval. This study adds further support for premise (3).

<sup>28</sup> See Jackson (1982).

<sup>29</sup> Paul also discusses this example at length in her (2014).

<sup>30</sup> Source: <http://www.cnn.com/2016/10/04/us/iyw-coss-marte/>, accessed 18 October 2016.

<sup>31</sup> See Miller and C’deBaca (1994).

<sup>32</sup> I should note that while I emphasize *stages* of persons in order to highlight the transformations that occur over time, this should not be understood as bringing with it any metaphysical commitment to punishing



---

stages of persons rather than persons themselves. Rather, we should consider, say, *the person at 14* and *the person at 49* when punishment is concerned.

<sup>33</sup> It should be clear that similar issues arise with respect to the death penalty. This point is made particularly vivid in a recent article in *The Atlantic* (<http://www.theatlantic.com/politics/archive/2016/11/a-deadly-question/508232/>, accessed 6 December 2016), “A Deadly Question: Have Juries Sentenced Hundreds of People to Death by Trying to Predict the Unpredictable?” The author, Abbie Vansickle, writes:

Before jurors could sentence someone to death [in Texas], they must first decide if the person will be a future danger.

The precise wording of the question is convoluted, asking jurors “whether there is a probability that the defendant would commit criminal acts of violence that would constitute a continuing threat to society.” At its core, it contains an incredible idea: Can we predict whether or not a killer will kill again?

Vansickle goes on to show that experts are in agreement that the answer to this question is a negative one:

Dr. Mark Cunningham, a Seattle-based psychologist, and Dr. John Edens, a psychologist at Texas A&M University, have devoted their professional lives to the question of whether we can predict the future dangerousness of those convicted of crimes. Both have published extensively on the topic. And both have reached much the same conclusion.

“Juries show absolutely no predictive ability whatsoever,” Cunningham said. “And, in fact, experts are similar.”

The American Psychiatric Association...has taken a similar position and implored the Supreme Court to ban the future dangerousness question in capital cases, saying in an amicus brief that “[t]he unreliability of psychiatric predictions of long-term future dangerousness is by now an established fact within the profession.” The APA concluded that the “future dangerousness” question relies on junk science, and found that experts are wrong in two out of three predictions of “future dangerousness.”

---

Given the possibility of transformations and transformative experiences, answers to the question about future dangerousness might be even more misguided, as past actions will not be a reliable guide to future ones when later selves are radically different from their earlier selves.

<sup>34</sup> See, for instance, Winston (1974).

<sup>35</sup> While Paul situates her argument within a decision theoretic framework, I do not wish to here take on board all aspects of such a framework. Instead, what is crucial for my purposes is the epistemic impoverishment involved when transformative experiences are concerned, which Paul's discussion very nicely develops.

<sup>36</sup> I should mention, however, that Paul does consider cases of parents deciding to give their child cochlear implants, which more closely parallels my argument here. (See Paul 2014, pp. 56-61.)

<sup>37</sup> Of course, this version of the puzzle arises only when the possibility of the person in question changing radically is *relevant* to my decision, which is clearly the case when sentences are being imposed.

<sup>38</sup> In addition, values can change dramatically over time, both at the individual and the collective levels, and these can be crucial dimensions of appropriate punishments.

<sup>39</sup> Source: <https://www.ncjrs.gov/txtfiles/bills.txt>, accessed 22 October 2016.

<sup>40</sup> Source: <http://www.politifact.com/truth-o-meter/statements/2016/aug/28/reince-priebus/did-hillary-clinton-call-african-american-youth-su/>, accessed 22 October 2016.

<sup>41</sup> Source: <http://www.msnbc.com/msnbc/clinton-admits-his-crime-bill-made-mass-incarceration-worse>, accessed 24 October 2016.

<sup>42</sup> Source: <http://www.npr.org/2014/09/12/347736999/20-years-later-major-crime-bill-viewed-as-terrible-mistake>, accessed 24 October 2016.

<sup>43</sup> Another very important factor here, especially for our purposes, is that social attitudes can change in light of *new evidence*. For instance, as scholars, such as Alexander in her (2012), show how mass incarceration is in part the result of racially driven practices, attitudes toward sentencing in the United States have changed.

<sup>44</sup> Source: <http://www.npr.org/2014/09/12/347736999/20-years-later-major-crime-bill-viewed-as-terrible-mistake>, accessed 24 October 2016.

---

<sup>45</sup> Source: <https://www.washingtonpost.com/news/post-politics/wp/2016/02/25/clinton-heckled-by-black-lives-matter-activist/>, accessed 24 October 2016.

<sup>46</sup> Source: <http://www.nytimes.com/2013/12/13/nyregion/ex-brooklyn-judge-seeks-reversal-of-his-verdict-in-1999-murder-case.html>, accessed 24 October 2016.

<sup>47</sup> For a classic defense of pragmatic encroachment, see Stanley (2005) and Fantl and McGrath (2009).

<sup>48</sup> See Reed (2010) for this view.

<sup>49</sup> It should also be noted that some decisions are made by groups or collectives, such as juries and states, and it is an open question how to understand transformations and transformative experiences within this framework. There are, for instance, metaphysical questions about what it means for groups to transform or have transformative experiences in the relevant sense. Does this mean merely that some individual members of the group have the experiences in question, or does something need to take place at the level of the group? There are also epistemic questions about how the possible changes in preferences of the member of a group bear on their decision-making. If, say, a small percentage of the group's members undergo a transformative experience that changes their preferences, does this then change the group's preferences? While these are fascinating questions, they lie beyond the scope of the present paper.

<sup>50</sup> I am understanding “punishers” here broadly to include lawmakers, judges/juries, victims' families, and society more broadly.

<sup>51</sup> See Mele (1995).

<sup>52</sup> See Fischer and Ravizza (1998).

<sup>53</sup> See Kapitan (2000).

<sup>54</sup> The example in the text involves popping a transformative experience pill *after* a punishable act, but similar questions might be asked about *preemptive* punishment. This calls to mind a scene from Philip Pullman's *The Amber Spyglass*:

“I propose to send a man to find her and kill her before she *can* be tempted.”

“Father President,” said Father Gomez at once, “I have done pre-emptive penance every day of my adult life. I have studied, I have trained—“

---

The President held up his hand. Pre-emptive penance and absolution were doctrines researched and developed by the Consistorial Court, but not known to the wider church. They involved doing penance for a sin not yet committed, intense and fervent penance accompanied by scourging and flagellation, so as to build up, as it were, a store of credit. When the penance had reached the appropriate level for a particular sin, the penitent was granted absolution in advance, though he might never be called on to commit a sin. It was sometimes necessary to kill people, for example: and it was so much less troubling for the assassin if he could do so in a state of grace.

(Pullman 2001, p. 75)

<sup>55</sup> Source: <http://www.nytimes.com/2012/08/25/world/europe/anders-behring-breivik-murder-trial.html>, accessed 8 October 2017.

<sup>56</sup> Source: <https://www.nytimes.com/2015/03/29/magazine/the-radical-humaneness-of-norways-halden-prison.html>, accessed 8 October 2017.

<sup>57</sup> Source: <http://www.nytimes.com/2012/08/25/world/europe/anders-behring-breivik-murder-trial.html>, accessed 8 October 2017.

<sup>58</sup> This paper draws on some of the arguments I made here:

[http://opinionator.blogs.nytimes.com/2016/02/01/the-irrationality-of-natural-life-sentences/?\\_r=0](http://opinionator.blogs.nytimes.com/2016/02/01/the-irrationality-of-natural-life-sentences/?_r=0).

<sup>59</sup> For very helpful comments on earlier versions of this paper, I am grateful to Nilanjan Das, Enoch Lambert, Lauren Leydon-Hardy, Trevor Nyman, Laurie Paul, Kathryn Pogin, Baron Reed, John Schwenkler, and audience members at the Pre-Pacific APA Conference on Transformative Experience in Seattle, WA, the 34<sup>th</sup> International Social Philosophy Conference at Loyola University, and the Ethics and Values Workshop at Loyola University.