

## Safer self-injury or assisted self-harm?

Kerry Gutridge

Published online: 16 March 2010

© The Author(s) 2010. This article is published with open access at Springerlink.com

**Abstract** Psychiatric patients may try (or express a desire) to injure themselves in hospital in order to cope with overwhelming emotional pain. Some health care practitioners and patients propose allowing a controlled amount of self-injury to occur in inpatient facilities, so as to prevent escalation of distress. Is this approach an example of professional assistance with harm? Or, is the approach more likely to minimise harm, by ensuring safer self-injury? In this article, I argue that health care practitioners who use harm-minimisation can be considered to be helping physical injury to occur, although they do not encourage the act. I consider why there are compelling reasons to believe that a patient who self-injures is not maximally autonomous in relation to that choice. However, I then move onto argue that allowing a degree of self-injury may enable engagement with psychotherapy (enhancing autonomy) and behavioural change. In these circumstances, allowing injury (with precautions) may not be harm, all things considered.

**Keywords** Assistance · Ethics · Harm · Psychiatry · Self-injury

In this article, I discuss the morality of a psychiatric hospital management strategy for self-injury that aims to reduce or limit, rather than stop, the behaviour. By definition, health care practitioners who use the approach allow (a degree of) self-injury to occur within an inpatient environment. To assess whether a practitioner should use this approach, I examine the terms used to describe the strategy to achieve a degree of conceptual clarity. I then consider whether allowing self-injury is right or wrong.

---

K. Gutridge (✉)

Centre for Ethics in Medicine, University of Bristol, 3rd Floor Hampton House, Cotham Hill,  
Bristol BS6 6AU, UK  
e-mail: Kerry.gutridge@bristol.ac.uk

The approach has been labelled ‘safer self-harm’ (Teresa Shaw, 2004, unpublished MA dissertation), ‘harm-minimisation’, or ‘assisted self-harm’ [1–3]. Unreflective adoption of any of these terms is problematic as they carry an implicit evaluation of the method and, by extension, the practitioners involved in its implementation. Safer self-harm and harm-minimisation are synonymous positively-evaluative phrases, implying the approach is good because it decreases the probability or magnitude of a negative event.<sup>1</sup> In contrast, assisted self-harm implies that those who use the approach help, in some way, a negative state of affairs to obtain. In this article, I undertake a detailed analysis of this third phrase, ‘assisted self-harm’, defining and determining the significance of its three constitutive terms in relation to clinical staff and patient descriptions of self-injury.

The article is divided into four subsections. In the first section, I explain what I mean by the term ‘self-injury’ and consider whether assistance in injury can be provided by an act or omission. I conclude that if the mental health care practitioner provides the patient with the means for self-injury or otherwise fails to take steps to prevent injury (e.g., by removing available means), then he or she contributes relevantly to the outcome, assisting with the act by helping it to occur (even if he or she does not encourage the act by initiating, maintaining, increasing, or promoting the behaviour). In the second section, I consider whether providing assistance means the practitioner is accountable for the outcome or whether the practitioner’s responsibility can be abdicated if the individual who injures himself or herself voluntarily undertakes the action, i.e., if the ‘self’ chooses the harm.<sup>2</sup> Through an analysis of autonomy and self injury, I reach the conclusion that the choice to self-injure may not be maximally voluntary,<sup>3</sup> but it also may not be completely non-voluntary. If the choice to self-injure falls somewhere between the two extremes, the practitioner who assists with the action and the patient who undertakes the action both contribute relevantly to the occurrence of the consequent event, and as such, they both bear a degree of responsibility for the outcome. In the third section, I assess whether the practitioner is partially responsible for a behaviour that is *immoral*. In doing so, I consider whether self-injury is necessarily harm and, if so, whether such harm is *bad*. I conclude that allowing self-injury does not cause *relative harm* (harm measured against an individual’s current ‘norm’, i.e., *not* harm as subjectively determined by the individual). Self-injury nonetheless does maintain the individual in a state of *objective harm* (a harmed-condition measurable independent of the subject’s perceptions, conceptions, *and* current state). In the final section, I suggest that the health care practitioner has an obligation to remove the individual from this harmed condition. This obligation entails the adoption of a successful strategy to enact change. Paradoxically, removing an individual from a harmed condition may depend on allowing self-injury to continue temporarily until certain autonomy enhancing psychological conditions can be reached that enable

<sup>1</sup> Harm is a ‘weak normative concept, carrying a presumption of evaluative negativity’ [4, p. 1034].

<sup>2</sup> For a detailed discussion of voluntariness and body modification in non-psychiatric cases, see [5].

<sup>3</sup> An individual may not be able to control their own choices maximally until defects in reasoning or information are remedied; see [6, pp. 196–213].

successful behavioural change (perhaps realizable through psychotherapeutic intervention).

### Assisting injury?

‘Self-injury’ is an umbrella term that can be used to describe many different behaviours. According to the National Patient Safety Agency, self-injurious behaviours encountered in an inpatient environment include self-cutting, burning, poisoning, and strangulation (using a ligature) [7, p. 42].<sup>4</sup> Following from research conducted by the Bristol Crisis Service for Women, in this article, the term ‘self-injury’ will be used to describe injuring behaviours carried out to overcome unpleasant affective mental states. These states include subjective feelings such as overwhelming emotional pain, self-hatred, anger, anxiety, neediness, or unreality [8, p. 13]. The injuring behaviours are used with the intention of helping the individual cope with these distressing feelings and as such can be distinguished from other self-injurious acts undertaken with suicidal intent or to cause a positively evaluated physical alteration.<sup>5</sup>

I will limit the current discussion of self-injuring behaviours to those which result in a loss of bodily integrity and can *relatively* easily be subject to harm-minimisation strategies, i.e., self-cutting and self-burning. In this context, a loss of bodily integrity should be understood as an alteration to the body’s wholeness which is likely to cause disfigurement (such as scarring) though not necessarily impairment (diminished physiological function). It is important to consider self-cutting in particular because it is the most common type of self-injurious behaviour [8, p. 6].

There are different ways in which a health care practitioner can be associated with self-injury in an inpatient environment. I will focus my attention on two scenarios in which self-injury might be permitted (subject to certain restrictions):<sup>6</sup>

1. The care team provides the patient with a hospital-owned cutting implement on request for private self-cutting.
2. The care team allows a patient, on request, to keep a self-acquired cutting implement (e.g., glass) for private self-cutting.

In both circumstances the implement is sterilized to reduce the chance of infection. It is established that the patient in question self-injures regularly (e.g., has visible scarring) and knows how to dress wounds, and the patient is then provided with the necessary materials to do so. To reduce the risk of impairment or accidental death, the patient agrees to speak with the care team if the team has any concerns about the severity of the injuries. The care team is aware of the early warning signs of the

<sup>4</sup> This description is not intended to be exhaustive.

<sup>5</sup> Cf. [5].

<sup>6</sup> These scenarios have been chosen as examples of two approaches that have been suggested for practice. The different approaches were described by participants in qualitative interviews that the author is currently conducting with health care practitioners and patients.

self-injury escalating beyond an agreed level and has the patient's permission to intervene if this occurs.<sup>7</sup>

According to Schoene-Seifert [4, p. 1034], injury<sup>8</sup> is 'ethically relevant only if it occurs or persists in consequence to human agency, be it by action or omission, from intention or negligence'.<sup>9</sup> What needs to be determined in this article is (1) whether the practitioners' behaviour (action/omission) is a means to the consequent injury (are they agents who are relevantly related to the outcome) and (2) whether their behaviour is undertaken because the practitioner is striving towards an appropriate goal (rather than acting carelessly or recklessly). In this section, I address these points by considering whether the care team is assisting the patient in either or both of the scenarios. (At this juncture I will not consider the role of the patient in the sequence of events.<sup>10</sup>)

Helpful examples of the different types of assistance that may exist can be found in criminal law; e.g., 'a person who aids, abets, counsels or procures the suicide of another' [10] essentially assists in that act. These terms mean, respectively, that individuals are judged to provide assistance if they help, advise, encourage, *or* bring about an event (even though another commits the last act).<sup>11</sup> I will use these criteria as a list of *alternative* meanings for the term 'assistance' against which the scenarios can be assessed. I will start by examining whether the health care practitioners help obtain a state of affairs (or bring it about). I will then discuss whether they encourage the patient to undertake the act. Although it is an interesting question, for brevity, I will leave aside analysis of the relevance of advice.<sup>12</sup>

In each scenario, do the practitioners help self-injury to occur or persist? All things being equal, the distinction between scenario 1 and 2 is between doing (action) and allowing (inaction or omission).<sup>13</sup> In scenario 1, the health care practitioner is active in relation to the harm (by providing the means), and in scenario 2, the practitioner is passive (by letting the patient perform the action, using an implement they already possess, without intervention) [12, p. 9]. It

<sup>7</sup> These conditions are the author's interpretation of a harm-minimisation care plan for self-injury used within the Sexual Abuse Service, South Staffordshire and Shropshire Healthcare NHS Foundation Trust.

<sup>8</sup> For consistency I have substituted the word harm with the word injury.

<sup>9</sup> Negligence can mean '(1) intentionally imposing unreasonable risks of harm (advertent negligence or recklessness) and (2) unintentionally but carelessly imposing risks of harm (inadvertent negligence)' [9, p. 154].

<sup>10</sup> I will pick up this question in the next section.

<sup>11</sup> My understanding of the different concepts of 'assistance' is based on common usage rather than legal interpretation. The arguments advanced hold regardless of whether these definitions follow legal developments.

<sup>12</sup> As a point of interest, NICE guidelines on self-harm suggest that it is appropriate for a practitioner to 'discuss harm minimisation issues and techniques' when an individual is likely to repeat self-injury, but the guidelines are silent on the appropriateness of allowing self-injury to occur in health care facilities [11].

<sup>13</sup> I will assume for current purposes that the provision of the implement by the team does not provide any additional opportunity for dialogue between the team and the service user which might affect their drive to self-injure.

continues to perplex philosophers<sup>14</sup> whether, ‘in certain contexts, failure to perform an act, with certain foreseen consequences of that failure, is morally less bad than to perform a different act which has identical foreseen consequences’ [14, p. 92]. In this article I am interested in whether *both* scenarios help an event to occur (i.e., bringing the event about). This latter question can be determined while leaving aside the additional question of whether helping by action or inaction is *worse*.

A relevant question when considering whether or not help occurs is whether the health care practitioner causes the injury, as causation and responsibility are linked. Put another way, an appreciation of an agent’s role in the chain of events leading to injury is necessary before one can assess whether it is reasonable to hold the agent accountable for the outcome. Scenario 1, in which the health care practitioner *acts*, is a relatively unproblematic example of an agent causing the outcome.<sup>15</sup> The provision of a hospital-owned cutting implement is an act that leads to an injury. It is more controversial whether scenario 2 describes a causal sequence, so the following discussion will focus on this scenario, drawing on McGrath’s review of causation by omission.

According to McGrath there is a question of whether omissions are causes: either there is no causation by omission or there is far more than common sense suggests [15, p. 126]. Arguments for the first possibility are as follows: there is no causation by omission because causal relations are events (and omissions are not events), or because cause and effect involves a physical connection (such as a transfer of energy) which is missing when an individual fails to act [15, p. 127]. These objections to omissions as causes focus exclusively on cause as a natural phenomena (cause and effect). However, there are different causal idioms that link conduct to subsequent events, e.g., cause and effect, cause and consequence, and cause and result [16, p. 172]. According to Feinberg, it is more natural to accept that omissions have consequences rather than effects [16, p. 173].

If one focuses on the relationship between cause and consequence, for an omission to be a cause, it is necessary to accept the following counterfactual proposal: an omission causes an event if and only if the event occurs when the omission obtains, and the event does not occur when a counterfactual act obtains (the counterfactual act *prevents* the event) [15, p. 132]. In the self-injury example a counterfactual act, such as removing the cutting implement, would prevent the particular injury in question from occurring.<sup>16</sup> Unfortunately, this explanation does not provide necessary and sufficient conditions for omissions as causes because there are many counterfactual actions that could prevent the event. The question remains how one determines that the particular omission is *the* cause; why failure of the specific agent to perform the specific preventative act is relevant.

Jonathan Bennett and David Lewis accept that omissions can be causes and, to avoid over-extension (the second possibility in McGrath’s dilemma), that *relevant*

<sup>14</sup> E.g., [13].

<sup>15</sup> If one assumes at this juncture that the service user’s action does not break the chain of causation.

<sup>16</sup> One could argue that there is still a probability that injury can occur by other means and that the likelihood of injury may in fact increase (e.g., because of increased distress), but these empirical facts are not relevant to the particular point being made.

causal omissions can be pragmatically limited to those that are salient (they ‘stand out as notably significant’) [17, p. 133; 18, p. 196]. McGrath suggests that normative conditions should be added to determine significance. She proposes that relevance is measured by what is *normal*, what an agent is *supposed to* do in a given situation: an agent relevantly causes an outcome if he or she omits to undertake an action that is its ‘normal would be preventer’ according to some standard determined by the actual situation [15, pp. 136–141]. In Feinberg’s terms, an omission is a relevant cause of the consequence if the agent was required to act (e.g., because of a moral obligation or a special duty in virtue of his or her profession), was in a position to act, and did not.

This understanding of relevant causes can be used to assess whether the mental health care practitioners help with the patient’s self-injury. I propose that their omission is a salient cause, because the health care practitioner has a duty of care in relation to the patient, and the practitioner’s inaction leads to a significant event not being prevented. The event is significant because the practitioner purposefully decides to provide an opportunity for injury by failing to undertake common precautions, such as removal of sharp objects and observation. The normative conditions extend the assessment from a consideration of whether the practitioner helped, to whether the practitioner *should* have helped, assessing moral culpability for the omission as well as causal responsibility. This assessment can be addressed more successfully, in the given context, within the analysis of harm.

Once one accepts that the health care practitioner’s omission is a causal factor in the consequent injury, and this means they help the event obtain (bringing it about), a question remains of whether they encourage the injury. To encourage the act, the practitioner needs to stimulate or spur the patient onto self-injury. I propose that encouragement needs to be measured against the explicit intentions of the practitioners when they allow self-injury to occur. Practitioners who advocate the use of the approach defend the strategy as a means to prevent escalation of injury, and, rather than simply allowing any injury, they suggest potential distraction techniques to avoid injury and promote safer injuring techniques.<sup>17</sup> These facts strongly suggest that the intention is neither to increase the incidence or degree of injury nor to maintain the injury beyond the time that it helps a patient to cope with distress. The aim that guides the (in)action is injury reduction rather than stimulation. If a practitioner acted in a way that encouraged injury, one can assume that the behaviour was reckless or careless in relation to the injury. A possible objection to my position is that help of itself provides tacit encouragement for the action. My response is that it is logically possible to help an event obtain while simultaneously holding a desire that said event not eventuate. The health care practitioner can communicate to the patient this desire that the injury not occur and that the patient’s behaviour is tolerated as a temporary coping strategy.

In conclusion, the practitioner assists in the self-injury, but this assistance is limited to helping someone undertake the behaviour, and does *not* include providing

---

<sup>17</sup> Determined at interview with practitioners (unpublished) and explicitly stated in [19].

encouragement intended to initiate, maintain, increase, or promote the behaviour.<sup>18</sup> An omission as well as an act can help injury to occur if the causal idiom being used is cause and consequence, not cause and effect. Although the health care practitioner is an agent who stands in a relevant relationship to the injury, the patient also plays a significant role in the injury's occurrence—the patient is considering *self*-injury. Bearing this in mind, I will now address the patient's role and its relevance to the practitioner's accountability.

### Self-chosen injury?

In the previous section, I temporarily put aside the relevance of the patient personally undertaking the final act—the *self* causing the injury. In this part of the article, I will address whether the patient *voluntarily* chooses to injure himself or herself and, if so, whether this means that the patient is sufficiently accountable that no responsibility need be attributed to the practitioner. According to Feinberg, voluntariness makes a difference to whether an action is judged to be wrong: 'One class of harms (in the sense of set-back interests) must certainly be excluded from those that are properly called wrongs... These include harms voluntarily inflicted by the actor upon himself' [16, p. 35]. I will now assess whether the patient's self-injuring actions are voluntary, i.e., whether the injury is a consequence of his or her volition (i.e., whether it is self-determined).

One of the conditions used to determine whether an individual is acting freely is the possession of certain mental capacities, such as the ability to act on the basis of one's own, self-chosen, desires or plans; that is, having 'the capacity for intentional action' [9, p. 100]. According to O'Neill, accounts of individual autonomy can be viewed as descriptions of the personal traits, capacities, or competencies necessary for successful functioning, i.e., of psychological abilities which individuals need (to a modest<sup>19</sup> degree) to ensure that they have the resolve and self-confidence to fulfil their obligations [20, p. 96]. These psychological abilities and their relevance for self-injury can be explored via an analysis of available accounts of procedural and substantive autonomy.

Procedural accounts can be divided into endorsement and historical accounts [21, p. 13]. Frankfurt's writing on freedom of the will is recognised as a fundamental endorsement account.<sup>20,21</sup> Frankfurt posits that people can form second order desires about their first order desires, where a first order desire is a desire to act and a second order desire is an evaluation of the first order desire (e.g., people can want to want an action) [24, pp. 64–66]. Frankfurt distinguishes between second order desires and second order volitions, which are second order desires that a person

<sup>18</sup> This does not entail inferring that encouragement goes beyond assistance but only that assistance can be provided with or without encouragement.

<sup>19</sup> Intuitively, I disagree with O'Neill's assertion that these abilities are only needed to a 'modest' degree.

<sup>20</sup> This is in spite of the fact that the account was originally intended to be a conceptualisation of the necessary features for a person. For discussion of Frankfurt in relation to autonomy, see [22, pp. 6–7].

<sup>21</sup> First order and second order endorsement conditions are also evident in [23, p. 20].

wants to be his will [24, p. 67]. He argues that an unresolved conflict between second order desires will prevent second order volitions, which ‘removes [a person] from his will so that his will operates without his participation’ [24, p. 71]. This destruction of the authentic will is pertinent for an examination of self-harm. If it can be shown that individuals who self-harm are not acting on a second order volition, then it could indicate that they are not acting (maximally) voluntarily in their self-harming actions. For example, the person who self-injures may simultaneously want to want to injure because of self-hatred, and wish not to want to injure, realising it is only a transiently effective solution for emotional pain.

There are objections associated with relying solely on an endorsement account to determine procedural failures in an individual’s motivational structure.<sup>22</sup> According to many authors, a hierarchical, critically reflective desire structure encounters an unavoidable problem of regress [21, p. 13]. Although some maintain that an infinite regress is circumvented because of the natural limitations of the human mind [23, p. 19], it is difficult to establish a definitive example of an action that demonstrates an unavoidably authentic desire structure. It might be accepted that the problem of regress does not fatally undermine the application of endorsement accounts if the aim is to establish cases where unresolved conflict reveals an inauthentic desire. However, it can be questioned whether an endorsement account is sufficient for this purpose because the account has also been charged with a (possibly intractable) problem of perfectionism. Endorsement may set too high a standard for autonomy if the account rules out the possibility of having an autonomous desire of which one doesn’t approve [25, p. 113]. Perhaps ‘we need a fine[r]-grained account of the kinds of inconsistencies among desires that can plausibly be said to be autonomy undermining’ [26, p. 104].

Another approach for procedural accounts is to establish historical criteria for autonomous desires.<sup>23</sup> Christman distinguishes his historical criterion from hierarchical structures by claiming that the key feature of autonomous desire formation is that the agent ‘did not resist the development of D [desire] when attending to this process of development, or P [the person] would not have resisted that development had P attended to the process’ [27, p. 347]. So, on this account, a desire is autonomous if it is chosen by the person on reflection without actual or counterfactual resistance. In relation to self-injury, the desire to injure would be judged to be autonomous if the individual would not have resisted the development of the desire to injure if he or she had attended to the process of the formation of the desire. The desire to self-injure is often formed as a reaction to abuse or neglect [8, p. 3]. One can postulate that a behaviour motivated by these events may be resisted on reflection.

Like endorsement accounts, historical accounts are also subject to objections. An objection that is levelled against both accounts is that they do not address the effects of oppressive socialisation on the development of desires [21, p. 13]. All accounts

<sup>22</sup> These objections hold in addition to any potential objections that could be posited based on the difficulty of assessing people’s inner worlds via their accounts of action or observed behaviour.

<sup>23</sup> According to [21], Dworkin also adds an historical dimension to his endorsement account by stating that autonomous agents can reflect on the acquisition of their desires and ‘attempt to change these in the light of higher order preferences or values’.



that rely on procedural conditions for the judgment of autonomy, as phenomenal accounts, position the judgments within the social and physical forces of nature, and as such, all reflective capacities used by a procedurally autonomous person may have been subject to oppressive social forces. It is open to debate whether these forces necessarily undermine autonomy or whether one can successfully step back and reflect on the social forces which inform one's development.<sup>24</sup>

If one accepts that, in spite of certain limitations, procedural accounts outline *necessary* features of individual autonomy, perhaps through the use of a principle of charity,<sup>25</sup> the question remains whether these features are *sufficient* to judge whether an action is autonomous. With respect to the question at hand, do these features describe all the relevant psychological characteristics that can be used to determine whether an individual who self-injures is acting voluntarily? If one turns to substantive accounts of autonomy, it is apparent that procedural accounts may not be sufficient for this purpose.

One aspect of individual psychology that procedural accounts neglect is the non-cognitive, those affective qualities that can be autonomy enhancing or undermining. One relevant substantive account of autonomy which addresses affect is provided by Benson in his examination of free agency and self-worth [29]. On this account, 'free agents must have a certain sense of their own worthiness to act, or of their status as agents, which is not guaranteed by their abilities to act freely by reflectively authorizing their wills and their action' [29, p. 650]. If individuals cease to trust themselves to govern their actions competently (e.g., as the result of psychiatric diagnoses that estrange them from their former sense of themselves as worthwhile agents or because of persistent self-critical mental states), then their ability to act autonomously can be affected [29, p. 657]. Benson's theory is substantive because it claims that 'holding certain attitudes, would preclude the necessary sense of worth (psychologically, if not logically)' [29, p. 664].

If this substantive account is used to supplement the procedural accounts, then agents need to meet criteria for endorsement, have freely-formed desires relative to the self-injuring action, *and* feel a sufficient degree of self-worth to be judged to possess enough individual autonomy to act voluntarily. The self-worth condition may be criticised as unnecessary by those who believe that insufficient self-worth undermines the ability to undertake cognitive reflection; however, it is logically possible that the cognitive abilities of an individual remain unchanged even if the individual mistrusts these abilities.

The intended self-injury of an individual who has failed to meet certain thresholds of endorsement, free-formation of desires, and self-worth may be 'substantially non-voluntary' [30, p. 8]. This is not to say that the choice is 'near the bottom of the scale of voluntariness' [30, p. 8], but just that it is possible for the choice to be *more* voluntary. Given this eventuality it seems that it would be inappropriate for the patient to be held *fully* responsible for the injury (even if they reach certain procedural standards such as legal capacity). However, if the patient's

---

<sup>24</sup> See [28].

<sup>25</sup> A principle of charity, suggested in [27, p. 350], is an acknowledgement that an account is the best, strongest, possible explanation given our current understanding (in spite of its limitations).

actions are not minimally voluntary and the practitioner's actions are voluntary, both the practitioner *and* patient cause the injury to some degree. This means they can both be held partially responsible for the outcome. The question remains whether they are both responsible for *harm*.

### Is injury harm?

It is uncontroversial to state that the outcome of the management strategy under consideration is physical injury, which is what is intended<sup>26</sup> by both parties and desired by the patient. The important consideration is whether allowing injury is necessarily an example of allowing harm. An important distinction can be made between proximate injury and harm—on a narrow conception of injury, a body-altering injury (physical damage) may not count as harm, all things considered [31, p. 279].

Feinberg specifies that harm should be conceived as 'the thwarting, setting back, or defeating of an interest' that is *also* a wrong [16, pp. 33–36]. An interest is defined as something one has a stake in, a component of one's well-being, and a wrong occurs when 'a person's indefensible (unjustifiable and inexcusable) conduct violates another's right' [16, p. 35]. Alternatively, the word harm can be used to indicate doing someone wrong, in the sense of violating someone's rights *or* having an adverse effect on his or her interests [9, p. 152]. For present purposes, I accept the latter distinction and put aside the question of rights, focusing instead on how self-injury relates to a person's interests. This approach means that a practitioner can be seen as having a *prima facie*, rather than an absolute, obligation not to harm others or impose indefensible risks of harm [9, p. 153].

According to Schoene-Seifert, there is 'an asymmetry between harm and benefit in the sense that harm pertains exclusively to the basics of well-being' [4, p. 1034]. Feinberg calls these basics the welfare interests, claiming that they are 'minimal but nonultimate goods' that are necessary as a means to achieving one's ulterior interests [16, p. 37]. The specific welfare interests are contestable, but perhaps few would argue that there are interests in an absence of physical suffering or malfunction and emotional suffering or dysfunction.<sup>27</sup> Feinberg includes the following relevant interests amongst the welfare interests: the integrity and normal functioning of one's body; the absence of absorbing pain and suffering or grotesque disfigurement; emotional stability; and a certain amount of freedom from interference and coercion [16, p. 37].

Feinberg, following Rescher, states that there are few if any trade-offs operative between interests as they 'make a chain no stronger than the weakest link' [16, p. 37]. However, he concedes that some individuals will undertake 'protective diversification' whereby they arrange their interests 'so that what harms some of

<sup>26</sup> In relation to the practitioner's intentions, I accept that it is open to debate whether they intend or merely foresee the outcome. In the particular circumstances under consideration the practitioner knows the injury is virtually certain to occur and accepts that allowing the injury serves a purpose that is (temporarily) valuable (as assessed by the patient).

<sup>27</sup> See Schramme in this special issue for an analysis of mental disorder as dysfunction.

them necessarily helps others...the interests so arranged will not sink below a certain level, even at the cost of an equally strong expectation that they will not substantially rise above that level either' [16, p. 41]. The problem with this arrangement is that in protecting the interests from falling below a certain level, the individual is necessarily damaging part of the chain, making it difficult to advance his or her net interests sufficiently to realise ulterior interests. As such, protective diversification may prevent incapacitating suffering, but it is not a sufficient condition for well-being [16, p. 41]. As I described earlier, some patients who self-injure cause physical disfigurement (and pain) to help them cope with severe emotional suffering. This can be viewed as an example of protective diversification.

If one assumes that these welfare interests are commensurable, protective diversification may not cause a net reduction in welfare, although it does prevent welfare advancement. So is the prevention of welfare enhancement harm? Harm can be viewed as a relative concept: 'whether one is harmed by an event is determined by reference to where he was before, and whether his position has improved or regressed' [16, p. 54]. Once harm is understood in this way, and if harm-minimisation strategies can successfully prevent injuries that cause impairment of function, allowing self-injury does not appear to cause *relative* harm. Moreover, if the individual is already disfigured (already has scarring) it is possible that this interest is only affected to a slight degree, while the emotional stability interest is significantly enhanced (albeit temporarily).<sup>28</sup>

However, there is another non-relative understanding of harm, according to which an individual can be judged against objective criteria as being in a harmed condition. On this view, the baseline against which people are judged is what is 'normal' for flourishing agents rather than for the specific individual. If one considers the effects of self-injuring on advancement of interests it is possible that the individual, even absent relative harm, remains at a level of welfare too low to successfully achieve ulterior interests. This is supported by the probability that while self-injuring offers a temporary relief from emotional pain, it neither addresses the reasons for emotional vulnerability nor provides an opportunity for emotional repair. Consequently, if the individual continues indefinitely to use self-injury as the *only* means of coping, the individual may never reach a 'normal' level of welfare. If this is the case, the final remaining question is whether allowing an activity that does not produce a relative harm but does leave an individual in a harmed condition is ever justifiable.

### Enhancing the self to prevent harm

Health care practitioners have a *prima facie* obligation to remove the self-injuring individual from a harmed condition to improve their welfare. What needs to be established is whether prohibition of self-injury or allowing injury is the best means to remove an individual from a harmed condition. Effective behavioural change, in part, involves helping the individual to develop their psychological capacities. As

<sup>28</sup> One act of self-injury relieves but does not end emotional suffering.

demonstrated earlier, in relation to self-injury, this involves helping the individual reflect on his or her beliefs, challenge any inconsistencies, and develop self-worth. Different authors have linked the development of judgment and self-worth to being afforded the freedom to act. According to Feinberg, 'If adults are treated [paternalistically with the state *in loco parentis*] they will in time come to be like children. Deprived of the right to choose for themselves, they will lose the power of rational judgment and decision. Even children, after a certain point, had better not be 'treated as children,' else they will never acquire the outlook and capability of responsible adults' [32, p. 24]. And in the words of Dworkin, 'our self esteem and sense of worth are bound up with the right to determine what shall be done to and with our bodies and minds' [23, p. 95]. It appears that prohibition of self-injury does not, in and of itself, assist the individual to develop the capacity to choose reasonably and with confidence.

Some may object that firm prohibitive boundaries are a necessary preliminary measure before persuasion can be used to effect change, i.e., that stopping self-injury is a necessary pre-condition for effective therapy. In response to this, I suggest that to accept boundaries without resistance, an individual may need to be at an advanced stage of psychological recovery.<sup>29</sup> Practitioners and patients report that prohibiting self-injury in inpatients can create an escalation of distress and can lead the patient to disengage from any psychotherapeutic intervention or to experience a disintegration of relations with their health care team [Chris Holley, personal communication]. If these claims are correct, then prohibition may prevent the necessary conditions for change from being achieved. Instead, respecting the patient's choice might offer a better opportunity for therapeutic dialogue to develop or evolve and, in turn, for alternative coping strategies to be explored. If so, allowing a (degree of) self-injury is a temporary measure that ultimately removes the patient from a harmed condition by providing an opportunity for the development of sufficient autonomy to protect physical *and* emotional interests.

## Conclusion and future considerations

If the proposed conceptual understanding of assisted self-injury and the empirical claims about behavioural change are accepted, one can conclude that health care practitioners are justified in allowing self-injury *if* they ensure relative harm does not occur *and* they work towards enhancing the patient's capacity to undertake voluntary choices. It can be predicted that with psychological recovery the desire to self-injure will diminish and, as such, objective harm will be reduced. Given these claims allowing proximate injury may not lead to harm, all things considered. Injury is permitted in the short-term to enable the patient to develop the capacities necessary for long-term advancement of interests. Although practitioners help to bring about a degree of injury, they do not encourage the event but suggest ways to minimise the degree of physical damage. As such, they assist with a patient's safer self-injury.

---

<sup>29</sup> This may be precluded by the fact that the person is sufficiently unwell to merit hospitalisation.

An important caveat needs to be acknowledged and addressed in future research. In the approach described in this article, injury is tolerated as a temporary means to engagement and therapeutic change. This does not mean that the practitioner is endorsing self-injury as an uncontroversial choice. When a practitioner begins psychotherapy with the patient, it is important that this fact is explicitly addressed. This is relevant because it is likely that psychological recovery rests in part on the patient developing an understanding that non-neglectful and non-abusive relationships require respect for another person's mind and body. It is important that the health care practitioner's temporary acceptance of injury, until other coping strategies are developed, is not interpreted by the patient as contradictory to this therapeutic aim.

**Acknowledgements** An earlier version of this article was presented at the 31st Congress on Law and Mental Health, IALMH, June 28–July 3, 2009. I am grateful to participants at that conference for their questions and comments. Thanks to my academic supervisors Dr. Richard Huxtable and Dr. Jonathan Evans, and staff and students at the Centre for Ethics in Medicine (Bristol), particularly Alex Calladine and Heather Bradshaw, for their helpful comments. Thanks are also due to members of Avon and Wiltshire Mental Health Partnership NHS Trust, Bristol Crisis Service for Women, the Sexual Abuse Services South Staffordshire, Shropshire Healthcare NHS Foundation Trust, Anne-Laure Donskoy, and Hanna Pickard for helping me develop my ideas and arguments. This work was supported by the Wellcome Trust [083278/Z/07/Z].

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Non-commercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

1. National Self-Harm Network. 2000. *Cutting the risk: Self-harm, self-care and risk reduction*. London: NSHN.
2. Pengelly, Nicky, Barry Ford, Paul Blenkiron, and Steve Reill. 2008. Harm minimisation after repeated self-harm: Development of a trust handbook. *Psychiatric Bulletin* 32: 60–63.
3. Hewitt, David. 2004. Assisted self-harm: Some legal considerations. In *New approaches to preventing suicide: A manual for practitioners*, ed. David Duffy, and Tony Ryan, 148–166. London: Jessica Kingsley Publishers.
4. Schoene-Seifert, Bettina. 2004. Harm. In *Encyclopedia of bioethics*, ed. Stephen Post, 1033–1038. New York: Thompson Gate.
5. Schramme, Thomas. 2008. Should we prevent non-therapeutic mutilation and extreme body-modification? *Bioethics* 22: 8–15.
6. Harris, John. 1985. *The value of life: An introduction to medical ethics*. London: Routledge.
7. Scobie, Sarah, Edana Minghella, Colin Dale, Richard Thomson, Paul Lelliott, and Kathryn Hill. 2006. *With safety in mind: Mental health services and patient safety. Patient safety observatory report 2*. National Patient Safety Agency.
8. Bristol Crisis Service for Women. 1995. *Women and self-injury: A survey of 76 women: A report on women's experience of self-injury and their views on service provision*. Mental Health Foundation.
9. Beauchamp, Tom, and James Childress. 2009. *Principles of biomedical ethics*, 6th ed. New York: Oxford University Press.
10. Suicide Act. 1961. s. 2 (1). 9&10 Eliz. 2, c.60. London: HMSO.
11. National Institute for Clinical Excellence. 2004. *Self-harm: The short-term physical and psychological management and secondary prevention of self-harm in primary and secondary care*. London: NICE.

12. Huxtable, Richard. 2007. *Euthanasia, ethics and the law: From conflict to compromise*. Oxon: Routledge-Cavendish.
13. Hope, Tony. 2000. Acts and omissions revisited. *Journal of Medical Ethics* 26: 227–228.
14. Glover, Jonathan. 1977. *Causing death and saving lives*. London: Penguin Books.
15. McGrath, Sarah. 2005. Causation by omission: A dilemma. *Philosophical Studies* 123: 125–148.
16. Feinberg, Joel. 1984. *Harm to others*. New York: Oxford University Press.
17. Bennett, Jonathan. 1995. *The act itself*. New York: Oxford University Press.
18. Lewis, David. 2000. Causation as influence. *Journal of Philosophy* 97: 182–197.
19. Holley, Chris. 2007. *Clinical guidelines for the management of self injurious behaviours: Adult and older peoples' mental health services*. South Staffordshire and Shropshire Healthcare NHS Foundation Trust.
20. O'Neil, Onora. 2002. *Autonomy and trust in bioethics*. Cambridge: Cambridge University Press.
21. Stoljar, Natalie. 2007. Theories of autonomy. In *Principles of health care ethics*, 2nd ed, ed. Richard Ashcroft, Angus Dawson, Heather Draper, and John McMillan, 11–17. Chichester: John Wiley & Sons Ltd.
22. Christman, John. 1989. Introduction. In *The inner citadel: Essays in individual autonomy*, ed. John Christman, 3–23. New York: Oxford University Press.
23. Gerald, Dworkin. 1988. *The theory and practice of autonomy*. Cambridge: Cambridge University Press.
24. Harry, Frankfurt. 1989. Freedom of the will and the concept of a person. In *The inner citadel: Essays in individual autonomy*, ed. John Christman, 63–76. New York: Oxford University Press.
25. Christman, John. 1988. Constructing the inner citadel: Recent work on the concept of autonomy. *Ethics* 99 (1): 109–124.
26. Stoljar, Nathalie. 2000. Autonomy and the feminist intuition. In *Relational autonomy*, ed. Catriona Mackenzie, and Natalie Stoljar, 94–111. New York: Oxford University Press.
27. Christman, John. 1991. Liberalism and individual positive liberty. *Ethics* 101 (2): 343–359.
28. Mackenzie, Catriona, and Natalie Stoljar. 2000. Autonomy reconfigured. In *Relational autonomy*, ed. Catriona Mackenzie, and Natalie Stoljar, 3–31. New York: Oxford University Press.
29. Benson, Paul. 1994. Free agency and self worth. *The Journal of Philosophy* 91 (12): 650–668.
30. Feinberg, Joel. 1987. Legal paternalism. In *Paternalism*, ed. Rolf Sartorius, 3–18. Minneapolis: University of Minnesota Press.
31. Sheldon, Sally, and Stephen Wilkinson. 1998. Female genital mutilation and cosmetic surgery: Regulating non-therapeutic body modification. *Bioethics* 12 (4): 263–285.
32. Feinberg, Joel. 1986. *Harm to self*. New York: Oxford University Press.