An Essay towards an Epistemology of Responsibility
: *A Probabilistic Approach*

Masaki Ichinose

Department of Philosophy, The University of Tokyo
Email: Ichinose *at* l.u-tokyo.ac.jp

1 Metaphysics and Epistemology

Responsibility is an unhappy concept in philosophy. It is definitely one of the most crucial concepts for us; nonetheless it is hardly focused upon primarily in comparison to its traditional partner, freedom. Let's take examples of Hume and J. S. Mill. They consider the problem of freedom or liberty so extremely seriously that they devote one chapter or one book to the discussion on freedom (i.e., chapter 8 "Of Liberty and Necessity" in Hume's *First Enquiry* and Mill's *On Liberty*), whereas responsibility is discussed by them only as a secondary issue as a derivative of freedom. Why is this? Probably it's because of the ambivalent character of responsibility. Let me explain.

Freedom or liberty has been supposed to be a metaphysical concept in the history of philosophy similarly to immortality of souls, God's existence, or personhood in that they are beyond observation. Certainly, the problem of freedom is raised in political contexts as well, but even in that case, freedom functions as an ideal to be aimed at, which cannot be empirically observed. Therefore, responsibility should be involved in this metaphysical flavour, since we usually suppose that responsibility is introduced as a result of free actions. This is represented by this famous expression, "could have done otherwise". If I could have done otherwise, I was free then so that I am responsible for my action I actually chose. If this is the case (which we rarely deny, as it's a very healthy understanding), then we have to regard responsibility as metaphysical too. Of course, we know that the phrase, "could have done otherwise", does not always establish the connection between freedom and responsibility, as Frankfurt's cases show, in which I intentionally do something under the condition that, even if I did not intend to do that, I would be forced to do that. In that case, we are not exactly free, but I did that by my intention, so I am responsible for that. However, apart

from those kinds of skillfully invented cases, freedom is, roughly speaking, tied with responsibility.   At least, the possibility of alternative actions is a sufficient condition (although not a necessary condition) for freedom and responsibility.

But, we must not stop the story here.   Responsibility appears much far more realistic and urgent in our daily life than freedom.   That is to say, if something wrong or harmful actually occurs in our society, we ought to make decisions, as far as possible, as regards how to attribute responsibility for that to deal with the harm.   The typical case is criminal responsibility. This aspect of responsibility is clearly beyond metaphysics.   We should collect evidence and search eye-witnesses to confirm what actually happened.   On the basis of those data, we come to know how to make decision on responsibility.   This process could be called "epistemological" rather than metaphysical, as it is exactly similar to the process of knowing something through perception or reasoning.   In short, responsibility is not only a metaphysical partner of freedom to be ideally aimed at, but also an epistemological object to be realistically known.   This double significance of responsibility probably has made it hard to scrutinize responsibility from a clear-cut point of view.   On second thoughts, however, its double character could be taken to be something positively helpful.   In fact, there is no clear-cut division between metaphysics and epistemology, as both are intimately connected, which is exemplified by such an epistemological question about the metaphysical issue of God as; "how could we know that God exists?" To such questions, we have to give some epistemological grounds.   In this respect, responsibility is a quite helpful issue to clarify how those two major areas of philosophy stand each other, since responsibility covers both areas by definition.

The epistemological character of responsibility can be etymologically confirmed.   The concept of responsibility is well known to originate from a Greek word, "αἰτία", which means not only responsibility but also "cause". Namely, according to this etymology, for instance, to ascribe responsibility for some harm is nothing but to assign a cause of the harm, which could be a particular human behaviour or an agent's condition and what not.   This synonymity is still alive in English.   If this is the case, it is very natural that responsibility constitutes an epistemological issue, since the causal relation must be one of the main topics in epistemology.   In addition, the synonymity between cause and responsibility provides another explanation as to why responsibility can be discussed also from a metaphysical viewpoint, for, as you know, causality is

studied as a subject of metaphysics in a certain context. Taking into account those mapping argued thus far, I will focus upon the epistemological aspect of responsibility here, particularly by taking the issue of criminal responsibility.

Perhaps, someone might oppose to my standpoint by saying that responsibility belongs to *quid juris* rather than *quid facti*, hence that is not appropriate to be discussed from an epistemological point of view. This opposition is exactly true. I totally agree to that. However, my intention is not that responsibility should be studied only from an epistemological viewpoint, but that responsibility has a crucial aspect to be analysed in an epistemological manner. In fact, *quid juris* and *quid facti* are not independent of each other. *Quid juris* should be judged at the basis of *quid facti*. So, I suppose that my epistemological strategy could be justified.

2 Intrinsic Uncertainty

However, even if we could sort out our point of view from which responsibility will be discussed, a problem might come up. That is to say, responsibility is so full of uncertainties that it appears impossible from the outset to attain some knowledge about responsibility, if we define knowledge as "a justified TRUE belief" in a traditional manner. If we must exactly follow the traditional definition of knowledge, we could have only a quite limited amount of knowledge, e.g. mathematical or logical knowledge, which does not cover our ordinary usage of the word, "knowledge" at all. All of our empirical knowledge or second-hand knowledge illustrates this situation. For example, we can say that we know 'a brain-dead person cannot be resurrected', or we know 'Adolf Hitler died in 1945', but, as you know, those kinds of knowledge cannot reach perfect certainty. Actually, Richard Jeffrey once stated that;

> Outside the realm of what we are sure of lies the puzzling region of probable knowledge －puzzling in part because the sense of the noun seems to be cancelled by that of the adjective (Jeffrey, 1992, p. 30).

Then, Jeffrey proceeds to put forward an idea of replacing the notion of knowledge by the concept of belief in developing epistemology. That's an entirely acceptable strategy in view of the contemporary context of philosophical debates on epistemology, as Bayesian epistemology symbolically represents. This situation applies to the epistemology of responsibility as well. It does not

matter which terminology to use, knowledge or belief.

Then, what kind of uncertainty is responsibility involved in? We can immediately raise at least three examples that represent uncertainties about responsibility. First, it appears that responsibility is not applied to the case in which responsibility must be applied most. By this I mean the paradoxical fact that in human history, a person who killed many people to win a war was often not attributed responsibility for murder but rather regarded as a heroic figure. This leads to a question. How many people must be killed to acquit the agents of their crime? We could probably answer this question only in a vague manner. This is also true in the case in which something wrong results from too many people's complicity (including those who have already died), a typical example of which is the bankruptcy of a country caused due to long-term inaction. We could also answer such a question in a vague manner as, 'how many accomplices should participate in a wrong behavior in order for them to escape from responsibility?'

Second, the probabilistic feature of responsibility can be easily confirmed by focusing on such a self-evident truth that past events to which responsibility is supposed to be attributed are not occurring now. We have no choice but to take into account probabilities for those past events to have actually occurred in evaluating responsibility, which is nothing but the job of historians or courts[1]. The issue of legal competence of a suspect probably strengthens my viewpoint on the probabilistic status of responsibility, because legal competence is mostly not judged either perfectly had or completely lacked, but it should be measured by considering the probabilistic degree. Of course, the English word "responsibility" is concerned with not only past events but also future possibilities, since responsibility often implies a type of duty towards the future. However, even so, the probabilistic feature of responsibility still holds, since we cannot be perfectly certain as to what we should do now to fulfill the responsibility for the very reason that it is concerned with the future.

Third, it is not clarified as to how we could assume responsibility for and deal with the aftermath of the issue in question. For example, Japanese people tend to quickly resign from their positions in order to take responsibility for their acts, similar to what was done by some former prime ministers of Japan. This tendency originates from the traditional ethics of Japan called *Bushido*, according to which *Seppuku* (or *Hara-kiri*, i.e., suicide by disembowelment) is 'a process by which warriors could expiate their crimes, apologies for errors, escape from

disgrace, redeem their friends, or prove their sincerity' (Nitobe, 2002, p. 107)[2]. However, objectively speaking, it is uncertain as to how resignation or suicide could work as a settlement of the aftermath.   Then, how else should we place responsibility?   We could probably answer this only in a vague manner.   In any case, those three examples strongly suggest that responsibility is intrinsically uncertain, as those three aspects of uncertainty about responsibility seem to be absolutely unavoidable.   I will eventually highlight two kinds of uncertainties of those, namely, a hopelessly self-destructive one so that it should be excluded from the substantial core of responsibility, and an unavoidably inbuilt one so that we must establish how to deal with it.

3 Criminal Responsibility

I will concentrate upon the issue of criminal responsibility, as I believe that the proper character of responsibility stands out the most in this issue.   First of all, let's confirm the general scheme of offence as to which criminal responsibility matters.   The scheme common to any form of offence could be shown in this way;
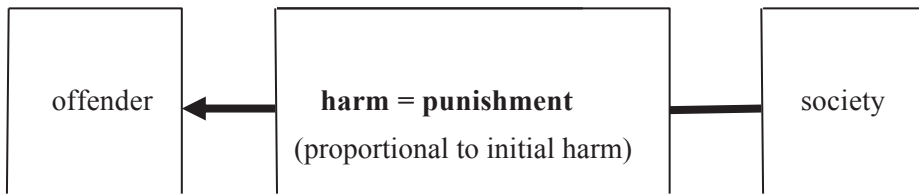
<The Scheme of Offences>
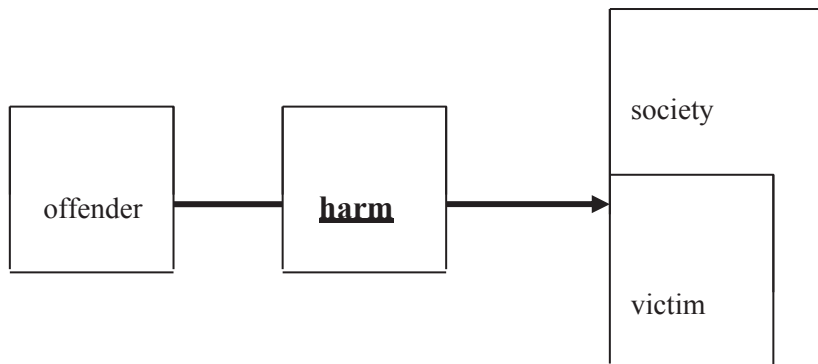
The arrow denotes the direction of causal relation.   In other words, offences is nothing but the phenomena in which an offender causes harm to a victim and our society, which means that an occurrence of HARM is the whole basis for the problem of offences.   The occurrence of harm triggers the problem.   It logically follows from this basic point that there are three aspects around harm, namely,

offender, victim, and society.

Generally speaking, we react to offences by applying the system of Criminal Justice, or to put it differently, Retributive Justice.　According to the system, the scheme of offences should exactly reverse when we deal with the aftermath of those offences like this;

\<The System of Retributive Justice\>

```
┌──────────┐     ┌──────────────────────────┐     ┌──────────┐
│          │ ◄─── │  harm = punishment        │     │          │
│ offender │     │ (proportional to initial harm) │     │ society  │
│          │     │                          │     │          │
└──────────┘     └──────────────────────────┘     └──────────┘
```

In principle, our society gives back the offender some amount of harm in proportion to the amount of initial harm caused by the offender.　It is called "retributive justice" because of the relation of giving back.　We usually suppose that this reaction is regarded as a solution of offences committed.　That is to say, this system of retributive justices is generally understood as making offenders take responsibility.　Of course, as you know, punishment can be justified by the utilitarian theory that considers consequences of punishment like prevention of crimes or rehabilitating offenders, as well as the retributive theory.　I do not deny the utilitarian justification of punishment; otherwise punishment has nothing meaningful, nothing constructive.　All that I want to notice here is that the retributive reaction is an absolutely necessary condition for initiating the system of punishment, and the utilitarian treatment should be given only under that condition.　This viewpoint of mine roughly conforms to John Rawls's arguments as to punishment, in which he states;

> One must distinguish between justifying a practice as a system of rules to be applied and enforced, and justifying a particular action which falls under these rules; utilitarian arguments are appropriate with regard to questions about practice, while retributive arguments fit the application of particular rules to particular cases (Rawls, 1955, p. 146).

It is quite clear why the retributive aspect is supposed to be indispensable, because in principle, we do not need to punish offenders themselves but we may give harm, for example, to a family member of those offenders if we justify the system of punishment only through utilitarianism, which sounds overwhelmingly counterintuitive.

In short, as far as the present system of justice is concerned, the system consists of both of these two perspectives; <u>retrospective</u> and <u>prospective</u> perspectives. Retribution only focuses upon past events, whereas the phase of punishment justified by utilitarianism considers the future benefits. This contrast, of course, corresponds to two connotations of the concept of responsibility that I mentioned earlier, namely, responsibility for past events and towards future possibilities. In a word, responsibility is the hybrid of two perspectives, which might also make responsibility a bit ambivalent. In any case, this situation reconfirms the point that there are intrinsic uncertainties as regards responsibility.

4 Two Practical Difficulties

However, any institutional system is imperfect. This is true of retributive justice, as a practical matter of fact. Rather, we should say that drawbacks of retributive justice are too serious to keep it as it has been thus far. In fact, it seems to me that what is needed before anything else is not to search the justification for that, but to renovate the system itself. I will raise two basic difficulties.

First, the fundamental problem as to retributive justice is the fact that this is the offender-centred system. As I described, the core of offences lies in the occurrence of harm which is caused by the offender to the victim and their society. There are three aspects which compose the offence or the occurrence of harm, that is, offender, victim, and society. However, the offender-centred system ignores the two other aspects, i.e., victim and society. This neglect is rather intentional. The system of retributive justice has been established as an institution of our modern society, in which revenge or lynch is prohibited. According to the classical story provided by the social contract theory as John Locke developed, our society must have been established by our mutual consents to quit the natural right to punish offenders and to resign the right up into the hand of the community (See Locke, 1960, section 87, p. 324). Therefore, to punish someone in a private manner should be excluded. This sounds reasonable as far

as a fair and impartial implementation of retributive treatment is regarded as the only goal of justice.   But, the question that I want to raise here is; how is the harm in question solved by punishing the offender?   Did the offender take responsibility to deal with the harm by being punished?   This sort of questions should be answered, as, as I have stated a couple of times, the core problem as to offences is nothing but the occurrence of harm.   The offender-centred system seems to largely miss this main point, probably because that system has given the top priority to the safety of the society.   To this argument, someone might say that victims are allowed to bring a civil action against the harm they suffered in order to gain compensation[3].   But, as you know, it is an onerous burden for victims to bring civil actions and gain reparations, as it takes a lot of costs, time, and labour for them.   In this respect, we should still say that there is a serious difficulty with retributive justice as the offender-centred system.

Second, the system of retributive justice recently has focused upon the subjective perceptions or mental state of offenders too much.   Of late, what is called affirmative defences doctrine such as self-defence, provocation, and duress 'permits the defendant asserting these defences to submit evidence about his or her own feelings and thoughts at the time of the offense' (Slobogin, 2006, p. 34). Its prominent example is the so-called "battered women syndrome".   That is 'the state of "leaned helpfulness" allegedly visited on women who suffer through cyclical battering from their spouse or significant other' (Slobogin, 2007, p. 23). Such women sometimes could harm their partner even when no violence is perpetrated by their partner.   This syndrome claimed by the defendant and clinical testimony based on the defendant's subjective feeling is used as a self-defence claim and tends to be admissible in the court at least in the US.   Of course, *mens rea*, which is the traditional requirement of culpability, is the basis of the contemporary subjectification trend.   Certainly, this trend appears to be acceptable if we contemplate what harm implies in the scheme of offence described above.   Offences cause harm by definition, but the harm essentially carries moral implications unlike harm caused by natural phenomena.   Anthony Duff correctly points out that;

> The victim of crime has been not just harmed, but *wronged*; he has suffered a wrongful, as distinct from a natural or merely unlucky (Duff, 2003, p. 46).

In other words, even though we are caused harm by others, we will not regard the

harm as offences if the harm is morally neutral at the level of the agent's mental state, a typical example of which is the case that I am pushed in standing in a crowded train to fall by other standing people because of a sudden breaking of the train without any morally bad intention of the people pushing me.   I might be harmed in this case, but the harm will not be supposed to be an offence.   On the contrary, even if nothing harmful occurs in a physical sense, we feel offended if we perceive some wrong intention in other's mental state like in the case of verbal abuse or sexual harassment.   Thus, causal relation between offender and harm is not purely physical, but includes some moral factors like *mens rea* or blameworthiness that have to do with agents' mental state.

So far, so good.   However, the more seriously we consider the agent's subjective perception or mental state as a standard of offences, the more clinical testimonies of psychiatrists or psychologists tend to matter in the court, as agents' subjective perceptions or mental state must be somehow testified to be an admissible claim.   Then, 'subjectifying blameworthiness in this way opens the door wide to psychological speculation' (Slobogin, 2007, p. 25).   This naturally leads to the situation that clinical testimonies or medical diagnoses as to mental disorders of the defendant are frequently required in the court, since mental disorder crucially affects subjective perceptions or mental state of offenders. However, this brings about quite perplexing difficulties.   What are these?

5 Mentally Disordered Offenders

Those difficulties can be described in the form of dilemma.   The issue of criminal responsibility of mentally disordered offenders is well known in terms of what is called "Insanity Defence".   There are lots of complicated histories about insanity defence, but it is clear that this defence was noticed as one of the central issues of criminal justice in terms of the case of Daniel McNaghten in England in 1843 as in the following;

> Daniel McNaghten, suffering from the paranoid delusion that the Prime Minister, Sir Robert Peel, was part of the plot to persecute and to destroy him, shot at and killed Peel's secretary believing him to be Peel (Ten, 1987, p. 124).

Such sort of delusion, being incomprehensible from an ordinary point of view, was regarded as a salient characteristic of insanity, by reason of which he was

acquitted on a charge of murdering Sir Robert Peel's private secretary.　This case leads to what is currently called "the M'Naghten Rule," which is virtually an origin of "Insanity Defence."　The Rule states that;

> To establish a defence of insanity it must be clearly proved that at the time of committing the act, the accused was labouring under such a defect of reason, from disease of the mind, as (a) not to know the nature and quality of the act he was doing; or (b) if he did know it, not to know he was doing what was wrong (Allen, 1991, p. 114).

As is clearly described here, the M'Naghten Rule refers only to cognitive conditions, namely, cognitive impairments in making the defence of insanity, whereas as a matter of fact, volitional impairments (i.e., inability to control their own will) and irrationality (i.e., irrational way of thinking or reasoning) are usually claimed as well in the contemporary contexts for the insanity defence.

This insanity defence is destined to be involved in a serious dilemma. Requirements for the insanity defence are conceptually tied with mental disorder or mental illness like hysteria or schizophrenia, although we should be very careful of the distinction between medical diagnoses and legal judgments, since medical diagnoses provide only clues for legal judgements.　Generally speaking, mental states of offenders in view of the insanity defence are analysed, as I mentioned above, by 'the two-pronged test (focusing on cognitive incapacity to appreciate the criminal or wrongful nature of the act, or volitional inability to conform one's conduct to the requirements of law)' (Grisso, 1996, p. 92).　In addition, some scholars stress irrationality or inability to employ practical reasoning as the reason for the insanity defence (See Schopp, 1991, pp. 108-131).

Those three inabilities or impairments are mostly supposed to arise in connection with mental disorders.　I do not intend to chase the problems as to details of mental disorders themselves too far, for my interest here only lies in the scrutiny on how clinical testimony works as to the issue of criminal responsibility.　However, the actual situation as to the insanity defence has been desperately chaotic and unstable.　At least, we can immediately find three phases to cause such instability.　Firstly and fundamentally, it is almost impossible to recognize past mental state of offenders at the time of the offence, which is nothing but the problem of historical knowledge in an epistemological sense. Second, as a matter of fact, the standard to support the insanity defence has

fluctuated with regard to whether volitional condition should be considered or not. Rather, it is sometimes questioned whether even the insanity defence itself should be retained or abolished. Third, a lot of so-called syndromes like black rage, cocaine-induced psychosis, battered women's, battered children's have been proposed so that relevant legal questions requiring clinical assessment associated with criminal responsibility has been expanded too much. As a result of this unstable legal history of the insanity defence, the contemporary subjectification trend is lost in a maze.

In fact, the concept of mental disorder or mental illness has been problematic from the outset. To confirm this point, we should remember the argument of Thomas Szasz, who is famous for his *antipsychiatry* movement. Probably, main points proposed by Szasz could be summarized as in the followings.

(1) Psychiatry is the study of personal conduct in terms of verbal communications between a patient and a psychiatrist.

(2) As far as psychiatry seeks to explain or change human behaviour, statements concerning goals and values must remain indispensable.

(3) Therefore, psychiatry cannot be empirical science like biology or physics, for their problems cannot be solved by medical methods.

(4) In modern medicine new diseases were *discovered* by observing a patient's bodily structure, whereas in modern psychiatry new diseases were *invented* by observing a patient's behaviour. Paresis was *proved* to be a disease; hysteria was *declared* to be one (See Szasz, 1974, pp. 1-13.)

(5) So, "labeling individuals displaying or disabled by problems in living as "mentally ill" has only impeded and retarded the recognition of the essentially moral and political nature of phenomena" (Szasz, 1974, p. 25).

Szasz was afraid that 'not only psychoanalysis but also much of traditional and modern psychiatric theory assumes that personal conduct is determined by prior personal-historical events'(Szasz, 1974, p. 5), because he felt that this deterministic understanding might negate our moral concepts like responsibility. His deepest anxiety arises from the extreme fact that 'psychiatrists have asserted malingering, too, is a form of mental illness. This presents us with the logical absurdity of a disease which, even when it is deliberately counterfeited, is still a disease' (Szasz, 1974, p. 13). Consequently, the concept of responsibility must

collapse together with the insanity defence if we theoretically pursue the current subjectification trend to its extreme.

6 Biological Fantasy

I will not totally agree with Szasz's position, because he seems to simply presume that medicine, unlike psychiatry, is genuinely a study of bodily structures in an organic sense, as I suppose that any category of bodily disease cannot be exempt from culture-laden characteristics, which is typically exemplified by the case of the deaf.　It depends upon our value or culture as to whether the deaf are ill or not.　However, in any case, Szasz's argument is definitely worth noticing and unignorable.　Then, how could psychiatry defend itself against his criticism?　There seems to be one possible route for psychiatry to avoid his attack.　That is the biologicizing mental disorder which results in making psychiatric diagnoses a kind of natural science.　For example, a Japanese psychiatrist, Akira Fukushima, has proposed the concept of murderer's insanity or homicidal insanity, and pointed out that almost all people belonging to this category are found, on the basis of medical tests about their brains, to have some slight variants of brain structures, in particular a frontal lobe.　He calls these variants "Minimal Brain Organic Variants" (See Fukushima, 2005, pp. 211-221). I imagine that there can be a plenty of similar proposals or reports in other countries.　If homicidal behaviours are caused by such sort of organic variants, psychiatry could be objectively founded by biology or brain science so that they can defend against Szasz's criticism.

However, as you easily perceive, this will introduce another horn of the dilemma.　In fact, this biologicization of homicidal insanity is nothing but a bit refined revival of Lombroso's phrenological idea of criminal man with regard to their theoretical structures (Lombroso, 2006).　Or, notorious histories around "born criminal", for instance, the Jacobs study (Hubbard & Wald, 1999, pp. 105-106) or the case of Juke family (Nelkin & Lindee, 2004, p. 25), are ultimately birds of a feather.　Exactly speaking, that's a biological fantasy, which has been not adequately confirmed but rather simply mistaken, or just statistically conjectured at most.　Nevertheless, the problem arises.　These contexts tend to motivate people to adopt biological or genetic determinism, as a result of which the concept of responsibility must be largely obscured.

I am not questioning the classical problem about whether determinism is compatible or not with freedom or responsibility.　Of course, I know that, even if

we adopt determinism, we could theoretically keep the concept of responsibility by proposing the viewpoint of a kind of compatibilism like Hume's. Rather, what I try to question is concerned with the uncertainty of responsibility which is my main interest here. Namely, if we attribute the cause of homicidal behaviour of a certain man to his brain organic variants, we could pursue the deeper cause of the variants into another level of physical phenomena, for example, some accidental damage given when the man was an embryo in his mother's womb or the man's parents' gene. The problem is that there is no definite criterion as to where we should stop such pursuits. Any stopping point is just arbitrary. Hence, we could not place responsibility anywhere. This seems to be hopeless uncertainty. As I mentioned, there is intrinsically uncertainty with regard to responsibility, but that uncertainty must not be the radical one that nullifies responsibility. The uncertainty must be the one that we can somehow deal with, as we need to keep the concept of responsibility to survive.

In any case, if we accept either the subjectification trend or the biologicization trend of mental disorder in view of insanity defence, the concept of responsibility is involved in a serious dilemma then to collapse. This is the first type of uncertainties that I mentioned earlier as so hopeless to be excluded. It seems to me that this situation reflects a crucial defect of the system of retributive justice itself.

7 Hybrid between Retribution and Restoration

Let's come back to the starting point. I want to reconfirm two basic points.

(a) The concept of responsibility stems from the concept of "cause", although the cause in this case implies not only physical phenomena but also mental state like *mens rea*.

(b) All problems of criminal responsibility begin with the occurrence of harm. The core element of offences is this occurrence of harm.

So, what we should do is to search a different system dealing with offences from retributive justice in terms of taking those two points into account. First, we should focus upon the occurrence of harm. As I pointed out, the offender-centred system, i.e. retributive justice, is (intentionally) omitting this basic point. According to my understanding, the system strayed into a labyrinthine dilemma because of this omission. If this is the case, the simple

and natural proposal to solve the dilemma is to change our attention straightforwardly to the occurrence of harm.   That is to say, we ought to offer the harm-centred system of offences and responsibility.

The system of restorative justice, which has been actually practiced in many countries, rather than retributive justice appears to be quite helpful in order to develop the harm-centred system, because restorative justice basically pays attention to the restoration of harm with regard to offences, which is a dramatically different concept from retributive justice.   But, exactly speaking, what is restorative justice?   The idea of restorative justice was invented by finding the present system of retributive justice unsatisfactory.   The next passage of Tony Marshall is often quoted as a definition of restorative justice.

> Restorative Justice is a process whereby parties with a stake in a specific offence collectively resolve how to deal with the aftermath of the offence and its implications for the future (Marshall, 1999. p. 5).

His next passage is also helpful to understand restorative justice.

> Restorative Justice is centrally concerned with restoration: restoration of the victim, restoration of the offender to a law-abiding life, restoration of the damage caused by crime to the community.   Restoration is not solely backward-looking; it is equally, if not more, concerned with the construction of a better society in the present and the future (Marshall, 1999, p. 7).

However, 'it raises a number of questions: Who are the "parties with a stake in the offender"?   How do they come to a collective resolution? What does it mean to "deal with the aftermath of an offence"?   What "implications for the future" should be considered?' (Van Ness, Morris, and Maxwell, 2001, p. 5)   To begin with, it is not clear what a restoration is.   Obviously it is easily imaginable to restore the harm in the case of simple theft of jewelry.   It would be restored if the jewelry was sent back. But, how about wounds to the emotions?   How about mementoes being stolen?   How about murdered people?   Second, how is restorative justice different from retributive justice if offenders should be, more or less, compelled to restore the harm they caused?   In fact, some people initially found the idea of restorative justice quite attractive and utopian, but then were disappointed to think about its feasibility, as a kind of compulsion should be

needed in carrying it out (E.g., see Acorn, 2004). That sounds to be eventually the same as retributive justice. As to this point, Anthony Duff objectively points out;

> Restoration is not only compatible with retribution: it requires retribution, in that the kind of restoration that crime makes necessary can (given certain deep features of our social lives) be brought about only through punishment (Duff, 2003, p. 43).

Those questions cannot be answered without considering the debate between two types of restorative justice, namely, the purist model vs. the maximalist model. The purist model principally focuses on the restoration of the relation between an offender and its victim through their direct conversation. This method is called "victim-offender mediation" (VOM). As is immediately understood, this method can be applied only to quite limited cases of offence, because victims are not always willing to meet offenders, and mentally disordered offenders could not communicate with victims. Therefore, we should suppose that the purist model of restorative justice is proposed not as an alternative of retributive justice but as a complement to that. Although I do not deny that the model helps very much to realize the peaceful and constructive solution with regard to some specific offences, the purist model cannot work well in view of my interest in mentally disordered offenders. This is the system thinking much of restoration of human relation rather than that of harm itself, so it is suitable to call that "victim-centred system", which is subtly different from the harm-centred system that I hope to establish. On the other hand, the maximalist model aims at radically challenging the system of retributive justice itself. Lode Walgrave, a supporter of the maximalist model, declares that;

> The maximalist version of restorative justice......aims at developing a fully fledged justice system that would be consequentially oriented towards doing justice through restoration and would be an alternative to the existing punitive or rehabilitative justice system (Walgrave, 2001. p. 18, footnote 1).

That is to say, the maximalist model of restorative justice replaces the infliction of punishment to offenders by the restoration of the harm with concerted effort between society, offenders, and victims. This system could be depicted in the

following diagram.

\<The Maximalist Model of Restorative Justice\>
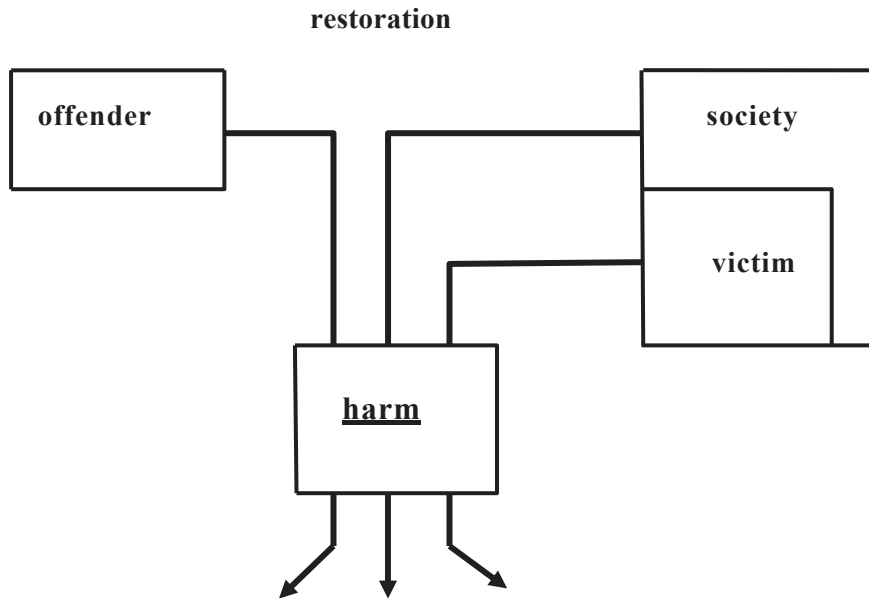
restoration

As far as this system aims at the practical system in our real society, the practice of restoration should be understood not literally but realistically. For example, 'offenders, in principle, give financial payment, work for the victim, etc.' (Marshall, 1999, p. 11) as far as possible.   If offenders cannot do that for some reason like in the case of seriously mentally disordered offenders, victims themselves should try to restore the harm by the assistance of the society.   As is understood in this description, this system unites criminal cases with civil cases. In addition, this system will include an occasion for offenders to recover their dignity through their participation in some practice of restorative justice, as 'offenders are prepared to be perceived as victims, even if they show a great reluctance to be labeled as "mentally disordered"' (Peay, 1997, p. 688). Certainly, being involved in hideous scene of murder as an offender with being spattered by the blood of his victim is not a happy life.   The offender suffers miserable time.

However, as Duff pointed out as I mentioned before, this system of restorative justice cannot be implemented without compelling offenders to take

part in the restoration, which seems to be very similar to punishment. As I discussed, as far as offences and their responsibility are concerned, a retributive viewpoint is indispensable at least at the starting point to ignite the system, otherwise the concept of responsibility must be involved in the chaos. Restorative justice principally tries to forgo the concept of punishment in discussing responsibility, but the justice could not work without the notion of retribution, apart from the question about whether punishment makes sense only in a retributive manner or not. In this respect, unlike Walgrave, what we should propose here is a kind of the hybrid between restorative justice and retributive justice, in which we regard restoration of harm rather than punishment as the main goal of justice by basically inflicting on offenders the duty or responsibility to restore that.

8 Sentence and Practice

This idea of the hybrid system can be also bolstered by the etymological significance of responsibility as a cause, which is one of two basic points I referred to. It seems to me that, if the insanity defence does not work reasonably because of its hopeless uncertainties, one promising strategy is to forgo such defence and to simply return back to the original position. Namely, we consider neither subjective mental states nor biological conditions of offenders, and focus only upon causal relation between offensive behaviours of a man and its harm in order to place responsibility. This is what I want to offer as "the harm-centred system of responsibility". Two points should be noticed. First, this causal scheme seems to be likely to stray into hopeless uncertainty regarding responsibility in the same way as the case of the biologicization trend of mental disorder, but it is not the case. The biologicization intends to think about human behaviours as natural phenomena, whereas the harm-centred system considers an individual man who can be an offender as a unit of responsibility from the outset, so deeper pursuits beyond the individual man should not be given. Second, this causal scheme is also involved in uncertainty between cause and effect, but such uncertainty could be dealt with in terms of probabilistic causality. This uncertainty is intrinsic and unavoidable, while the uncertainty the biologicization faces is so hopeless as to entirely destroy the concept of responsibility. In any case, this harm-centred system in terms of simple causation initially has to focus upon an offender to place responsibility, so in that sense the system has an affinity with the idea of mixing retributive justice.

However, what should we think about the case of seriously mentally disordered offenders?   Couldn't they be engaged in restorative activities?   Such hard cases are possible indeed.   Should they be acquitted of restoration eventually?   But, if they should be acquitted, the causal scheme could not be retained and we would turn back to the dilemma of the insanity defence.   In order to avoid this, I want to propose the distinction between a sentence and a practice of restoration.   That is to say, offenders are sentenced about whether they have responsibility to restore the harm or not only by focusing upon *actus reus*, i.e., bad behaviours, without any consideration at all for their subjective mental state, mental disorder, or biological conditions.   That is to say, the insanity defence is completely abolished in the substantial court proceedings.   Only after those sentences, offenders are given assessment of their actual competency to practice the restorative activities by taking into account medical diagnoses with regard to their subjective and biological conditions.   If they are assessed to seriously lack the competency to practice the restoration because of mental illness, they have to be medically treated in the hospital, but even in that case their responsibility remains.   When they recover from the illness, they ought to immediately conduct the restorative activities.   In short, these two phases of restoration are introduced.

---

responsibility = sentenced as a cause of the harm

legal competency = competency to practice the restoration

Perhaps, someone is afraid that this idea is nothing but the revival of an ancient idea of "strict liability".   But, it is not true, because this system distinguishes between sentences and practices.   It has a cushion when it actually inflicts the responsibility of restoration on offenders.   My causal version of the harm-centred system is quite close to a type of sentence, "guilty but mentally ill", which is currently popular in some area, although in my case the same idea should be expressed as "responsible for the harm but mentally ill" instead of using the word, "guilty".   In addition, my version shares the basic idea with Slobogin's Integrationism.   My causal version of the harm-centred system of responsibility will result in the equation of mentally disordered people with healthy people as regards responsibility of offences in that, no matter how

disordered an offender is, responsibility of restoration is equally inflicted on the offender as far as the person was the cause of the harm.   Likewise, Slobogin's Integrationism basically treats mentally disordered offenders in the same way as normal offenders by focusing upon the existence or not of *mens rea*.   He says 'An integrationist regime would excuse people with mental disability only on the grounds available to *all* offenders' (Slobogin, 2006, p. 15).

9 Three Levels of Uncertainty

The main purpose to make responsibility return back to simple causation in my harm-centred system is to objectify the legal judgement about responsibility of offences as far as possible, because causal relations between human behaviours and harm, which are observable events in principle, are basically able to be tested in a scientific manner.   But, in spite of this, we could not escape from its intrinsic uncertainties as I have repeated.   Hence, in order to firmly establish the causal version of the harm-centred system of responsibility, we should face the problem of how to measure "degrees of responsibility" by considering that intrinsic uncertainties.   Those uncertainties arise in three different levels.   First two levels of uncertainties are these.

(α) There is uncertainty with regard to reliability of evidence (i.e., physical evidence, testimony of eyewitness, and so on) to testify the fact that the defendant truly did the behaviour in question.

(β) We must face uncertainty about how the behaviour caused the harm.

Uncertainty (α) is easy to understand.   Legal epistemology is usually discussed with regard to this reliability of evidence, for instance, through Bayesian ideas (See Bovens & Hartman, 2003, chap. 3 & chap. 5).   But, uncertainty (β) might sound a bit strange, as it seems to be not clear how (β) is different from (α).   Uncertainty (β) is proposed by considering the case that, although the behaviour eventually resulted in the actual harm, the probability for the behaviour to cause that specific harm was low.   For instance, please suppose the case that I glared fiercely at my rival in love with intending to kill him, and the friend was so extremely scared of my glaring that he suddenly died of heart failure.   Or suppose that I bawled at my rival so fiercely hoping to kill him that he was scared to die of heart failure.   In these cases, my glaring or my bawling caused my rival's death as a result, but it had been supposed to have very low probability for

such causal relation to actually hold.　In such cases I am scarcely responsible for my rival's death, because expected harm (considering both the probability to cause the harm and the amount of the harm) in such cases is not high. Otherwise, in fact, glaring at or bawling at someone must be regarded as an attempted murder even if nothing happened, which sounds counterintuitive.　Of course, there are multiple layers of degrees of this kind of uncertainties between innocence and being manifestly responsible, hence there must be borderline cases which are difficult to be judged.　That leads to the problem of vagueness.

However, as I paid special attention when I quoted Duff, harm in the case of responsibility is not just like natural disasters, but implies being wronged[4].　So, we must also consider *mens rea* when we decide how responsible an offender is. Thus, the third level of uncertainty appears.

(γ) We must deal with uncertainty as regards how true it is that the defendant did the behaviour in question from *mens rea*.

We could not avoid this uncertainty, but, as you immediately understand, the assessment of *mens rea* is desperately hard.　In fact, to pursue the subjectification trend too much should result in, as I pointed out, the collapse of responsibility.　Nevertheless, even if we should avoid adopting the insanity defence based on medical diagnoses about mental disorders, *mens rea* requirement is indispensable by definition as far as we discuss responsibility of offences.　But, how could we measure that uncertainty about *mens rea*?　In particular, what matters in this case is *mens rea* in past mental states of the agent at the time of offence.　It is epistemologically difficult to verify the truth of propositions in the past tense in general as problems of historical knowledge show, much worse it is to verify the truth of MENTAL state in the past.　How could we achieve that hopeless task?

10 Relevance Ratio

Of course, if we aim at attaining the perfect truth, we should immediately give up that kind of task and enjoy scepticism at once.　But, we are talking about the realistic problems and the public policy, so what we should do is to look for the better methodology even though not perfect.　In this respect, we have one promising way.　That is the idea of "Relevance Ratio" which was proposed by Lyon and Koehler.　They mentioned the situation that expert testimony in court

proceedings is often criticized as junk science, and against such criticism, they put forward the idea of relevance ratio in order to make such expert testimony more objective and scientific. Relevance they mention is a legal concept about evidence. It is usually prescribed that all evidence must be relevant to be admissible. According to Lyon and Koehler, 'Federal Rule of Evidence 401 defines "relevant evidence" as evidence that has "any tendency to make the existence of any fact that is of consequence to the determination of the action more probable or less probable than it would be without evidence"' (Lyon & Koehler, 1996, p. 46). They develop their argument on this notion of relevance by illustrating it in the context of expert testimony in child sexual abuse cases. They say that;

> One can determine whether the presence of symptoms increase the chance that abuse occurred by considering two proportions: the proportion of abuse cases in which the symptom occurs, and the proportion of nonabuse cases in which the symptom occurs. If the proportion of abuse cases exhibiting the symptom is greater than the proportion of nonabuse cases exhibiting the symptom, then the symptom is relevant for proving that abuse occurred (Lyon & Koehler, 1996, p. 46).

That is to say, relevance ratio is given by comparing the relevant proportion in both the abused and noabused population. Let "H" denote the hypothesis or the sentence that a child was abused, and "E" the sentence that the child exhibits this symptom, then this ratio can be expressed by the usual notation of conditional probability like this;

$$\frac{P(E|H)}{P(E|\sim H)}$$

This is the same as what is called "likelihood ratio" in the Bayesian scheme. It is known that Bayes's theorem is practically easy to apply in terms of the odds-likelihood ratio form. In fact, this odds-likelihood ratio form is highly recommended to use in the context of medical decision making (See Weinstein & Fineberg, 1980, chap. 4). The odds-likelihood ratio form is expressed by the next formulation of Bayes's theorem.

$$\frac{P(H|E)}{P(\sim H|E)} = \frac{P(H)}{P(\sim H)} \times \frac{P(E|H)}{P(E|\sim H)}$$

P(H|E) / P(H|~E) is called "posterior odds" and P(H) / P(~H) is called "prior odds" in Bayesian contexts. That is to say, if we know both the prior odds and the likelihood ratio between particular evidence and particular hypothesis, we can deduce the posterior odds on them, i.e., we can get the posterior probability of the particular hypothesis on the basis of the particular evidence. That's a common story of Bayesianism.

However, Lyon & Koehler introduce the term, "relevance ratio" instead of "likelihood ratio" in order to free the likelihood-based approach from its Bayesian baggage and to make their approach more fitted with legal context using the notion of relevant evidence. The reason why they hesitate to make commitment to Bayesian reasoning is the fact that there are serious controversies over the feasibility of identifying prior unconditional probabilities. Relevance ratio skips the consideration for prior probabilities, so that it will be much easier to practically use than the Bayesian scheme. In addition, they emphasize another merit of adopting the relevance ratio, that is, what would be mistakenly judged if we focus only on alleged symptoms in isolation could be objectively assessed. For instance, there might be a prosecutor who wishes to introduce expert testimony that a particular child suffered from low self-esteem to support his contention that the child was abused. But, even if data are available which indicate that many abused children suffer from low self-esteem, it may also be true that many nonabused children suffer from low self-esteem. Neither fact in isolation determines whether low self-esteem is relevant for proving abuse. But, relevance ratio works here. If a higher percentage of abused children exhibit low self-esteem than nonabused children, then this symptom is relevant for proving that abuse has occurred. However, in fact, many studies indicate that the relevance ratio for low self-esteem is near unity, so that this symptom fails a relevance test (Lyon & Koehler, 1996, p. 47).

Another remarkable deduction that can be easily drawn by adopting the relevance ratio scheme is concerned with the absence of symptoms. They illustrate this point in terms of the relevance of rape trauma syndrome for the

occurrence of abuse.   They explicitly declare that 'if statistics demonstrate the relevance of rape trauma syndrome for proving that abuse occurred, those same statistics demonstrate the relevance of a failure to exhibit rape trauma syndrome for proving that abuse did not occur' (Lyon & Koehler, 1996, p. 71).   This can be proved easily as follows: The relevance ratio associated with evidence E is

$$\frac{P(E|H)}{P(E|\sim H)} = \frac{X}{Y}$$

and the relevance ratio associated ~E is

$$\frac{P(\sim E|H)}{P(\sim E|\sim H)} = \frac{1-X}{1-Y}$$

If E is relevant for proving H, then X > Y.   Furthermore, if X > Y, then, as Algebra simply suggests, - X < -Y, then 1-X < 1-Y, therefore 1-X / 1-Y < 1. Consequently, the relevance ratio associated with ~E < 1, that is, ~E is probative of ~H.

11 Past Mental State

This relevance ratio analysis is remarkably simple and helpful.   However, since it's too simple, at least two doubts immediately arise.   First, could we completely ignore prior probability when we assess evidence?   It is often reported by psychological studies that ignoring prior probabilities results in serious mistakes about the posterior probability of the hypothesis in question (E.g. Tversky & Kahneman, 1973).   In fact, even though the relevance ratio of rape trauma syndrome for proving that abuse occurred is very high, for instance, more than 10, the posterior odds could be very low if the prior odds is extremely low, for instance, less than 0.01, which is really possible, for rape trauma syndrome can be extremely rare in safe countries like the past Japan.   In that case, the exhibition of rape trauma syndrome does not confirm that abuse occurred, at least if we follow the Bayesian confirmation theory.   However, this problem is not so serious, for it seems that the prior odds on most alleged syndromes or evidence is mostly unknown in our realistic situation[5].   If this is the case, we have no choice but to allocate 1 to prior odds, admitting the principle of indifference.   Namely,

as a practical matter of fact, the relevance ratio analysis is supposed to work well.

However, the second doubt seems to be more serious.   How could we apply the methodology of relevance ratio to the test on *mens rea* as past mental state, which is the main problem in this context?   Slobogin expresses deep concern about this after mentioning Lyon and Koehler's relevance ratio analysis.   He introduces his concern like this;

> Consider, however, the difficulty of gathering information necessary to calculate the relevance ratio where psychological evidence about past mental state is involved. For instance, when self-defence is asserted in a battered woman case, one would need to determine the proportion of battered woman who kill their spouse believing they have no alternatives and the proportion of nonbattered woman who kill their spouse believing they have no alternatives.   Without any other information, the most plausible conclusion is that, of the two groups, the battered women are more likely to feel that they have no option. That kind of thinking, though is not science; the scientific method would require large enough samples of both groups being compared (Slobogin, 2007, p. 45).

This criticism sounds quite reasonable.   Obviously it must be extremely hard, or nearly impossible, to appropriately collect data of PAST MENTAL states, in that the condition of *ceteris paribus* could not be satisfied and various elements depend on each other.   Certainly, in the child abuse example, it's possible to observe the existence or nonexistence of abuses and their symptoms mostly in an objective way, whereas in the case of *mens rea*, it seems to be hopeless to do so. However, we must be careful to recognize what the issue is.   What matters here is concerned not with the existence of some syndromes but with the existence of *mens rea*, which is nothing but what my version of the harm-centred system of responsibility questions.   I agree with Slobogin that some (newly invented) syndromes cannot be well treated by the relevance ratio analysis.   But it is not necessarily so in the case of *mens rea*.   We should not suppose that we cannot perceive *mens rea* in any sense at all, because, if we cannot, nobody can judge whether the defendant had *mens rea* or not, so that *mens rea* could not be the legal issue.   Intuitively speaking, we usually perceive other people's *mens rea* or morally bad intentions through such symptoms as their countenances, their ways of behaviours, their voices, and their ways of offending.   There are lots of data

accumulated concerning those.   Therefore, we can compare the proportion of people's action with *mens rea* who exhibit such and such symptoms with the proportion of people's action without *mens rea* who exhibit the same symptoms, irrespective of questions about particular syndromes.   This line of argument is analogues to Frank Ramsey's idea of understanding degrees of belief in terms of behaviours of betting.   In this sense, I think that the relevance ratio analysis works here.

12 Degrees of Responsibility

Nevertheless, we could not avoid imperfectness with regard to the relevance ratio analysis in respect of prior probabilities and the way of collecting data.   So, we should try to optimize the analysis in terms of combining other perspectives. I do not have a definitely excellent idea, but, for the time being, I will propose the next weighted average as a measurement of degrees of *mens rea*.

$$\frac{(\frac{RR(MR)}{1+RR(MR)}) \times q + Pe\,(MR) \times s + Pv\,(MR) \times v}{q + s + v}$$

RR (MR) denotes the value of the relevance ratio on the existence of *mens rea*, so RR (MR) / 1 + RR (MR) signifies the probability of the existence of *mens rea*. Then, Pe (MR) denote the probability given by expert testimony about the existence of *mens rea*, and Pv (MR) the probability given by victims or the bereaved family about the existence of *mens rea*.   Small letters, "q", "s", and "v" are weights, which can be determined according to each particular context.   The reason why I insert as an item the assessment given by victim's side is clear.   It depends upon victims' perception as to whether the offender had *mens rea* or not, as the case of sexual harassment shows.   *Mens rea*, as a matter of fact, is at least partially an interpersonal state, so that if others felt no *mens rea* even if the offender subjectively had that, the *mens rea* does not matter so much.   Then, probability that appears here is basically subjective one although people use empirical data about frequencies or some other evidence in subjectively assessing, so that the probability can be treated by Bayesian method.   Therefore, if new evidence about the behaviour of the offender is offered, we could update the probability through Bayesian confirmation theory.

　　　But, my arguments thus far are only concerned with the uncertainty (γ) that I mentioned above.　In order to propose the reasonable way of measuring degrees of responsibility, we should take uncertainties (α) and (β) into account as well.　I leave the uncertainty (α) up to Bayesians now.　Here I will focus upon the uncertainty (β).　Initially I formulate expected harm.　Expected harm of act*i* at the time *t* is defined like this:

$$
\begin{aligned}
\mathrm{EH}_t(act_i) &= \mathrm{AH}_t(r_1)\mathrm{P}_t(r_1) + \mathrm{AH}_t(r_2)\mathrm{P}_t(r_2) + \cdots\cdots + \mathrm{AH}_t(r_n)\mathrm{P}_t(r_n) \\
&= \sum_{i=1}^{n} \mathrm{AH}_t(r_i)\mathrm{P}_t(r_i)
\end{aligned}
$$

$$(\mathrm{P}_t(r_1) + \mathrm{P}_t(r_2) + \cdots\cdots + \mathrm{P}_t(r_n) = 1)$$

AH denotes "amounts of harm".　I follow such my commonsense that degree of responsibility must be proportional to amounts of harm in principle.　A small letter, "r", signifies possible results caused by the particular "act", hence the sum of probabilities of all possible results must be 1.

　　　Next, using this notion of EH, I give the next formulation by which we can measure degrees of responsibility (DR), although uncertainty (α) (that is expressed REL following the notation of Bovens & Hartman) is still left untouched.

$$
\mathrm{DR}(actj \ \& \ rm) = REL \times \frac{EH\!t(actj)\mathrm{AH}(rm)}{(mah)^2} \times \frac{(\frac{\mathrm{RR(MR)}}{1+\mathrm{RR(MR)}}) \times q + \mathrm{Pe}(\mathrm{MR}) \times s + \mathrm{Pv}(\mathrm{MR}) \times v}{q + r + v}
$$

Three terms correspond to three levels of uncertainties on responsibility for restoring harm.　The notation, "mah", means "the maximal amount of harm" that is virtually the same as the harm caused by hideous murder.　The DR(*actj & rm*) formulated here should be read as "degrees of responsibility in the case that a man did the *actj* at the time *t* and the *resultm* actually occurred under the conditions of some degrees of reliability about our knowledge that the *actj* was actually done by the offender, some degrees of causal liability, and some degrees of *mens rea*. In other words, the offender must take the responsibility to restore the harm in such DR.

　　　Of course, the DR is only concerned with the sentence of responsibility of restoration.　There remains a task as regards the practice of restoration, which is quite realistic, for the practical problems arise with respect to mental disorder or other impairments.　That needs to be discussed in a totally different context,

referring to psychiatric diagnoses.   That is another story.

13 Dangerousness

My arguments thus far discuss the problem of responsibility for past events, i.e., the retrospective phase of responsibility.   However, as I mentioned earlier, there is another phase of responsibility, namely, the prospective phase of responsibility concerning future possibilities.   Applying this phase to the case of offences, we have to face the problem of uncertainty with regard to recidivism or dangerousness of offenders.   Probably, this problem might have a part to do with the task as regards the practice of restoration that I mentioned just above.   In fact, this field has currently achieved remarkable progresses.   A typical outcome of those progresses is what is called "Violence Risk Appraisal Guide" (VRAG) which was proposed by Quinsey, Harris, Rice, and Cormier.   That is the purely statistical methodology based upon detailed data which are divided into 12 terms. These are the next 12 terms.

1. Lived with both biological parents to age 16
    Yes = -2, No = +3
2. Elementary school maladjustment
    No problems   = -1
    Slight or moderate discipline or attendance problems
                                                      = +2
    Severe behaviour or attendance problems
                                                      = +5
3. History of alcohol problems
    Alloy one point for each of the following: alcohol abuse in biological parents, teenage alcohol problem, adult alcohol problem, alcohol involved in a prior offence, alcohol involved in the index offence.
    0 point = -1, 1 or 2 points = 0, 3 points = +1, 4 or 5 points = +2
4. Martial status
    Ever married (for at least 6 months)   = -2
    Never married   = +1
5. Criminal history score for convictions and charges for nonviolent offences prior to the index offence (from the Cormier-Lang system I omit)
    Score of 0   = -2
    Score of 1 or 2   = 0

Score of 3 or above = +3

6. Failure on prior conditional release

No = 0, Yes = +3

7. Age at index offence

$\geqq$ 39 = -5

34-38 = -2

28-33 = -1

27 = 0

$\leqq$ 26 = +2

8. Victim injury (index offence only)

Death = -2

Hospitalized = 0

Treated and released = +1

None or slight = +2

9. Any female victim (for index offence)

Yes = =1, No = +1

10. Meets *DSM-III* criteria for any personality disorder

No = -2, Yes = +3

11. Meets *DSM-III* criteria for schizophrenia

Yes = -3, No = +3

12. Hare Psychology Checklist-Revised Score (PCL-R) (that I omit)

$\leqq$ 4 = -5

5-9 = -3

10-14 = -1

15-24 = 0

25-34 = +4

$\geqq$ 35 = +12

Thus, according to them, probability of violent recidivism can be assessed by the VRAG score that is got by summing up all scores in the above following the next table (Quinsey, Harris, Rice and Cormier, 2006, pp. 283-286).

VRAG category	VRAG score	Probability of recidivism	
		7 yeas	10 years
1	$\leqq$ -22	0.00	0.08

2	-21 to -15	0.08	0.10
3	-14 to -8	0.12	0.24
4	-7 to -1	0.17	0.31
5	0 to +6	0.35	0.48
6	+7 to +13	0.44	0.58
7	+14 to +20	0.55	0.64
8	+21 to +27	0.76	0.82
9	$\geqq$ +28	1.00	1.00

Based upon this, for instance, Governments can rule that offenders assessed from 7 to 9 VRAG categories should be involuntarily committed to a mental hospital for a longer period than assessed only from a retrospective perspective of responsibility, awaiting the time for the offender to recover to be able to conduct restorative activities.

14 Conclusion

This is the amazing system indeed.   However, the fact that this system is empirically precise is one thing; the question of whether we should adopt this system as a legal and moral norm is another thing[6].   This system regards people in an extremely naturalistic manner.   They were natural phenomena like weather.   I understand that such treatment is absolutely necessary as a public policy in a certain context, but I am not 100 % convinced that we should adopt the system.   To be honest, I really hesitate to adopt that.   Problem is how we should face future risks.   Should we achieve the society that is severely controlled but safer, or could we accept a certain amount of risks?   Or is a medium state possible?   My optimistic hope is that we could reduce risks of offences to a certain extent by establishing the fair and objective system regarding the retrospective phase of responsibility as I discussed above, although we cannot be perfectly exempt from risks.   Responsibility, in any case, must be the crucial device that we can rely upon.

---

1 E.g., see Stigler 1999 or McCullagh 1984, chap. 3.   As shown by these studies, the probabilistic approach to historical knowledge is actually popular.

2 This traditional tendency explains why Japan has still retained the system of death penalty despite international pressure for its abolishment.   It appears to be inherently

difficult for Japanese people to accept the abolishment of the death penalty, regardless of whether it is morally correct or not from a contemporary viewpoint.

3 In fact, if we should be fair to Locke, he himself considered both restraint and reparation as two reasons of the system of punishment (Locke, 1960, section 8, p. 272 et al.).   Namely, he included two phases, i.e., criminal trials and civil trials, in a single system of punishment unlike our contemporary institutions.

4 Victor Tadros also put forward the similar viewpoint by saying: 'We cannot simply determine what a person did without engaging in an enquiry about morals and an enquiry about states of mind'.   Nevertheless, that is not to say that there is a separation between legal causation and natural causation.   Rather, according to Tadros, 'the best legal account of causation is broadly along the lines of that used in the natural world' (Tadros, 2005, p. 181).   Obviously, Tadros's argument basically conforms to my line of thoughts.

5 We must add the viewpoint that the notion of unconditional probability is quite dubious.   As to this viewpoint, see Hájek, 2003, pp. 195-201.

6 However, I do not suppose that descriptivity represented in the *is* sentences is completely incompatible with normativity appearing in the *ought to* sentences.   See Ichinose (forthcoming).

References

Acorn, A. 2004. *Compulsory Compassio; A Critique of Restorative Justice*. UBC Press.

Allen, M. J. 1991. *Textbook on Criminal Law*. Blackstone Press Limited.

Bovens, L. and Hartmann, S. 2003. *Bayesian Epistemology*. Oxford University Press.

Duff, A. 2003, 'Restoration and Retribution'. In *Restorative Justice and Criminal Justice*, eds. Andrew von Hirsch, Julian Roberts, Anthony E Bottoms, Kent Roach, and Mara Schiff. Hart Publishing.

Fukushima, A. 2005. *An Introduction to Criminal Psychiatry*. Chuko-shinsho. (in Japanese)

Grisso, T. 1996. 'Pretrial Clinical Evaluation in Criminal Cases'. *Criminal Justice and Behavior*. 23.1: 90-106.

Hájek, A. 2003. 'Conditional Probability Is the Very Guide of Life'. In

Probability Is the Very Guide of Life, 2003, eds. H. E. Kyburg and M. Thalos, Open Court.

Hubbard, R. & Wald, E. 1999. *Exploding the Gene Myth*. Boston: Beacon Press.

Ichinose, M. (forthcoming). 'Normativity, Probability, and Meta-Vagueness,' *Synthese*, forthcoming (DOI: 10.1007/s11229-015-0950-7).

Jeffrey, R. 1992. *Probability and the Art of Judgment*. Cambridge University Press.

Locke, J. 1960. *Two Treatises of Government*. ed. P. Laslett, Cambridge University Press.

Lombroso, C. 2006. *Criminal Man*. Translated by M. Gibson & N. H. Rafter. Duke University Press.

Lyon, T. D. and Koehler, J. J. 1996. 'The Relevance Ratio: Evaluating the Probative Value of Expert testimony in Child Sexual Abuse Cases'. *Cornell Law Review*. 82.1: 43-78.

Marshall, TF. 1999. *Restorative Justice: An Overview*. Home Office.

McCullagh, C.B. 1984. *Justifying Historical Descriptions*. Cambridge University Press.

Nelkin, D. & Lindee, M.S. 2004. *The DNA Mystique: The Gene as a Cultural Icon*. Ann Arbor: The University of Michigan Press.

Nitobe, I. 2002. *Bushido: The Soul of Japan*. Tokyo: Kodansha International Ltd., originally published in 1900. Philadelphia: The Leeds & Biddle Co.

Peay, J. 1997. 'Mentally Disordered Offenders'. In *The Oxford Handbook of Criminology*. Oxford University Press.

Quinsey, V. L., Harris, G. T., Rice, M. E., and Cormier, C. A. 2006. *Violent Offenders: Appraising and Mapping Risk*. Second Edition. American Psychological Association.

Rawls, J. 1955. 'Two Concepts of Rules'. *Philosophical Review*. 64: 3-32. Quoted from *Theories of Ethics*, 1967, ed. P. Foot, Oxford University Press.

Schopp, F. S. 1991. *Automatism, Insanity, and the Psychology of Criminal Responsibility*. Cambridge University Press.

Slobogin, C. 2006. *Minding Justice: Laws that Deprive People with Mental Disability of Life and Liberty*. Harvard University Press.

———— 2007. *Proving the Unprovable: The Role of Law, Science, and Speculation in Adjudicating Culpability and Dangerousness*. Oxford University Press.

Stigler, S.M. 1999. 'John Craig and the Probability of History', in S. M. Stigler, *Statistics on the Table*. Cambridge: Harvard University Press.

Szasz, T. 2003. *The Myth of Mental Illness*. Revided Edition. Perennial.

Tadros, V. 2005. *Criminal Responsibility*. Oxford University Press.

Ten, C. L. 1987. *Crime, Guilt, and Punishment*. Oxford University Press.

Tversky, A. and Kahneman, D. 1973. 'Availability: Aheuristic for judging frequency and probability'. *Cognitive Psychology*. 5: 207-232.

Van Ness, D., Morris, A., and Maxwell, G. 2001. 'Introducing Restorative Justice'. In *Restorative Justice for Juveniles*, eds. A. Morris and G. Maxwell, Hart Publishing.

Walgrave, L. 2001. 'On Restoration and Punishment'". In *Restorative Justice for Juveniles*, eds. A. Morris and G. Maxwell, Hart Publishing.

Weinstein M. C. and Fineberg, H. V. 1980. *Clinical Decision Analysis*. W. B. Saunders Company.