

# Forgetting ourselves: epistemic costs and ethical concerns in mindfulness exercises

Sahanika Ratnayake,<sup>1</sup> David Merry<sup>2</sup>

<sup>1</sup>School of Philosophy,  
Australian National University,  
Canberra, Australia  
<sup>2</sup>Institute for Philosophy,  
Humboldt-Universität zu Berlin,  
Berlin, Germany

## Correspondence to

David Merry, Institut für  
Philosophie, Humboldt-  
Universität zu Berlin, Unter  
den Linden 6, 10099 Berlin,  
Germany;  
[david.merry@gmail.com](mailto:david.merry@gmail.com)

Received 10 February 2017  
Revised 8 December 2017  
Accepted 5 February 2018

## ABSTRACT

Mindfulness exercises are presented as being compatible with almost any spiritual, religious or philosophical beliefs. In this paper, we argue that they in fact involve imagining and conceptualising rather striking and controversial claims about the self, and the self's relationship to thoughts and feelings. For this reason, practising mindfulness exercises is likely to be in tension with many people's core beliefs and values, a tension that should be treated as a downside of therapeutic interventions involving mindfulness exercises, not unlike a side effect. Clients ought to be informed of these metaphysical aspects of the exercises, and mental health providers ought to take them into account in assessing which course of treatment to recommend. Given these concerns, the casual way in which mindfulness exercises are presently distributed by mental health providers to the general public is inappropriate.

## INTRODUCTION

Mindfulness exercises are widely used in psychotherapy. They form part of mindfulness-based cognitive therapy, dialectical behavioural therapy (DBT), acceptance and commitment therapy (ACT) and Jon Kabat-Zinn's mindfulness-based stress reduction (MBSR) programme. Mindfulness is often described as a particular mode of engaging with the present moment, often involving a lack of judgement, evaluation or particular kinds of conceptualisation.<sup>1</sup> These exercises are being widely distributed through self-help books, apps and websites. As well as being recommended for serious psychological problems, they are marketed to people for stress, workplace resilience, becoming happier, or indeed, for not becoming happier but living with meaning.<sup>2</sup> Presently, mindfulness is taught in schools, universities and workplaces.

Although mindfulness exercises originated in close connection to various philosophical and spiritual traditions, their proponents claim that they are universal or non-denominational.<sup>1</sup> The thought is that, in the context of contemporary psychotherapy, mindfulness exercises lose whatever spiritual or philosophical meaning they may have had, in the same way that a yoga pose might in the context of physiotherapy.<sup>ii</sup>

We aim to show that this portrayal of mindfulness exercises is problematically misleading. We will argue that, as they occur in contemporary psychotherapy, mindfulness exercises are tightly connected with philosophical claims about the self incompatible with a range of philosophical, religious and spiritual beliefs. As such, mental health practitioners and their clients ought to take this into account when considering their use.

We do not wish to deny the benefits of mindfulness exercises, nor that they are often overall helpful. Instead we wish to point out that, as with many beneficial medical interventions, mindfulness exercises have potentially negative aspects which need to be taken seriously.

In the first part of the paper, we show how, in the context of contemporary psychotherapy, certain mindfulness exercises are metaphysically loaded with philosophical claims about the self—namely claims that the self is separate from one's thoughts, feelings, values and beliefs.<sup>iii</sup> Some mindfulness exercises, known as cognitive defusion exercises, were developed by contemporary psychologists quite explicitly to be loaded in this way; others, taken over from Buddhist practice, vary in how explicit they are regarding their metaphysical commitments. We focus on the cognitive defusion exercises developed by Steven Hayes for ACT and the mindful breathing exercises popularised by Jon Kabat-Zinn in MBSR by way of illustration; the former is a particularly apparent case of metaphysical loadedness while the latter is noteworthy given its popularity and incorporation into other therapies as well as more generalised mindfulness programmes.

In the second part of the paper, we argue that this metaphysical loadedness ought to be taken into account in a way similar to side effects of other interventions. Clients ought to be informed of the metaphysical loadedness as part of informed consent, and mental health providers ought to take it into account in selecting courses of treatment. Our concern is not primarily the effect that the mindfulness exercises may have on a person's reasonable core beliefs about the self. Rather, we

poses, but we offer this example because we think it illustrates the thought process well.

<sup>iii</sup>The term 'feeling' here is taken up by proponents of mindfulness exercises as a translation of the Buddhist term *vedanā*. Putting aside complex interpretive issues, it can be understood roughly as sensations that are either pleasant, painful or neutral, and is among other things, included in the list of *skandhas* (see footnote vi). The relationship between *vedanā* and emotions is unclear, though *vedanā* may be an important component of emotions.<sup>17</sup>

<sup>i</sup>Linehan states mindfulness is 'non-denominational'; Kabat-Zinn claims mindfulness is 'universal', that it is 'just a particular way of paying attention'.<sup>15 20</sup> Similar reassurances are found in popular discussions of mindfulness.<sup>21–23</sup> The assumption that mindfulness exercises had been made philosophically neutral was a central part of the argument of a recent study of ethical issues connected to mindfulness exercises.<sup>24</sup>

<sup>ii</sup>We are sceptical that the case is straightforward with yoga



**To cite:** Ratnayake S,  
Merry D. *J Med Ethics* Epub  
ahead of print: [please  
include Day Month Year].  
doi:10.1136/  
medethics-2017-104201

## Extended essay

argue that simply practising these exercises will be in tension with many people's core beliefs and values.

These exercises offer an important case study because psychotherapies are frequently seen, in contrast to medication, as interventions without significant downsides. It is not obvious that the downsides of mindfulness exercises discussed here are less serious, say, than the side effects of antidepressant medication. Although what we say here does not apply straightforwardly to other kinds of psychotherapy, it does show that one cannot assume the harmlessness of a psychotherapeutic approach.

### MINDFULNESS EXERCISES AND THEORIES OF THE SELF

In this section we argue that, contrary to the sanitised presentation they are often given, mindfulness exercises are *metaphysically loaded* in that they encourage practitioners to view and conceptualise their mental states in a particular manner; a manner with which many would quite reasonably disagree. In what follows, we will clarify what this amounts to. To do so is easiest through consideration of how this way of seeing mental states connects to debates about the self to which the ancient Buddhists were a party, and in which contemporary psychologists have joined, in particular through the conception of self in ACT, which also draws inspiration from the earlier Vedantic conceptions of the self.<sup>iv</sup>

In describing the accounts of the self and their relationship to modern mindfulness exercises, we will focus on two sets of parallels. First, we will show that the arguments used by ACT to establish the account of the self accompanying cognitive defusion exercises originate in Vedantic philosophy and are similar to the arguments that establish the opposite conclusion in Buddhism. This not only undermines the claim that cognitive defusion exercises are non-denominational, but as we shall see in the next section, suggests that the justification for ACT's account of the self is more controversial than often presented.

Second, we focus on the two types of mindfulness exercises themselves, showing that they bear a striking resemblance to the mindfulness meditations of early Buddhism as described in the *Satipatthana Sutta* (SS).<sup>v</sup> Translated variously as 'Discourse establishing Mindfulness' and 'Frames of Reference' the SS is found both in the collection of middle-length discourses of the Buddha (*Majjhima Nikāya*) and in expanded form in the longer discourses (*Digha Nikāya*) which are components of the early Pāli Canon of Buddhist writings. As we shall see shortly, the influence of the SS in contemporary psychotherapy is readily apparent when we consider the texts that the likes of Kabat-Zinn cite within their own work. In mindfulness breathing exercises the resemblance to the meditation practice described in the SS is immediately apparent and explicit. In the case of the cognitive defusion exercises, we will show that they bear structural similarities to this practice.

ACT, unlike many other psychotherapies, is explicit about the fact that it comes bundled with a theory of the self which is referred to variously as the 'transcendental self' or 'self as context'.<sup>vi</sup> The argument establishing this account of the self,

as espoused by Steven Hayes, the founder of ACT, takes the unusual form of a behaviourist analysis of the term 'spirit'.<sup>3</sup> Apart from the methodology, Hayes takes his account of the self, more or less wholesale, from Deikman.<sup>4</sup> One of the central claims of this account of the self is that in spite of all the changes individuals undergo, there is one constant—the 'perspective, locus or point of view' from which one observes and interacts with the world:

If you are asked to remember what it was like to be 10 years old—to remember looking out at the world at that age and you are asked 'Was that you then?' or 'Deep down, is the you that is here now the you that was there then?' again the obvious answer is yes.<sup>3</sup>

The self espoused by Deikman and Hayes is characterless, bare and lacking any form whatsoever. It is simply the ongoing locus or context from which one is aware of, or can observe, the world and mental phenomena such as perceptions, thoughts and feelings. This is somewhat at odds with our ordinary first personal understanding of the self. While we might be prepared to agree that the above accurately describes our experience of 'looking out at the world', we tend not to think of ourselves as simply a vantage point from which we observe mental phenomena such as our own feelings and thoughts. We shall say more about how accounts of the self such as this sit in tension with core beliefs we might already have about the self in the next section.

In addition to the appeal from continuity of perspective, Hayes and Deikman argue that once we peel back the phenomena traditionally associated with the self, something still remains. Deikman appeals to Eastern sources to establish this point. The original source of the ACT account of the self is the school of views referred to as Vedanta. There are a number of early Indian philosophies that advance similar views of the self, which are variously considered by Deikman, but the majority of his sources, including the source below, fall under the Vedantic school. Here, Deikman describes the 'Who am I?' exercise of Ramana Maharshi:

If I lost my arm, I would still exist, therefore I am not my arm. If I could not hear, I would still exist. Therefore I am not my hearing. And so on, until finally, 'I am not this thought,' which leads to a radically different experience of the self.<sup>4</sup>

This radically different experience is, in ACT, the self as the continuing locus from which all the other phenomena that we traditionally associate with the self are observed.

Opposed to Vedantic accounts is the Buddhist no-self doctrine. The no-self doctrine is best understood as a negative metaphysical thesis. According to this doctrine, there are no substances or persisting entities which are the appropriate metaphysical referent to the term 'self', or related terms such as 'I'. Although this radically differs from the Vedantic view, it is motivated by some of the same claims, ones reflected in contemporary mindfulness exercises.

Much like the ACT view, the doctrine of no-self is at odds with commonplace understandings of the self. What the doctrine of no-self means for a first personal or phenomenological perspective is that individuals are deeply deluded with regard to how they understand themselves. While we tend to conceptualise ourselves as persisting subjects or selves, who have thoughts

arguments Hayes uses in this paper are basically the same as those used in philosophical discussions of the self.

<sup>iv</sup> Hayes claims relational frame theory offers 'a scientific account of how it [ACT's account of the self] emerges'.<sup>17</sup> Nonetheless, the core arguments are similar to those used in philosophical discussions of the self.<sup>25</sup>

<sup>v</sup> All references to the SS are sourced from Thanissaro.<sup>26</sup> We have chosen this translation as it is significantly less interpretively laden than certain other translations.

<sup>vi</sup> While Hayes is frequently explicit about the philosophical roots of ACT's theory of the self, he claims that Relational Frame Theory offers 'a scientific account of how it emerges'.<sup>25</sup> His core arguments for this theory of the self, however, are found elsewhere.<sup>4</sup> As we will see, the

and feelings, and so on, the doctrine states that there is no such underlying subject or self. There are thoughts and feelings, but no underlying entity to whom they can be said to belong.

A common argumentative strategy used to establish the Buddhist doctrine of no-self resembles strongly the approach used above to establish the opposite conclusion. This is unsurprising given that the Buddhist account of the self is a challenge to the earlier Vedantic view. Perhaps the most well-known instance of this argumentative strategy is found in the *Milindapanha*, where a king, Milinda, gets the following mysterious reply to the simple question, 'What is your name?':

Your majesty I am called Nāgasena, my fellow monks, your majesty, address me as Nāgasena: But whether parents, give one the name Nāgasena, or Sūrasena, or Virasena, or Sīhasena, it is nevertheless your majesty, just a counter, an expression, a convenient designator, a mere name, this Nāgasena; *for here there is no person to be found*.<sup>5</sup>

Milinda then asks what exactly the other monks in the order are referring to when they use the name 'Nāgasena'. He asks in turn whether the name refers to each of a person's physical constituents (eg, heart, lungs, brain) and whether it refers to each of a person's physical or mental constituents as identified in Buddhist doctrine (*skandhas*).<sup>vii</sup> Each time, Nāgasena replies in the negative.

This is the same argumentative strategy encountered previously; an exhaustive search for the self is conducted, resulting in a philosophical problem: if the self is not your thoughts, feelings, body, and so on, then is it something else? Vedantic philosophers answered yes, offering a number of candidates. Hayes gives an ambiguous answer: the self is a perspective, but he denies that this is an entity.<sup>6</sup> Here, when explicitly asked whether the self is to be found apart from the things considered, Nāgasena answers no. In doing so, he is appealing to the 'Principle of Lightness', a much older appeal to parsimony, not unlike Occam's razor.<sup>7</sup>

The philosophical problem motivating these theories of the self stems from a particular understanding of mental states and their relationship to the self. The problem would not arise if mental states were seen as partial constituents of the self, or its products, as they seem to be in many commonplace understandings of the self. We will find that mindfulness exercises in contemporary psychology are loaded with this controversial manner of conceptualising mental states as somehow separate from, or unrelated to, the self.

Apart from parsimony, a number of arguments support the doctrine of no-self. Using the strategy above, potential candidates for the self are examined exhaustively and their impermanence noted. This is significant as it is generally assumed that whatever the self is, it must be abiding, or continuous. Buddhists deny the permanence of the self both in the sense that the self is eternal in a way that survives death, and more pertinently for our ordinary beliefs, in the sense that underlying the changes individuals undergo, there is an abiding subject to be found. The argument from impermanence becomes extremely important as we turn to Buddhist meditation.

Of the many varieties of Buddhist meditation, the one that matters for present purposes is mindfulness meditation, as espoused in the SS of the Pāli Canon. Deikman was likely familiar with the SS through a popular secondary text and the

primary commentary by Buddhaghosa in the *Visuddhimagga*.<sup>8</sup> Further, there is abundant evidence that Kabat-Zinn was familiar with the SS and relied heavily on it in devising the MBSR mindfulness exercises; some of his sources cite the SS,<sup>9-14</sup> while others reproduce its text and discuss it extensively.

In a manner reminiscent of the argumentative strategy encountered above, in the SS, meditation practitioners are asked to consider themselves from various perspectives or frames of reference. For instance, they are asked to consider their body, their feelings, the skandhas, and so on. In thus observing themselves, practitioners are to *label* the phenomena under consideration and observe its quality, for instance, with regard to breathing:

... Always mindful, he breathes in; mindful he breathes out. Breathing in long, he discerns, 'I am breathing in long'; or breathing out long, he discerns, 'I am breathing out long.' Or breathing in short, he discerns, 'I am breathing in short'; or breathing out short, he discerns, 'I am breathing out short.'

This resembles the instructions given by Kabat-Zinn in an introductory breathing exercise:

Bring your attention to your belly, feeling it rise or expand gently on the inbreath and fall or recede on the outbreath. Keep the focus on your breathing, 'being with' each inbreath for its full duration and with each outbreath for its full duration, as if you were riding the waves of your own breathing.<sup>15</sup>

He goes on to state, in a manner that mimics even more closely the language of the SS: '...you should be aware of it [(the breath)] and feel the sensations associated with it and attend to their changing qualities.'

The SS progresses by describing how the same procedure is to be used, *mutatis mutandis*, with regard to other mental phenomena such as feelings:

And how does a monk remain focused on feelings in and of themselves? There is the case where a monk, when feeling a painful feeling, discerns, 'I am feeling a painful feeling.' When feeling a pleasant feeling, he discerns, 'I am feeling a pleasant feeling.' When feeling a neither-painful-nor-pleasant feeling, he discerns, 'I am feeling a neither-painful-nor-pleasant feeling.'

Apart from labelling and identifying the phenomena, practitioners are to note the way in which phenomena come to arise and pass, thereby coming to an understanding of their *impermanence*:

... he remains focused on the phenomenon of origination with regard to feelings, on the phenomenon of passing away with regard to feelings, or on the phenomenon of origination and passing away with regard to feelings.

In the initial directive to label, the practitioner considers their feelings in a first personal sense. Some distance is introduced when the practitioner is asked to observe these phenomena only with regard to their impermanence. By the end of the directive for each frame of reference, the SS instructs the practitioner to simply note, 'There are feelings': they are no longer the practitioner's feelings, simply feelings. There is a shift in the meditators understanding of themselves and their mental states, from recognising that they are personally experiencing certain phenomena such as feelings and thoughts, to impersonally noting that certain phenomena are the case, thereby relinquishing any claim to them. This suggests the same *separability*, or distance, between the self and mental phenomena

<sup>vii</sup> *Skandhas* (translated variously as heaps, aggregates or bundles) refer to the collection of components that are thought to exhaustively constitute sentient beings according to Buddhist doctrine; both physical and mental components are included in this taxonomy.<sup>7</sup>



## Extended essay

we encountered earlier. In stating impersonally that there are feelings, one is making a similar claim to 'I am not this thought,' encountered previously in the arguments supporting various accounts of the self.

There is some controversy about how to understand Buddhist meditation. Some have argued that the role of meditation in Buddhist practice and in understanding Buddhist metaphysical claims is exaggerated,<sup>12</sup> while others go so far as to claim that many early Buddhist texts are not expounding metaphysical claims but rather are for the most part describing what meditators can expect to experience during their practice.<sup>12 13</sup> There are also various textual and translation issues.<sup>viii</sup> We have waded into the midst of this debate by drawing parallels between arguments for the metaphysical doctrine of no-self and certain directives found in the SS. We do not mean to defend a particular interpretation of the SS here. Rather, we shall show that, regardless of how we might best interpret the original texts, the fact that contemporary mindfulness exercises encourage practitioners to conceptualise their mental states in a manner drawn from various accounts of the self is readily apparent, indicating their metaphysical loadedness.

The features of labelling, impermanence and separability that we encountered in the SS are obvious in the instructions for the MBSR sitting meditation which introduces mindfulness of thoughts and feelings in addition to the breath; Kabat-Zinn provides a list of the thoughts one is to identify based on their content. Practitioners are also directed to note feelings and moods. In addition to labelling, the emphasis on impermanence is clear:

When your attention is relatively stable on the breath, try shifting your awareness to the process of thinking itself. Let go of the breath and just watch thoughts come into and leave the field of your attention... Try to perceive them as 'events' in your mind... Note that an individual thought does not last long. It is impermanent. If it comes, it will go. Be aware of this... Note feelings and moods as they come and go.

The directive to perceive thoughts and feelings as 'events' closely resembles the directive in the SS to view oneself, and one's mental states, from an impersonal perspective. This involves taking a stance towards one's mental states which reflects a particular relationship with oneself. Namely that if there is a self, it is somehow separable from, and not to be found among, one's mental states. The relationship between the exercises and the background account of the self is explicit in Kabat-Zinn's wider remarks. Kabat-Zinn describes a process he refers to as 'selfing' where phenomena such as experiences, thoughts and feelings are constructed as 'my experience', 'my thoughts' and 'my feelings'.<sup>15</sup>

What we hope will be apparent by now to the reader, given the historical background, is the extent to which the view expounded by Kabat-Zinn is based on the no-self doctrine and that this is likely to be profoundly at odds with practitioner's pre-existing core beliefs about the self. While MBSR is explicit about the Buddhist roots of these exercises, their significance and association with the Buddhist account of the self is downplayed.<sup>15</sup> Furthermore, more generalised mindfulness programmes as well as other therapies, such as DBT, that have adopted these exercises often misleadingly present them as not subscribing to any particular religious or spiritual view whatsoever.

In the case of the cognitive defusion exercises developed by Hayes and adopted by psychotherapies such as DBT, where

they are classified under the umbrella of mindfulness, the parallels with the SS are less immediately obvious.<sup>16</sup> In Hayes' case, some of the cognitive defusion exercises explicitly establish the ACT account of the self in addition to displaying these aforementioned characteristics of labelling, impermanence and separability, so their metaphysical loadedness is explicit, though their controversial nature is not.<sup>3</sup> Our interest is in the exercises adopted by DBT which are ostensibly more innocent, but also adopt the same characteristics. In DBT, there is often a 'divide and conquer' approach used with regard to these characteristics. Labelling and noting the quality of an experience is used with breathing but more notably with emotions, where the practitioner is directed to identify and describe their emotions.<sup>ix</sup> For instance, in addition to labelling the emotion, practitioners are asked to consider what colour one's emotion might be.<sup>16</sup>

Noting impermanence and the separability of the self and mental states is reserved predominantly for targeting thoughts, in the context of visualisation exercises such as the following:

- Imagine sitting in a field watching your thoughts float away on clouds.
- See your thoughts written in the sand and then watch the waves wash them away.
- Envision yourself driving a car and see your thoughts pass by on billboards.<sup>16</sup>

The similarity to the SS in terms of impermanence should be apparent. Note also that the practitioners are positioned to observe their thoughts from a distance, as if they were separate and apart from them. In addition to the imagery adopted, separability and impermanence are often an explicit part of the instructions:

Just be sure that your idea captures the purpose of this exercise, which is to visually watch your thoughts come and go without holding on to them and without analyzing them.<sup>16</sup>

In spite of how they are presented to the public, both the breathing exercises of Kabat-Zinn and the cognitive defusion exercises of Hayes bring philosophical content with them and cannot be extracted fully from their original contexts which involve various Eastern accounts of the self. Their metaphysical loadedness is apparent where the practitioner is encouraged to conceptualise their mental states as separable from the self, a claim found in various background theories of the self.

As noted throughout, these background accounts of the self and the claim that individuals are somehow separable from their mental states are likely to sit in tension with a range of common and important core beliefs that practitioners may hold about the self. The most obvious is the tension between the various Eastern accounts of the self which underlie mindfulness exercises and any belief in an immortal soul with attributes of a person's personality or moral character; belief in such an immortal soul is a core doctrine in many other religions. Mindfulness exercises also conflict with widely held fundamental secular beliefs, such as the idea that there is a stable set of attributes that make up who we are, that our thoughts and feelings are reflections of this, and that we can see ourselves in them and judge our worth. These

<sup>viii</sup>For example, particular translations make the connection with the doctrine of no-self far more explicit.<sup>14</sup>

<sup>ix</sup>If emotions are construals (as Robert C. Roberts suggests)<sup>27</sup> or judgments (as Martha Nussbaum holds),<sup>28</sup> dealing with them in this way may also be suspect, as it asks people to desensitise themselves from what is valuable. It also may lead to seeing emotions as episodic, where the ongoing nature of emotions has been emphasised.<sup>29-32</sup> We thank an anonymous reviewer for *Journal of Medical Ethics* for raising these points.

beliefs are strongly at odds with the directive found both explicitly and implicitly in mindfulness exercises to consider thoughts and feelings as separable from the self. We shall discuss these tensions in more detail; in the next section we shall also consider the ethical and epistemic concerns raised as a result.

### THE EPISTEMIC COST: IMPLICATIONS FOR PRACTICE

In this section, we argue that mental health providers and their clients should consider the metaphysical loadedness of mindfulness exercises in assessing treatment options, in a way analogous to possible side effects of medication. We believe that in order to meet prevailing standards of informed consent, clients must be informed of the metaphysical loadedness of mindfulness exercises, and given appropriate space and support to assess whether the benefits of the treatment justify their use. This also suggests that mental health providers should consider whether the patient's condition is serious enough to justify a metaphysically loaded course of treatment, in light of available alternatives.

In saying that the metaphysical loadedness of mindfulness exercises should be treated like a side effect, we do not claim that metaphysical loadedness necessarily leads to further negative effects. We shall begin by considering this possibility and the implications for practice. However, we are more centrally concerned that engaging in the practice itself may be in tension with a person's core beliefs and values. After considering this possibility and what it entails, we will close the discussion by addressing some plausible objections to the argument presented here.

Although it is not our central worry, we think some concern about effects and alterations to a person's core beliefs is justified. The core beliefs we are concerned about are those which are reasonable and valued parts of a client's world view. As we saw, mindfulness exercises encourage practitioners to see their mental states as events that happen to the self, rather than its products or constituents. According to some Christians, the self is a soul tainted by original sin. Various mental states are thought to evidence this tainted state, such as lust or greed. If a course of therapy were to remove or weaken this belief by persuading individuals, for instance, that feelings of lust were not essentially tied up with one's nature, that would be a cost: Christian clients may rightly feel they had been led by therapy into heresy.<sup>x</sup> Similarly in a secular sense, people may reasonably enough see themselves as being partly constituted by various aspects of their personality, such as kindness or impulsiveness, which are correspondingly accompanied by various mental states such as feelings of empathy and reckless judgements. Since mindfulness exercises encourage people to see mental states as events that do not belong to the self, they teach people not to see their mental states as evidence for what their self is like. They thus oppose deep beliefs people may have about who they are. Mental health providers should be reluctant to interfere with people's reasonable core beliefs and clients may quite reasonably see tensions with them as reasons not to undergo a course of treatment.

The effects of mindfulness on core beliefs have not been studied, so are unknown. But there are intuitive reasons to suspect they might have some effect, and the burden of proof usually lies on those claiming that something is safe. At the very least, the discussion thus far is an invitation to conduct further research.

We do not intend to deny that these exercises might have positive effects on people's core beliefs about themselves. Through building metacognitive abilities, mindfulness exercises may help people develop insight that some of their core beliefs are delusional. But there is no tension between thinking that a course of therapy may remove delusions about the self, and that it may also affect reasonable and valued core beliefs about the self, just as there is none in worrying about damage to the brain in removing a brain tumour. The potential costs and benefits must be weighed against each other, and frequently, the assessment will involve a difficult judgement call.

Psychology is a frequently surprising subject, and it might surprisingly turn out that mindfulness exercises have little impact on people's core beliefs about themselves, or that they *only* affect delusional beliefs. For this reason, we see the core of our argument as not being the effect mindfulness exercises have on core beliefs. Instead, we are most interested in the tensions that might exist between the very practice of these exercises, and the core beliefs and values people hold.

Many people would reject the views of the self involved in mindfulness exercises. One kind of tension this introduces is that, for these people, mindfulness exercises could be a consoling fantasy. Suppose an atheist was having trouble coping emotionally with a natural disaster in which there were many casualties. They might object to imagining the people who died in the disaster frolicking in heaven, even if doing so would help them cope. They do not believe in heaven, thus the imaginative exercise is in tension with their core beliefs. Mindfulness exercises involve a similar tension with the core beliefs of those who strongly reject the metaphysical claims they involve, for example, for those who see aspects of their personality as partly constitutive of who they are and those of particular religious persuasions. These beliefs are likely to be of importance to those who hold them even when they bring about significant suffering, as they may, for instance, in the aforementioned religious case where one could be overwhelmed by guilt about one's flawed nature. Practising mindfulness exercises for such an individual in order to relieve suffering would be participating in a consoling fantasy akin to the atheist case.

Different people will rate the importance of tensions between one's core beliefs and the views of the self associated with mindfulness exercises differently: a practitioner might accept the tension involved in using such imaginative exercises to overcome severe problems in interpersonal relationships, but not be willing to do so to simply relieve suffering. Such value judgments need to be acknowledged and respected when deciding whether a course of treatment is appropriate.

A different tension may arise between a person's *epistemic* values and the *pragmatic* use to which mindfulness exercises are put. Since mindfulness exercises are metaphysically loaded, they are appropriate objects of critical exploration, as part of self-discovery. In therapy, they are adopted as a tool to relieve suffering and increase functioning. This goes against common epistemic values, the sort of values that lead one to accept uncomfortable truths, and to be honest, even brutally honest, with oneself. A common example of a tension between pragmatism and epistemic values is the lack of consensus about whether one would like to know if one's partner was cheating on them; some prefer the 'ignorance is bliss' approach while others defend the importance of truth even when it brings about distress. In the same way that there is legitimate disagreement about how the tension between pragmatically avoiding suffering and epistemic values in this case, patients might legitimately have different views regarding exercises which may bring their epistemic values into question. Note

<sup>x</sup>In presenting this paper, an audience member mentioned being unable to continue with meditation because of a conflict with a different Christian theory of the self.

## Extended essay

that this tension may arise even if one basically agrees with the metaphysical claims involved in these exercises, since one gives up the appropriate critical perspective.

This tension has particular relevance to ACT. Hayes presents the transcendental self in ACT as an assumption inherent in language, and claims talk of the transcendental self is justified from a scientific perspective.<sup>17</sup> This may give clients the impression that they are accepting a claim about the self on good scientific authority. Yet the theory of the self in ACT is defended, as we saw, using similar considerations as the Vedantic theory of the self. These considerations are contestable, and can equally well support the Buddhist denial of the self. Providers of ACT should be clear that the transcendental self is a philosophically controversial position, and practitioners are being asked to adopt it for pragmatic rather than epistemic reasons.

It may be objected that the epistemic costs of mindfulness exercises are importantly different from physiological side effects. One difference is that paradigmatic side effects such as stomachaches and the collateral damage that may accompany certain medical procedures are questions of bodily integrity. We will return to the accusation of disanalogy shortly but note for now that this difference might seem particularly important if you support the waiver model of informed consent, which construes informed consent as a waiver of otherwise prevailing moral or legal norms, and so only required when an action would violate them.<sup>18</sup> For instance, without consent from the patient, surgery would be physical assault. Since laypeople are allowed, the objection goes, to offer meditation exercises and philosophical arguments to each other, mindfulness exercises do not violate any prevailing norms.

But there is a relevant norm in professional contexts: professionals must respect the religious, spiritual and ethical beliefs of their clients.<sup>xi</sup> People should reasonably expect that professionals will not try to convert them, for example, and will respectfully acknowledge when a recommended course of action is morally unacceptable to the patient. Maintaining this norm is partly an acknowledgement of the power imbalance that exists between professionals and their clients: patients may feel considerably less able to raise objections or adopt a critical stance towards what the professional says given the perceived asymmetry in expertise. As such it is particularly important for professionals to stay within the realms of their expertise; they should not use their expertise in law or medicine to give weight to their moral or religious convictions. Finally, it is a question of trust. People who seek services from professionals often seriously need them and they will not seek help if they expect these professionals to lead them away from their religious convictions, or pressure them into doing things that they find morally problematic. The possible tensions detailed earlier between mindfulness, core beliefs and values show that recommending mindfulness exercises potentially violates this professional norm, so there is good reason to think a waiver would be required in order to meet the standards of informed consent.

A second possible difference is that the metaphysical issues involved in mindfulness exercises are harder to explain to people than other side effects, and that trying to explain it to people may alarm them unnecessarily. One might think that this is another

way in which 'typical' side effects differ from the metaphysical loadedness of mindfulness exercises. We believe however that the difficulty of explaining the metaphysical loadedness reveals an analogy with more common side effects; in both cases, the professional reveals something about the treatment that is not immediately obvious to the patient. In the same way that a patient is likely to be unaware that a particular medication will cause nausea, they will be unaware of the metaphysical implications of mindfulness exercises.

We believe that the difficulty of explaining these issues is in danger of being overstated. Although the details of the metaphysics are delicate and difficult to explain, the general issue is not. Most people have spent some time reflecting on questions of the self: whether something of them will survive the body, whether they are their brain, whether they are the same person they were 10 years ago. Mental health professionals can, in the first instance, alert patients that there is an issue of this sort. For some patients, this will suffice to establish either that they are comfortable with the therapy, or that they are not.

Other patients might only consent to the therapy if they understood the metaphysics better, a task that may be onerous and time consuming. Mental health providers might develop resources that will help patients understand the metaphysical issues. Such reluctance may, however, be a reason to consider an alternative therapeutic approach. If patients are not willing to proceed when they are aware that there may be metaphysical implications of the treatment that they do not understand, it is not acceptable to get them to accept the course of treatment by keeping quiet about these implications.

Another possible objection is that almost every therapeutic intervention will come with some metaphysical commitments.<sup>xii</sup> As such, mindfulness exercises and the background theories of the self that accompany them may only differ in degree rather than kind from other interventions. For example, a psychopharmacological intervention for depression presupposes that a person's mood can be influenced by changes to the body, which could be accompanied by particular metaphysical commitments regarding the nature of causality and the body. It might be an important part of a person's self-conception, however, to see their mood as influenced by what is going on in their life, rather than the state of their body. Taking the drug might lead them to see their mood as more influenced by the state of their body than they would like to think. Consequently, one might think that it is unclear whether mental health service providers ought to discuss these potential metaphysical implications given that there are likely to.

There are two responses we would make to this objection. The first is that mindfulness exercises seem to be an unusual case in that many other interventions such as psychopharmacological ones do not involve a representation of the world as being a certain way, while mindfulness exercises do. In mindfulness exercises, practitioners are encouraged to represent to themselves a particular account of the self that they may feel uncomfortable with. They actively engage with this representation by conceptualising and in some cases acting out (via visualisation exercises) a particular relationship between themselves and various phenomena such as thoughts. Psychopharmacological interventions may presuppose that the mood can be influenced by changes in the body, but insofar as their effectiveness exceeds that of a placebo, they do not depend on the patient actively engaging with representations of things being

<sup>xi</sup>Our argument here also puts considerable pressure on the idea that mindfulness interventions are compatible with liberal neutrality, as argued by Schmidt.<sup>24</sup>

<sup>xii</sup>We thank an anonymous referee for *Journal of Medical Ethics* for raising this objection.



this way. Even when practitioners are not asked to see these representations as accurate, they might reasonably object to using them to produce clinical outcomes. There is an uncomfortable parallel between participating in the rituals of another religion or belief system that one objects to and practising mindfulness exercises that is very salient in the case of mindfulness exercises that is potentially less obvious with other interventions.

Second, if it were the case that mindfulness exercises are not a special case and that other interventions came with metaphysical commitments that the clinician felt would be at odds with their patient's core beliefs we do believe that this should be made apparent to the patient and discussed openly in selecting a treatment plan. It seems highly unlikely that a clinician could discuss every putative metaphysical commitment of an intervention but following the precedent set by other areas of medical ethics, we suggest that these issues be discussed where there is good reason to believe that they will sit in tension with a particular patient's core beliefs and values, in particular moral beliefs. For instance, while a doctor would not normally feel the need to discuss at length the necessity and benefits of a blood transfusion, this will not be the case where the patient objects to them based on religious grounds. Similarly, it seems plausible to think that in the case of specific patients, for instance, those holding a Christian world view, it is more pressing to discuss metaphysical commitments. Furthermore, where the course of treatment raises metaphysical concerns for the patients as they are undergoing it, clinicians should be open to discussing rather than dismissing these concerns. For instance, if it were the case that a psychopharmacological intervention distresses the patients as they are uncomfortable with the belief that their mood is not as responsive to life events than they had first believed this should be discussed rather than deemed to be outside the purview of the treatment.

We close with some further remarks about practical implementations that expand on the observations made above. First, the requirement for informed consent has been supplemented in many practices by a requirement for shared deliberation, in which clients discuss the pros and cons of therapeutic approaches with mental health providers, and aim to come to a shared decision.<sup>19</sup> If mental health providers enter such discussions believing that mindfulness exercises have no metaphysical implications, or that it is never reasonable to refuse therapy for these reasons, then they may under-rate patients' legitimate concerns. This is particularly problematic in light of the power asymmetry discussed earlier. Part of training in therapies involving these exercises should involve learning about their metaphysical loadedness, and the variety of resultant concerns patients may have both before and as a part of ongoing treatment. Furthermore, as we saw immediately above, it is possible that metaphysical commitments which sit in tension with a patient's core beliefs could encompass a number of interventions beyond just mindfulness exercises. As such, it is more pressing to interrogate the current model of the therapeutic relationships with an eye towards improving the procedure of shared deliberation.

Second, it should be apparent that we do not think mindfulness exercises such as these should be abandoned altogether. Apart from informing clients, there may be avenues for making mindfulness exercises less metaphysically loaded. One can pay attention to the present moment in more and less metaphysically loaded ways, and if this is the mechanism through which mindfulness assists people in dealing with mental illness, then it may be possible to develop exercises that are more metaphysically innocent. Alternatively, one might develop a range of exercises drawing from different spiritual traditions, so as to be able to offer exercises that line up with a client's metaphysical convictions.

The final remark concerns the dissemination of mindfulness exercises to the general public. Decisions to distribute mindfulness widely should be undertaken only to achieve specific outcomes which are significant enough to outweigh the downsides discussed here. Importantly, such efforts must make it clear that mindfulness exercises encourage practitioners to conceptualise the relationship between their mental states and their self in a way many will find objectionable. We urge mental health providers to take a more balanced approach that considers the epistemic costs of mindfulness exercises, and communicates them openly.<sup>xiii</sup>

**Contributors** SR contributed the initial idea for the paper and primarily wrote section 1. DM primarily wrote section 2. The content in both sections was refined through considerable discussion by both authors.

**Funding** This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** None declared.

**Provenance and peer review** Not commissioned; externally peer reviewed.

© Article author(s) (or their employer(s) unless otherwise stated in the text of the article) 2018. All rights reserved. No commercial use is permitted unless otherwise expressly granted.

## REFERENCES

- Kabat-Zinn J. *Wherever you go, there you are: mindfulness meditation in everyday life*. New York, NY: Hyperion Books, 1994.
- Harris R. *The happiness trap*. Wollombin: Exisle Publishing, 2007.
- Hayes SC. Making sense of spirituality. *Behaviourism* 1984;12:99–110.
- Deikman A. *The observing self: mysticism and psychotherapy*. Boston, MA: Beacon Press, 1982.
- Anon. *The questions of King Milinda*. Oxford: Clarendon Press, 1890.
- Hayes SC. Buddhism and acceptance and commitment therapy. *Cogn Behav Pract* 2002;9:58–66.
- Siderits M. *Buddhism as philosophy*. Surrey: Ashgate Publishing, 2007.
- Rahula W. *What the buddha taught*. New York, NY: Grove Press, 1959.
- Gunaratana H. *Mindfulness in plain English*. Somerville, MA: Wisdom Publications, 1993.
- Hanh TN. *The miracle of mindfulness: a manual on meditation*. Boston, MA: Beacon Press, 1987.
- Thera N. *Satipatthana: The heart of Buddhist meditation: a handbook of mental training based on the Buddha's Way of mindfulness*. London: Rider, 1962.
- Dreyfus G. *The sound of two hands clapping: the education of a tibetan buddhist monk*. Berkeley, CA: University of California Press, 2003.
- Shulman E. *Rethinking the buddha: early buddhist philosophy as meditative perception*. Cambridge: Cambridge University Press, 2014.
- Jokita U. Dahmminda U trans. Mahāsātipatthāna Sutta: the greater discourse on steadfast mindfulness [Internet]. 1986 <http://nrcvee.iitd.ac.in/files/sutta-study/mahasatipatthana-english.pdf>
- Kabat-Zinn J. *Full catastrophe living*. New York, NY: Bantam Dell, 2005.
- McKay M, Wood J, Brantley J. *The dialectical behavior therapy skills workbook*. Oakland, CA: New Harbinger, 2007.
- Anālayo. *Satipatthāna: the direct path to realization*. Cambridge: Windhorse Publications, 2004.
- Manson NC, O'Neill O. *Rethinking informed consent in bioethics*. Cambridge: Cambridge University Press, 2007.
- Liégeois A, Eneman M. Ethics of deliberation, consent and coercion in psychiatry. *J Med Ethics* 2008;34:73–6.
- Linehan M. *DBT skills training manual*. 2nd edn. New York, NY: The Guilford Press, 2015.
- Kozlowski E. Can christians practice mindfulness? [Internet]. Huffington Post: 2013 [cited from 2017 Jan 18]. 2013 [http://www.huffingtonpost.com/eden-kozlowski/mindfulness-and-religion\\_b\\_3224505.html](http://www.huffingtonpost.com/eden-kozlowski/mindfulness-and-religion_b_3224505.html)
- Center for Mindfulness. History of MBSR [Internet]. 2014 <http://www.umassmed.edu/cfm/stress-reduction/history-of-mbsr/> [cited 18 Jan 2017].
- Mental Health Foundation. Be mindful [Internet]. <http://bemindful.co.uk/> [cited 18 Jan 2017].
- Schmidt AT. The ethics and politics of mindfulness-based interventions. *J Med Ethics* 2016;42:450–4.

<sup>xiii</sup> Our thanks to Chris Siver, Sally Merry, Alan Merry, some of the psychologists at Headspace (Frankston, Australia), an audience at the 2016 NZAP in Waikato, and two anonymous reviewers for *Journal of Medical Ethics* for helpful comments and discussions, as a result of which the paper was greatly improved.

## Extended essay

- 25 Fletcher L, Hayes SC, *et al.* Relational frame theory, acceptance and commitment therapy, and a functional analytic definition of mindfulness. *JRE&CBT* 2005;23:315–36.
- 26 Thanissaro trans. Satipatthana Sutta: frames of reference [Internet]. 2008 <http://www.accesstoinsight.org/tipitaka/mn/mn.010.than.html>
- 27 Nussbaum MC. *Upheavels of thought: the intelligence of emotions*. Cambridge: Cambridge University Press, 2001.
- 28 Roberts RC. *Emotions: an essay in aid of moral psychology*. Cambridge: Cambridge University Press, 2003.
- 29 Goldie P. *The emotions: a philosophical exploration*. Oxford: Oxford University Press, 2000.
- 30 Goldie P. Emotions, feelings and intentionality. *Phenomenol Cogn Sci* 2002;1:235–54.
- 31 Ben-Ze'ev A. Emotion as a subtle mental mode. Solomon RC, ed. *Thinking about feeling: contemporary philosophers on emotions*. Oxford: Oxford University Press, 2004:250–68.
- 32 Ratcliffe M. The Feeling of Being. *J Consious Stud* 2005;12:43–60.





# Forgetting ourselves: epistemic costs and ethical concerns in mindfulness exercises

Sahanika Ratnayake and David Merry

*J Med Ethics* published online March 9, 2018

---

Updated information and services can be found at:

<http://jme.bmj.com/content/early/2018/03/09/medethics-2017-104201>

---

*These include:*

## References

This article cites 7 articles, 2 of which you can access for free at:

<http://jme.bmj.com/content/early/2018/03/09/medethics-2017-104201#ref-list-1>

## Email alerting service

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

---

## Notes

---

To request permissions go to:

<http://group.bmj.com/group/rights-licensing/permissions>

To order reprints go to:

<http://journals.bmj.com/cgi/reprintform>

To subscribe to BMJ go to:

<http://group.bmj.com/subscribe/>