

Journal of Buddhist Ethics

ISSN 1076-9005

<http://blogs.dickinson.edu/buddhistethics/>

Volume 23, 2016

Facing Death from a Safe Distance: *Samvega* and Moral Psychology

Lajos Brons

Nihon University & Lakeland University

Copyright Notice: Digital copies of this work may be made and distributed provided no change is made and no alteration is made to the content. Reproduction in any other format, with the exception of a single copy for private study, requires the written permission of the author. All enquiries to: cozort@dickinson.edu.

Facing Death from a Safe Distance: *Samvega* and Moral Psychology

Lajos Brons¹

Abstract

Samvega is a morally motivating state of shock that—according to Buddhaghosa—should be evoked by meditating on death. What *kind* of mental state it is exactly, and how it is morally motivating is unclear, however. This article presents a theory of *samvega*—what it is and how it works—based on recent insights in psychology. According to dual process theories there are two kinds of mental processes organized in two “systems”: the experiential, automatic system 1, and the rational, controlled system 2. In normal circumstances, system 1 does not believe in its own mortality. *Samvega* occurs when system 1 suddenly realizes that the “subjective self” will inevitably die (while system 2 is already disposed to affirm the subject’s mortality). This results in a state of shock that is morally motivating under certain conditions. *Samvega* increases mor-

¹ Department of Philosophy, Nihon University, and Lakeland University, Japan Campus, Tokyo, Japan. Email: mail@lajosbrons.net.

tality salience and produces insight in suffering, and in combination with a strengthened sense of loving-kindness or empathic concern both mortality salience and insight in suffering produce moral motivation.

Introduction

The Pāli/Sanskrit term “*saṃvega*”—literally meaning (fearful) trembling—is used in Buddhist writings to denote a religiously and morally motivating state of shock or agitation.² For example, *Aṅguttaranikāya* 4.113 describes four kinds of persons being shocked (*saṃvijjati*) into a state of *saṃvega* by stories or experiences of suffering; Śāntideva writes that “the virtue of suffering has no rival, since, from the shock [*saṃvega*] it causes, intoxication falls away and there arises compassion for those in cyclic existence, fear of evil, and a longing for the Conqueror [i.e., the Buddha]” (*Bodhicaryāvatāra* 6.21); and Buddhaghosa suggests that the (repeated) experience of *saṃvega* decreases attachments (*Visuddhimagga* III.58), and increases loving-kindness (XIII.35) and vigor (XIV.137).

Samvega is (obviously) a disturbing experience—disturbing enough to result in a profound change in the attitudes, beliefs, and/or values of the person experiencing it—but its beneficial effects are reason to seek rather than to avoid it. Towards this end, Buddhaghosa recommends meditating on the subject of death. One may accidentally experience *saṃvega* in the course of ordinary life (in encountering death and suffering particularly), but meditation is the primary means to *intention-*

²See, for example, Brekke, and Heim. Coomaraswamy translates “*saṃvega*” as “aesthetic shock” and quotes a number of occurrences in non-Buddhist texts, where it means (fearful) trembling. The term is often translated as “sense of urgency,” but other translations, including “shock” and “terror,” can be found.

ally achieve *saṃvega*. In the section on death as a meditation subject (VIII.1-41), Buddhaghosa describes a number of meditation exercises focusing on the absolute unavoidability, finality and irreversibility of death. *Saṃvega* is the aim and purpose of these exercises: as VIII.5-6 makes clear, meditation on death is successful only if it results in *saṃvega*.³ (See also III.58.)

Death is one of two general meditation subjects—the other is a gradually extending loving-kindness—which “are needed generally and desirable owing to their great helpfulness” (III.57-59). Hence, meditating on death is not some kind of curiosity or extravagance, but is one of the most basic and most important exercises. And by implication, the experience of *saṃvega* is—at least for Buddhaghosa—among the most basic and most important goals of meditation. This, of course, conflicts with the image of (Buddhist) meditation as stress-reduction spread by the mindfulness industry, but as pointed out by Donald Lopez (108) and others, that image is very wrong. The goal of (at least some forms of) Buddhist meditation is *stress induction*, rather than stress reduction. Or, as Paul Williams remarks, “the spiritual path is not one of comfortable feelings and acceptance. It is deeply uncomfortable” (xxv).

There are very many different types of Buddhist meditation,⁴ but traditionally, two main kinds are distinguished. Tranquility (Pāli: *sa-*

³The same passage also clarifies that *saṃvega* should not be confused with something like sorrow.

⁴This variety suggests that “meditation” is a functional rather than a substantive category, which is also supported by Buddhaghosa’s refusal to define “meditation” because it “is of many sorts and has various aspects” and an attempt to definition would only “lead to distraction” (III.2). “Meditation,” then, is not defined by some substantive properties that all activities called “meditation” have in common, but by those activities’ function, by what meditation is intended to establish. However, even a functional definition of “meditation” would probably have to be disjunctive, as there does not

matha/Sanskrit: *śamatha*) meditation aims at achieving states of concentration called *jhānas/dhyānas* in which all attention is focused on its object. Tranquility meditation is associated with mindfulness (*sati/smṛti*) and is primarily a collection of techniques for improving the ability to concentrate, which is necessary for the second kind of meditation. That second kind of meditation is insight (*vipassanā/vipaśyanā*) meditation. Hence, except perhaps in Chan/Zen (which takes its name from the Chinese translation of “*dhyāna*”), tranquility meditation is not an end in itself but means towards the end of achieving insight. Insight does not result from mere concentration—as Lopez remarks, insight meditation “often entails a process of reflection, and even discursive reasoning, in which a thoroughgoing mental search is launched by the mind” (88).⁵ The insight to be achieved is of a threefold nature: insight into (or understanding of) impermanence (*anicca/anitya*), insight into no-self (*anattā/anitya*), and insight into suffering (*dukkha/duḥkha*).

It is not immediately clear how meditation on death fits into this classification. The explicit aim and purpose of meditation on death is *saṃvega*, which may be conducive to other kinds of meditation and religious activity in general, but which is neither concentration nor insight itself. On the other hand, *saṃvega* is not exactly not concentration or not insight either (pardon the double negations). It is not not concentration, because one of its (supposed) effects is to focus (i.e., concentrate) the mind on religious and moral goals; it is not not insight, because death

seem to be a single purpose, goal, or function that all activities called “meditation” share.

⁵While this conflicts with the popular view of meditation, it is very much in line with the English term’s etymology. The English term “meditation” comes from Latin “*meditari*” meaning cogitation, contemplation, pondering, thinking deeply, and so forth. Until fairly recently, occurrences of the term and its cognates in European languages referred to meditation in this sense. An obvious example is that of Descartes’s *Meditations on First Philosophy*.

has obvious relations to impermanence, no-self, and suffering, and insight into death is inseparable from insight into those. This ambiguity is closely related to the obscurity of the notion of *saṃvega* itself. That notion raises a number of puzzling questions to which Buddhaghosa offers no answers: What kind of state is *saṃvega*? How does it produce its supposed effects? How can a shocking realization of one's mortality be morally motivating? It is these questions that this article aims to answer.

It should be noted that these questions about a certain mental state and its effects are psychological questions primarily, even if they are of hermeneutical interest as well. By implication, this article is not concerned with Buddhaghosa's theory of *saṃvega* (if he had one) except insofar as elements thereof contribute towards the project of answering these questions by means of a (more or less) naturalistic account of that mental state. Consequently, this is not an article in Buddhology or in philosophy as traditionally conceived, but in theoretical moral psychology (or in the kind of "natural(istic) philosophy" that rejects boundaries between philosophy and theoretical science). Nevertheless, the article's topic has thus far been of interest to philosophers and religious thinkers mainly, and even if it does not belong to philosophy in some (overly) strict sense, it does aim to contribute to philosophy of religion and to moral philosophy—including Buddhist ethics—as well as to moral psychology.⁶

⁶To what extent a naturalistic account of what *saṃvega* is (or could possibly be) and how it could motivate—given all we know about the human mind—*can* contribute to Buddhist thought is debatable, because it is conceivable that naturalistic and Buddhist accounts of *saṃvega* turn out to be incompatible. One of the most central doctrines of Buddhism is the rejection of annihilationism (*ucchedadṛṣṭi*—the view that the self is annihilated at death) and eternalism (*śāśvatadṛṣṭi*—the view that the self survives death) in favor of a "middle way," but that middle way depends on commitments to mind/body dualism, *kamma/karma*, and rebirth, and none of these commitments are compatible with contemporary science. Rather, contemporary science is uncompromis-

From Buddhaghosa's writings (and other occurrences in Buddhist literature mentioned above) a number of essential features of *saṃvega*—or a “*samvegic*” mental state—can be inferred. As a rough definition, I propose the following:

Samvega (or a “*samvegic*” state) is

- (a) an occurrent (i.e., non-persistent) mental state with relatively short duration (better measured in minutes than in longer units of time),
- (b) that is shocking/agitating and absorbing (i.e., it is a state that disturbs and that completely takes over the conscious mind),
- (c) that is related to some kind of realization of one's mortality (such that either it *is* that realization or it is caused by it), and
- (d) that is religiously and/or morally motivating.

In the Western philosophical tradition, states similar (or identical, perhaps) to *saṃvega* have been mentioned by Epictetus and Søren Kierkegaard,⁷ and more recently by Thomas Nagel, Mark Johnston, and James

ingly annihilationalist (as understood here), and by implication, so is the account of *saṃvega* presented in this article. I don't think that this implies an incompatibility with Buddhist accounts of *saṃvega*, however, or with Buddhaghosa's writings in particular, but I will make no attempt to defend that belief here.

⁷Søren Kierkegaard (*Sickness, Anxiety*) suggested that a kind of existential shock or dread could be a religious motivator, and the teachings of Epictetus include some (cryptic) references to a shocking or agitating mental state related to the fear of death (e.g., *Discourses* II.1.10). According to Epictetus, this mental state should not be avoided or repressed, but one should learn to not let oneself be (outwardly) disturbed by it. A passage in *The Manual* suggests that it should even be cultivated because a daily reminder of one's inevitable death is a powerful motivator to do the right thing (§21).

Baillie. However, in all five cases it is uncertain whether the states mentioned actually satisfy all four criteria mentioned. Furthermore, with the exception of Baillie, these accounts offer little explanation of what kind of state *saṃvega* is (assuming that the states they mention are *saṃvegic* states indeed) and how it works. Baillie describes a state in which “rational capacities are immobilized and one is engulfed in inarticulate terror” (S188) and that satisfies the first three criteria, but possibly not the fourth. About his experience of that state he writes:

I entered into a state of mind unlike anything I had experienced before. I realized that I will die. It may be tomorrow, it may be in fifty years time, but one way or another it is inevitable and utterly non-negotiable. I no longer just knew this theoretically, but knew it in my bones. . . . It was as if the blinders had been removed, and I was the only person to have woken up from a collective dream to grasp the terror of the situation. (S189)

Clearly, Baillie’s experience satisfies criteria (a), (b), and (c). He does not describe it as morally (or religiously) motivating, but he does claim that it “was a pivotal moment in [his] life, with a significant influence on [his] future attitudes and actions” (S190). That—I think—is close enough to the spirit of criterion (d) to classify his experience as *saṃvega*, especially if it is taken into consideration that (d) is the supposed *effect* of *saṃvega* and that there is no reason to assume that *saṃvega* has the *exact* same effect on everyone and in all circumstances (implying that criterion (d) should not be taken as a necessary condition). Furthermore, regardless of whether Baillie experienced *saṃvega* in some strict sense, he offers a theoretical explanation of the state he experienced that explains what kind of state *saṃvega* is, although some amendments are necessary (largely due to the naturalistic commitment mentioned above). Because he did not consider his experience morally motivating, Baillie has noth-

ing to offer with regards to the second question—how *saṃvega* works—however, but Terror Management Theory gives a partial answer to that question.

The following three sections give an answer to the question of what kind of state *saṃvega* is based on Baillie’s theory, which is summarized first, followed by the identification of a number of minor problems and amendments to fix those. In the section after that I investigate whether *saṃvega* is necessarily related to “the full appreciation of our mortality” (Baillie S188), and conclude that this is probably the case. How *saṃvega* works—that is, how *saṃvegic* shock produces moral motivation—is discussed in the section after that. I argue in that section that the combination with loving-kindness meditation recommended by Buddhaghosa is essential for its morally motivating effects. The final section summarizes this article’s findings before turning—briefly—to the practical question whether these two meditation subjects, death and loving-kindness, should indeed be recommended (as Buddhaghosa does).

James Baillie on the “Expectation of Nothingness”

In “The Expectation of Nothingness,” James Baillie attempts to answer a puzzling epistemological question about a sense of terror or panic resulting from “the full appreciation of our mortality” (S188). Borrowing a term from Thomas Nagel, he calls this state the “expectation of nothingness,” but as suggested above, I will assume that the state he experienced and tries to explain is *saṃvega* and will call it such.

Baillie’s question is: Given that I already knew that I will one day die, how can I be shocked by that fact? In his answer he builds on ideas by Thomas Nagel and Mark Johnston, but most of all on Tamar Szabó Gendler’s notion of “alief” (“Belief,” “Action”). Gendler introduced the

alief/belief contrast in an attempt to explain the discrepancy between our explicit beliefs and what our behavior seems to indicate about our beliefs in (what came to be called) “Gendler cases.” One of her recurring examples is the Skywalk, a glass-floored horseshoe-shaped structure above the Grand Canyon. No one would enter the Skywalk without believing it is safe, but many people feel very unsafe seeing the abyss below their feet, and many have considerable difficulty not to act in accordance with that feeling. According to Gendler, we *believe* that we are safe, but we *alieve* (something like) “Really high up, long, long way down. Not a safe place to be! Get off!!” (“Belief” 635). In the case of the Skywalk example and other Gendler cases, aliefs are *belief-discordant*: there is a discrepancy between the subject’s aliefs and (explicit) beliefs.

Gendler provisionally characterizes aliefs as “associative, automatic, and arational”; as “states we share with nonhuman animals”; as “developmentally and conceptually antecedent to other cognitive attitudes”; and as “typically also affect-laden and action generating” (“Belief” 641; italics in original). “A paradigmatic *alief* is a mental state with associatively linked content that is representational, affective, and behavioral, and that is activated—consciously or nonconsciously—by features of the subject’s internal or ambient environment” (642). In such a state, the subject alieves *R-A-B*, in which *R* is “the representation of some object or concept or situation or circumstance,” *A* is “the experience of some affective or emotional state,” and *B* is “the readying of some motor routine” (643). Aliefs can be occurrent or dispositional. The characterization of a “paradigmatic alief” as an activated mental state characterizes an *occurrent* alief. A subject has a *dispositional* alief with content *R-A-B* “when there is some (potential) internal or external stimulus such that, were she to encounter it, would cause her to occurrently alieve *R-A-B*” (645).

James Baillie explains the *samvegic* experience of “the expectation of nothingness” as involving an alief that I (i.e., the alieving subject) will

(some day) die. He distinguishes three kinds of mental states, *A-states*, *N-states*, and *E-states*, that differ in the nature (and content) of the belief-like mental states involved. In an *A-state* the subject's belief that she will die is dispositional. This is our ordinary state in which we are absorbed in everyday tasks and do not think about our mortality, but in which we would admit that we are mortal when asked. At *that* point we would enter an *N-state*. In an *N-state* the subject occurrently believes that she will die. What distinguishes an *E-state* from an *N-state* is that in an *E-state* a belief-concordant *alief* that I will die is added. This, however, is only part of Baillie's analysis, as he further observes a difference in the content of the aliefs and beliefs that I will die, namely: the "I" of the alief is not the "I" of the belief.

The "I" of the belief that I will die is (more or less) Thomas Nagel's "objective self." It is the publicly accessible person referred to by me in the first person singular that is a part of objective reality; it is a somewhat abstract and relatively emotionally neutral "I." The "I" of the alief, on the other hand, is what Mark Johnston calls "the one at the center of this arena of presence and action" (139); it is the subjective "I." While the objective "I" of the belief is a perceived "I" (or "me"), the subjective "I" of the alief is the subject of perception (i.e., that what perceives) itself. And while the belief is about an objective future world minus the human being that is me, the alief is about the cessation of the subjective world: the alief that I die represents the fact that *my* world (i.e., this arena of presence and action) will cease to exist. In other words, contrary to the belief, the alief confronts me with an absolute and unavoidable finality, with the fact that it does not even make sense to speak of "nothingness" beyond death because there is no "beyond" death.

Beliefs about the objective "I" or self give rise to the "'not me' phenomenon." Arguments about my death are framed by my mind in

terms of the publicly accessible, objective “I,” and “this restriction creates the ‘not me’ phenomenon, where even the explicit avowal of the belief that I will die is kept at a distance, as about someone in the objectified world, who is seen as accidentally related to . . . the subjective ‘I’” (S200). Because of this phenomenon, “*N* is an essentially dissociated condition” involving “both believing that I will die, and also *not* believing it,” and states *E* and *N* differ in “the levels of psychological integration involved” (S201).

Although Baillie remains unclear about the exact subject of the dissociative belief, his theory seems to imply that the “I” involved is the subjective self. In addition to this ambiguity, he offers no argument for why this state would have to be a *belief* rather than an *alief*, which appears to be at least equally likely. Taking these ambiguities into account, Baillie’s theory can be summarized as follows:

In all three states—*A*, *N*, and *E*—the subject has a dispositional belief that the objective self will die, and a dispositional belief (or alief) that the (subjective) self will not die.

In the *N*-state, the subject has an occurrent belief that the objective self will die (realizing the first of the two dispositions).

In the *E*-state (i.e., *saṃvega*) the subject has an occurrent alief that the subjective self will die, temporarily suppressing the dispositional belief (or alief) to the contrary (i.e., the second of the two dispositions).⁸

⁸That the *E*-state only suppresses the dispositional belief (during that state) in the immortality of the subjective self and does not erase it, is implied by the fact that *saṃvega* is an occurrent state of relatively brief duration (see criterion a in the rough definition

In addition to the ambiguity of the dissociative belief in the *N-state*, there are some other weaknesses in Baillie's account. Mostly these appear to be related to Baillie's framing of the issue as epistemological rather than psychological. He attempted to explain how one could be shocked by what one already *knows*, rather than how one can be *shocked* by what one already knows. As a consequence of this framing, his perspective is exclusively informed by philosophical literature and ignores relevant psychological theories. The most obvious example hereof is the unnecessary adoption of Gendler's controversial (philosophical) theory of alief,⁹ while (psychological) dual process theories offer a much more established alternative that can be used toward the same explanatory end, as will be shown in the next section.

The epistemological framing of the issue also produced another important oversight. The focusing on the shock of the *known* takes that knowledge for granted. That is, the death-affirming *A* and *N-states* are normal *by definition* because of this framing. However, while they may indeed be normal states for the average, non-religious, academic philosopher, for the vast majority of human beings they are not normal states at all. The normal state—according to anthropology or sociology of religion, for example—is not characterized by a belief that I will die, but by the very *denial* of that fact, by what Ernest Becker famously called “the denial of death.” We'll return to this issue below.

given in the introduction) and that after the experience the subject returns to the *A* or *N-state*.

⁹See, for example, Schwitzgebel; Mandelbaum; or Albahari.

Dual Process Theory and *Samvega*

Since the 1980s, a large body of theoretical and empirical research in social psychology and related fields of inquiry has resulted in (and has been guided by) *dual process theories* of (aspects of) human cognition and decision-making. The defining characteristic of these theories is that they classify mental processes into two classes, one of which is (more) automatic, while the other is (more) controlled. Although there is considerable disagreement between these theories about the details (hence the plural), the core idea that there are such different mental processes or “systems” is supported by an abundance of evidence and has become a cornerstone of contemporary social psychology.¹⁰

Much of the disagreement between adherents of different dual process theories concerns the nature, scope, and workings of the two systems. Some are limited in scope in the sense that they just apply to certain processes while others are more general. Some propose a radical divide between the two systems, while others suggest that they are mere extremes on a gamut, or that one is realized in the other. The most conspicuous difference between theories, however, is terminological: there are no universally accepted labels for the two systems.¹¹ I’ll adopt the relatively neutral “system 1” and “system 2” (introduced by Stanovich and West) here.

¹⁰For a recent overview and discussion of dual process theories, see Gawronski and Creighton. For a less academic, more accessible, and well-written introduction into one particular dual process theory, see Kahneman (*Thinking*).

¹¹Typically the two systems are named by means of pairs of adjectives. These adjectival name-pairs include heuristic vs. systematic, experiential vs. rational, associative vs. rule-based, associative vs. propositional, reflexive vs. reflective, and impulsive vs. reflective.

The two systems are often described by means of contrasting keywords. For example, system 1 thinking is (usually) faster, while system 2 thinking is (usually) slower. In addition to the fast/slow contrast, other contrasts often mentioned are automatic/controlled (often considered to be the defining contrast), associative/rule-governed, unconscious/conscious, slow-learning/flexible, effortless/effortful, concrete/abstract, emotional/neutral, arational/rational, and parallel/serial. Different dual process theories do not agree which of these contrasts apply, however.

Although there are exceptions, philosophers typically assume that (their) thinking is characterized by the keywords mentioned for system 2 (i.e., those mentioned after the slashes in the previous paragraph), but psychological research has shown that people—including philosophers—mostly depend on system 1, and that system 2 is activated only in case system 1 cannot find a satisfactory solution to the problem at hand. Although system 1 is efficient and gets it right most of the time, it has its biases and limitations. For example, in *Thinking, Fast and Slow*, Daniel Kahneman writes that system 1 “creates a coherent pattern of activated ideas in associative memory” and towards that end, “neglects ambiguity,” “suppresses doubt,” “invents causes and intentions,” and is “biased to believe and confirm” (105).

Among the keywords describing system 1 (i.e., those before the slashes two paragraphs back) are “associative,” “automatic,” and “arational.” These are exactly the same keywords Tamar Szabó Gendler used to characterize aliefs.¹² Hence, aliefs appear to be the system 1 equivalent of beliefs, but whether this identification makes sense depends on one’s

¹²The similarity between Gendler’s belief/alief model and (some) dual process theory (or theories) was previously observed in Kriegel and J. Nagel.

theory of beliefs and on where and how one draws the boundary line between systems 1 and 2, and neither of these is uncontroversial.

Belief-like states are mental states (possibly in a loose sense of “state”) that are associated with representational or reconstructive contents and behavioral dispositions.¹³ It is often assumed that beliefs are propositional, but belief-like states do not just include believing that *p*, but also taking *p* for granted, unconsciously assuming that *p*, consistently acting as if one believes that *p*, as well as various other belief-like attitudes towards *p*. Furthermore, *p* in “believing (that) *p*” and so forth is a propositional *representation* of whatever is believed by the subject—in this loose sense of believing—but does not necessarily exhaust the content of that belief. What someone believes may be much richer and/or much more indistinct than any propositional representation can capture. If we assume that both systems have beliefs in this loose sense (we’ll turn to the main alternative below), then the beliefs of system 1 are associative, prototypical (or exemplar-based), and directly based on experience, while the beliefs of system 2 involve more reflection (on experience) and reasoning, and may be more abstract(ed).

The Skywalk case (see previous section) and other examples of belief-discordant aliefs given by Gendler can just as well—if not better—be explained as conflicting beliefs of system 1 and system 2. In the Skywalk case system 1 believes there is danger; system 2 doesn’t. And in the case of unconscious racism (Gendler “Epistemic”), the subject’s system 2 believes that all races are equal, but system 1 is biased against certain races. In either case, the automaticity of system 1 is what makes it hard

¹³Although according to mainstream analytic philosophy beliefs are representational, the notion of representation and its role in mental states and contents is controversial. To me, the notion reeks too much of what Donald Davidson called, “the essentially incoherent picture of the mind as a passive but critical spectator of an inner show” (52). The mind does not represent but reconstruct or simulate.

to control its influence on behavior, and the automatic/controlled contrast defines the two systems in many (if not most) dual process theories. Similarly, the aliefs and beliefs involved in Baillie's theory of *samvega* (or the "expectation of nothingness") can be reinterpreted as being beliefs of systems 1 and systems 2, respectively.

An immediate advantage of this reinterpretation is that it solves an ambiguity in the original theory. The ambiguous state at the root of the "not me" phenomenon was represented as the dispositional *belief* that the subjective self will not die (in the *A* and *N-states*), but is more plausibly attributed to system 1 than to system 2. The subjective self or "I" is what Mark Johnston called "the one at the center of this arena of presence and action" (see above). It is (at) the center of subjective reality, and the limits of the subjective self are the limits of subjective reality. The subjective self is an experiential notion; *all* experience—by necessity—is experience of the subjective self (because all experience takes place in the arena of which the subjective self is the center). The subjective self is there in *all* experience, and cannot not be there, or even imagined not to be there (or to not be). The notion of the subjective self is experiential, automatic, and automatically reinforced by continuing experience. In other words, it is a paradigm case of a system 1 notion.¹⁴ And

¹⁴Daniel Kahneman (*Thinking* chapter seven) points out that system 1 only makes use of activated/retrieved ideas. System 1 is biased to assume that "what you see is all there is." Importantly, the experiencing self is always there. It isn't literally *seen*, of course, but is seeing, and makes its presence felt in the act of seeing. What *I* see is all there is, but *I* am always there, seeing, and "seeing" myself. Kahneman (*Thinking* part V) also suggests a subtly different conceptualization of the two selves or "I" notions involved. He distinguishes the *experiencing self* from the *remembering self*. The experiencing self is (more or less) Johnston's "center of the arena" or Baillie's "subjective 'I'." The remembering self is a construction by system 2, but is heavily dependent on system 1, and the distinction between the two selves in general does not coincide with the two systems. The remembering self constructs one's life story; it is in that sense very similar to the notion of the "narrative self" proposed by Daniel Dennett and others, and further de-

likewise, the unavoidability of the subjective self in all experience generates and continuously reinforces the system 1 belief that the subjective self will always be there. A *samvegic* shock occurs when system 1 suddenly (i.e., occurrently) believes in its own finality: the arena *will* cease.

Amending the previous section's summary of Baillie's theory in accordance with the foregoing (i.e., the dual process reinterpretation and the reclassification of the dispositional death-denying belief) results in the following:

In all three states—*A*, *N*, and *E*—the subject has a dispositional system 2 belief that the objective self will die, and a dispositional system 1 belief that the subjective self will not die.

In the *N-state*, the subject has an occurrent system 2 belief that the objective self will die (realizing the first of the two dispositions).

In the *E-state* (i.e., *samvega*) the subject has an occurrent system 1 belief that the subjective self will die, temporarily suppressing the dispositional belief to the contrary (i.e., the second of the two dispositions).¹⁵

veloped in Narrative Identity Theory within personality psychology. It is this constructed, narrative self that plays the "1" role in the system 2 beliefs in Baillie's theory: those beliefs concern the death of my narratively constructed, remembered self.

¹⁵The occurrent system 1 belief only temporarily suppresses the dispositional system 1 belief in the immortality of the subjective self, however, and does not (and probably cannot) erase it, because that dispositional belief is continuously reinforced by experience. It is for this reason that *samvega* has a relatively short duration (see criterion (a) in the introduction of this article).

Obviously, this interpretation depends on the assumption that both systems 1 and 2 are capable of having beliefs, but as mentioned above, that assumption is controversial. Superficially, it may seem undeniable that both systems must have (or be able to be in) belief-like states in the loose sense of beliefs adopted here. Certainly Kahneman's depiction of the two systems requires the attribution of belief-like states or attitudes to *both* systems, and the common idea that system 1 is prone to certain biases—such as a bias to infer or invent causal patterns (i.e., causal beliefs!)—and that system 2 is capable (in principle) of correcting these makes sense only if both systems can have or be in belief-like states. Nevertheless, there is reason to suspect that at least one of the two systems does not have sufficiently belief-like states.

Based on extensive reviews of available psychological evidence, Peter Carruthers concludes that system 2 is incapable of having beliefs. System 1 is a heterogeneous class of unconscious processes; system 2 is conscious and takes place in working memory. However, evidence shows that all conscious thought is sensory-based, and that all amodal thought—such as abstract and conceptual thought—is unconscious. This means that beliefs, which are supposed to be amodal, cannot be conscious, and therefore, that only system 1 can have beliefs; and that what *appear* to be system 2 beliefs are really episodes of inner (or outer) speech (or other kinds of sensory—and therefore, modal—states).¹⁶

Carruthers's conclusions are as problematic for Baillie's alief-based theory of *saṃvega* as they are for the dual process reinterpretation

¹⁶The relevant writings by Carruthers are *The Opacity of Mind*, in which he argues that our minds are not transparent to ourselves, or at least not nearly as transparent as we'd like to believe; "On Knowing your Own Beliefs" in which he argues most explicitly that system 2 cannot have beliefs; and *The Centered Mind* in which he argues that beliefs and decisions are unconscious and that all conscious thought is sensory-based (and thus not amodal).

thereof a few paragraphs back (as well as for much of the philosophy of mind and some other branches of philosophy, but that is of no concern here) because both assume that we have conscious amodal beliefs. The *N-state* is defined as involving a conscious amodal belief that the objective self will (eventually) die, but if there are no conscious amodal states, then this is impossible. Nevertheless, this doesn't necessarily mean that the belief-like states involved in Baillie's theory don't exist; what it does mean is that some of them may be (considerably) less belief-like than thus far assumed.

The dual process reinterpretation of Baillie's theory involves four different supposedly belief-like states: (D2) a dispositional system 2 belief that the objective self will die (in all three states); (D1) a dispositional system 1 belief that the (subjective) self will not die (in all three states, albeit suppressed in *E*); (O2) an occurrent system 2 belief that the objective self will die (in the *N-state*); and (O1) an occurrent system 1 belief that the subjective self will die (in the *E-state*). Two of these, (O2) and (D2), are supposed to be system 2 beliefs. (O2) is the occurrent counterpart of (D2), or in other words, (D2) is the disposition to have the occurrent state (O2) in the right circumstances. (O2) cannot be a belief-state, however, because occurrent, conscious "beliefs" are not beliefs, but sensory-based mental events such as episodes of inner speech (i.e., silently talking to oneself). Hence, (O2) is an event of telling oneself (and/or others) that one will (eventually) die, and (D2) is the disposition to do so (in the right circumstances). The disposition (D2) is itself unconscious, and thus not a system 2 state, even if it is a disposition to produce system 2 events.¹⁷

¹⁷This disposition (D2) can have various origins—it can be rooted in episodic memories of earlier occurrences of (O2), in an episodic memory of telling oneself that one should believe in one's own mortality (because this "belief" follows from or best coheres with

The other two belief-like states, (D1) and (O1), are—supposedly—states of system 1, which can have (amodal) beliefs. (D1) is an unconscious belief of system 1 that the self (in some relevant sense) does not die. As argued a few paragraphs back, this belief is reinforced by continuous experience. (O1) is the belief-like state that defines *samvega*, and as *samvega* is a conscious experience, (O1) must be conscious as well. However, system 1 states are unconscious, and consequently, (O1) cannot be a system 1 state, but if (O1) is a system 2 state (which is the only alternative), then it is too similar to (O2) to explain *samvega*. The solution to this problem is to give up the implicit and unwarranted assumption that (O1) is *wholly* conscious. More plausibly, (O1) is a complex state that involves both an unconscious system 1 belief, and conscious system 2 events. The system 1 belief is an occurrent and fleeting (D1)-suppressing belief in one's own mortality. The actual, conscious experience of *samvega* is the system 2 counterpart of the unconscious mental stress resulting from the suppression of (D1) and of the occurrent system 1 belief in the mortality of the subjective self.¹⁸

The foregoing can be summarized in the same format as Baillie's theory and the both-systems-have-beliefs dual process reinterpretation presented above as follows:

In all three states—*A*, *N*, and *E*—the subject has a disposition to consciously affirm its own mortality (in the form of an inner/outer speech event), and an unconscious, dis-

one's other "beliefs"), and/or in an unconscious belief that one should believe in one's own mortality.

¹⁸It seems likely that causality runs both ways and/or that there are complex interactions between the occurrent system 1 belief or beliefs in the mortality of the subjective self and the occurrent, conscious mental events of system 2. For example, Buddhaghosa's suggestions on how to achieve *samvega* suggest that (O1) is caused by conscious—and therefore, system 2—events.

positional system 1 belief that the subjective self will not die.

In the *N-state*, the subject has a conscious (system 2) event realizing the first of these dispositions.

In the *E-state* (i.e., *saṃvega*) the subject has an occurrent, unconscious (system 1) belief that the subjective self will die, temporarily repressing the contrary dispositional belief (i.e., the second of the two dispositions), as well as conscious (system 2) events caused by the ensuing unconscious (system 1) mental stress.

A comparison of the two dual process theories of *saṃvega* (thus summarized) points at three important differences. Firstly, in the second theory, unconscious processes play a much more central role than in the first. Secondly, where the first theory speaks of system 2 beliefs, the second speaks of system 2 episodes of inner/outer speech. And thirdly, in the second theory the *E-state* or *saṃvega* involves both systems rather than just system 1. Although these differences are important, they do not entail a *radical* difference between the two theories. The jury may still be out on which variety of dual process theory is correct (although I think that Carruthers's theory is more extensively supported than any of its competitors), but both varieties can accommodate *structurally identical* theories of *saṃvega*—that is, theories that involve the same states with sufficiently similar “contents.” In either case, *saṃvega* or the *E-state* is characterized by a terrifying system 1 belief that the subjective self will die, which temporarily suppresses a death-denying dispositional system 1 belief, while system 2 is already disposed to affirm the subject's mortality. Nevertheless, while this gives us a reasonably well-developed and empirically grounded theory of what kind of state *saṃvega* is, it does not yet address the second weakness in Baillie's theory mentioned above, and more importantly, it offers no clues on how *saṃvega* produces its

morally motivating effects. These are among the topics of the following sections.

States of Denial and Transcendence

Contrary to Baillie's suggestion, for most people the death-affirming *A* and *N*-states are not the normal states. Rather, most people deny their deaths. Of course, they don't deny their *biological* deaths, but most people believe that their essential, subjective self (i.e., their soul, spirit, etc.) will somehow survive death—that is, most people believe in some kind of afterlife, rebirth, or some other form of essential immortality. Ernest Becker called this kind of belief the “denial of death.”

By implication, there is a state preceding the *A*-state and its occurrent counterpart *N*, namely a state of denial, which I shall call the *D*-state. In the *D*-state the subject has a dispositional system 2 belief that the self (in some relevant sense) will not die, or has a disposition to consciously deny its own mortality. Like the *A*-state, the *D*-state has an occurrent counterpart, *D*^{*}, in which the subject realizes this disposition by denying death.

The *A* and *N*-states are the secular counterparts of the more common *D* and *D*^{*}-states, but unlike *A* and *N*, *D* and *D*^{*} do not involve dissociation. According to Baillie, “*N* is an essentially dissociated condition” involving “both believing that I will die, and also *not* believing it,” but this dissociation is wholly absent in *D*. The greater “psychological integration” that *E* is supposed to establish is already established in *D*. Of course, this psychological integration comes at the cost of rejecting or disbeliev-

ing the *fact* of mortality,¹⁹ but that matters little to the subject. Furthermore, that *E* involves greater psychological integration is plausible only if after experiencing *saṃvega* the subject does not revert to the dissociative *A* or *N*-state, and that is rather unlikely.

According to Seymour Epstein (“Cognitive,” “Integration”), system 1 beliefs (or beliefs of the *experiential* system in his terminology) (can) change under the influence of repeated experience.²⁰ This raises the question whether there could be a state—let’s call it the *B*-state—that is brought forth by repeated experience of *saṃvega* (i.e., the *E*-state). The *B*-state would be the dispositional counterpart of the occurrent *E*-state (in the same way that *A* and *D* are the dispositional counterparts of *N* and *D**, respectively). In the *B*-state the subject would no longer have the dispositional death-denying system 1 belief that is present in *all* other states, but would have replaced that by a belief in the mortality of its subjective self, thus transcending the dissociation of states *A* and *N*. Considering that Baillie and others (see Baillie for examples) describe the *E*-state as involving a kind of panic and terror that drives out all capability of rational thought, if being in the *B*-state would be anything like being in the *E*-state, then the price of transcending dissociation may very well be that of leaving behind sanity. However, this would assume that a dissociative, occurrent state is representative (at least in the relevant respects) for a non-dissociative, dispositional state, and that assumption stands in need of independent support. If the panic and terror in *E* result from the *sudden* and *unexpected* realization of one’s mortality, then there is no ground

¹⁹For a thorough review of theories of “surviving death” and why they fail—thus establishing that it is an undeniable fact that we are mortal—see Johnston. See also Brons.

²⁰Epstein and Pacini write that the experiential system 1 “changes with repetitive or *intense* experience” (463; emphasis added). *Saṃvega* certainly seems to qualify as an intense experience.

for panic and terror in the *B-state*, and consequently, *B* would be nothing like *E*.

A more fundamental question than what it would be like to be in the *B-state* is whether that hypothetical state is even possible. As mentioned above, the death-denying system 1 belief (which is present in *all* other states) is continuously reinforced by experience. *All* experience is experience of the subjective (or experiencing) self, reconfirming its existence and—more importantly—its existential coincidence with experienced reality. It does not seem plausible that an occasional experience of *samvega* can compete with that. Hence, while Baillie’s theory and/or its dual process reinterpretations must certainly be extended in one direction—by adding the death-denying *D-state* and its occurrent counterpart—there probably are no other additional states.

On the Possibility of Other Causes of *Samvega*

Buddhaghosa (IV.63) mentions eight grounds of *samvega*—birth, aging, sickness, death, suffering of loss, past suffering caused by rebirth, future suffering caused by rebirth, and present suffering caused by the search for food—but significantly, meditation on death is the *only* exercise that is (explicitly) intended to bring about *samvega*. Even if the other seven are “grounds” (*vatthu*) of *samvega*, apparently they are so to a lesser or different extent, but that is to be expected if the theory of *samvega* presented above is right. *Samvega* then requires a stable, dispositional belief of system 1 and a terrifying, occurrent system 1 belief that contradicts this dispositional belief. This scenario, summarized below, is specific enough to suspect that it does not have very many instances.

In all states, the subject has a dispositional system 1 belief that *p*.²¹

In the *D* and *D**-states, the subject has a dispositional system 2 belief that *p*, or a disposition to consciously affirm that *p*.

In the *D**-state, the subject consciously realizes this second disposition.

In the *A*, *N*, and *E*-states, the subject has a dispositional system 2 belief that not-*p*, or a disposition to consciously affirm that not-*p* (or deny *p*).

In the *N*-state, the subject consciously realizes this third disposition.

In the *E*-state, the subject has an occurrent system 1 belief that not-*p*, temporarily suppressing the first dispositional belief.

James Baillie sees only two plausible scenarios similar to the “expectation of nothingness” (i.e., *saṃvega*) and he is somewhat skeptical about the second. The first is the recognition of the fact that I might not have existed. The fact that my existence is radically coincidental is not nearly as threatening or terrifying as the fact of my mortality, however. That my existence is coincidental cannot undo that existence, but death *will* do so.

²¹“*p*” should not be read as a placeholder for a single specific proposition here, but as a placeholder for some member of a loosely bounded set of very similar propositions, which includes both the proposition that the *subjective* self will die and the proposition that the *objective* self will die, for example.

The second case Baillie considers is that of the cessation of some *other* “arena of presence and action” that is close to mine. This suggestion is problematic for a number of reasons. Firstly, I have no system 1 beliefs (or aliefs, in Baillie’s theory) about the subjective self of the other. There is no level of intimacy that gives me access to (the thing at) the center of the other’s arena. My beliefs about that arena—and thus about the other’s subjective self—are inferred beliefs by *my* mind. Neither do I have system 1 beliefs about the (necessary, continuing) presence of the other of similar strength as my system 1 beliefs about the (necessary, continuing) presence of myself. “I” am always there (seeing, experiencing), but the other is not, and was not always there either.²²

Secondly, it is far from obvious that what shocks or terrifies me in or about the sudden realization of the inevitable death of the other that I am so intimate with (assuming that it does) is the cessation of “that unique point of view” (S202). Alternative interpretations of such a shock (if it occurs) are that it is a selfish shock born from the realization that the other will cease to be present in *my* arena, or that it is an empathic shock in which I feel (something like) what I envision the other to feel if the other would be in the *E-state*. In the latter case—let’s call it *empathic samvega*—I am experiencing *samvega* (in diluted form, perhaps) on behalf of the other: the envisioned experiences of the other are envisioned as one’s own. By implication, empathic *samvega* involves the same kinds of beliefs (or other mental events), and therefore, is not really a *different* kind of *samvegic* state. (And I doubt that “selfish shock” about the other’s death, the other alternative interpretation, can be fitted in the pattern summarized above.)

Given Baillie’s oversight of the “denial of death” (i.e., the *D-state*), it is no surprise that his suggestions for alternative *samvegic* scenarios

²²This may be different for conjoined twins, but I’ll ignore that rather uncommon case.

start from the dissociative *A-state*, but Buddhist thought suggests another case that—at least potentially—involves all the states distinguished above: the realization (in the *E-state*) that there is nothing, no *thing*, at the center of the arena. Although this insight is gradually gaining ground in Western science and philosophy, it remains controversial outside Buddhist philosophy.²³ I'll ignore that controversy here, and will assume that it is true that there is no thing at the center. The default belief, however, is that there *is* something at the center of my “arena of presence and action,” and that that something is me or my “self.” That is, the default belief (in the *D-state*) is that there are self-defining essences. Contrary to that belief, Buddhist philosophers, but also John Locke and David Hume, for example, argued that there is no single part of a self that must remain the same for someone to count as the same person. Selves are just bundles of experiences (Hume) bound together by memory (Locke), psychological continuity and connectedness (Parfit *Reasons*), or the stories we tell about ourselves (e.g., Dennett). Hence, rather than the thing *at* the center, the self is the center itself: it is the point of view and the “center of narrative gravity” (Dennett); and the notion of a “self” is—in Buddhist terms—nothing but a “convenient designator.”

²³Notwithstanding the mainstream belief in essential selves, there are quite a few prominent defenders of a no-self view, including Parfit (*Reasons*); Dennett; and Metzinger. An interesting question is *why* the belief in essential selves is so strong in Western thought. Obviously, that question is outside the scope of this article, but there is a suggestion by Mark Fisher and Jeremy Gilbert that is worth mentioning. They argue that submission to the dehumanizing work environments typical of capitalism can only be legitimated and made tolerable by the idea of an essential, authentic self that is isolated from and unaffected by the person-as-worker. Charles Lindholm makes a similar point, but traces the history of the belief in an authentic self further back (and better documents his case). If they are right, then the belief in a self-defining essence is ideology (in the Marxian sense of that term), which would explain the belief's pull and pervasiveness.

In the *A-state* and its occurrent counterpart *N*, the subject “knows” all that. That is, system 2 realizes and understands that there is nothing at the center (or is disposed to say so), but system 1 does not: system 1 continues to believe that there is something, some *thing*, at the center and that that thing is its essential being. This belief is originated in and reinforced by the biases of system 1, particularly by its bias to construct an easy story based on nothing but available, activated “evidence” (i.e., experience). *I* am always there in my experience; therefore, *I* am; therefore, there is some thing that is me at the center of my subjective universe (my arena). In the *E-state*, which can be brought about by meditation, but perhaps also in other ways, there is an occurrent system 1 belief that contradicts this dispositional system 1 belief in the self.

Thus far, this alternative scenario seems to fit in the above-summarized framework of states involved in *samvega* perfectly. The only thing missing is that the *E-state* is a state of shock (that is what “*samvega*” literally means), and that this state is religiously and morally motivating. There appears to be some evidence for both, however. Miguel Farias and Catherine Wikholm cite statistical and anecdotal evidence for the occurrence of anxiety and other shock-like states as a result of meditation. And in a psychological study of the effects of meditation retreats in Thailand, Tipawadee Emavardhana and Christopher Tori point out that “the thought of selflessness can lead to severe anxiety” (195) and is even potentially dangerous if not guided appropriately. Emavardhana and Tori also found that the experience significantly and positively affects religiosity (i.e., is religiously motivating).

Whether the anxiety and potential danger Emavardhana and Tori perceive is similar to the shock of *samvega* is debatable, however. They warn that “if the deconstruction of the ego can occur within a structured environment, which provides supportive philosophical reasons for the attainment of selflessness, this process can be safe for those without se-

rious psychiatric problems” (203), but this is not a warning for *saṃvega*, but for a situation in which the subject has not progressed beyond the state of denial *D* and suddenly jumps to the occurrent belief typical of *E*. It is a warning for a state *E** (rather than *E*, which follows *A* rather than *D*) in which there is no rational framework of beliefs to explain the shocking system 1 experience. In case of *saṃvega*, however, the *E*-state follows the *A*-state (or *N*), and therefore, the subject’s system 2 has *already* accepted the falsity of the claim that there is a self or that I will not die, and the shock in the *E*-state is caused *only* by an occurrent belief of system 1 that (more or less) conforms with what the subject (’s system 2) already professes to believe. Moreover, Emavardhana and Tori’s claim that “the deconstruction of ego” is “safe” if backed by “supportive philosophical reasons” (etc.) appears to mean that if the *N*-state is achieved first, then the occurrent system 1 belief that is typical for state *E* or *E** does not lead to shock (or “severe anxiety”). In other words, if an *E*-like (i.e., *E* or *E**) state follows *N* as in the *saṃvegic* pattern, then there is no shock.

The anecdotal evidence of anxiety and shock resulting from meditation presented by Farias and Wikholm is also better explained as jumping from *D* to *E** than as *saṃvega*. They describe cases of adherents to pop-Buddhism or “mindfulness” in search of their true selves who are confronted with the unexpected system 1 belief that there is no self. In other words, in these cases, the subject has a firmly held belief in *p* (i.e., is in state *D* or *D**) and then—in meditation—has an experience that leads to a system 1 belief that not-*p*. The state of shock and anxiety is caused by the sudden disruption of a firmly held belief in the self, necessitating a revision of a subjectively important part of the subject’s worldview. In case of *saṃvega*, on the other hand, the shocking occurrent belief only conflicts with system 1 and is already fully accepted—and perhaps even anticipated—by system 2 (or by the subset of the subject’s beliefs that are routinely expressed and/or affirmed by system 2).

Another way to look at this difference is in terms of dissociation versus psychological integration. In the meditative shock cases, the subject moves from a non-dissociative (integrated) *D* or *D*^{*}-state to a dissociative *E*^{*}-state, but in case of *saṃvega* it is the other way around: the subject moves from a dissociative *A* or *N*-state to a non-dissociative (integrated) *E*-state. It should be clear that these are two very different kinds of shock, and therefore, that even if the meditative experience of no-self is or can be shocking and religiously motivating, it does *not* involve *saṃvega*.

Perhaps, it is no coincidence that the only technique Buddhaghosa advises to achieve *saṃvega* is meditation on death. *Samvega* is the sudden, shocking realization by system 1 of its own mortality *after* the subject has already acknowledged that fact. That is a very specific scenario indeed, and it seems likely that no other set of beliefs can have the same effects.

***Samvega* and Moral Motivation**

Moral motivation is a motivation to do the right thing, but as Michael Smith has pointed out, “the right thing” is ambiguous—it can be read both *de dicto* and *de re*. Read *de dicto*, being motivated to do the right thing is being motivated to do what is supposed (or said, or believed) to be right; read *de re*, it is being motivated to do what really *is* right. To clarify the difference, suppose there is a culture that values fierceness and in which violence of men towards women is morally right.²⁴ Further suppose that such violence really is morally wrong and that it really is right to protect women (as well as other people) from harm. Then, a

²⁴According to Chagnon, the Yanomamö are such a culture, but this claim is controversial.

male member of that culture who is motivated to do the right thing *de dicto* is motivated to be violent to women, but a male member of that same culture who is motivated to do the right thing *de re* is *not* motivated to be violent to women and to protect them from harm instead.

The notion of *de re* moral motivation presupposes that there are things that “really are right.” Cultural relativists (as well as subjectivists) would reject this, of course, and claim that there is only *de dicto* rightness (i.e., things that are said to be right by some belief system or culture). And even among moral theorists that agree that there are things that “really are right” there is widespread disagreement about what exactly is right—ethical egoists, utilitarians, and virtue ethicists, for example, have rather different opinions about this. Furthermore, rightness relative to a moral theory is not *de re* rightness, but just another variety of *de dicto* rightness. *De re* rightness is not necessarily the same as moral objectivity, however—there is no *prima facie* reason to believe that intersubjectivity is insufficient. And if there are things about which virtually everyone would agree that they are right (given sufficient time for reflection if necessary), then those are “really right” in the here relevant sense.

There are at least two plausible candidates for intersubjective general principles of *de re* rightness. Firstly, in most influential moral theories in both Western and Asian philosophical traditions goodness tends to be associated with (moderate) altruism, and moral psychologists also often understand moral motivation as a motivation to behave pro-socially (i.e., altruistically). And secondly, Derek Parfit (*Matters* 565-569) points out that virtually everyone agrees that undeserved suffering is bad. Although this may be sufficient for intersubjective wrongness, it seems to me that a much stronger claim can be made about the moral status of suffering (in a broad, non-technical sense that may overlap with *dukkha* but is not identical to it): all suffering is inherently, objectively (and not just intersubjectively) bad. Anyone who denies this either

does not understand what suffering is or what “bad” means.²⁵ If the foregoing is right, *de re* moral motivation is (or includes) the motivation to be moderately altruistic and/or to reduce suffering (and conversely, a *de re* motivation to do the wrong thing would be a motivation to be egoistical and/or to increase suffering).

The *de re/de dicto* distinction is not just of theoretical interest, but matters here because *samvega* is related both to a motivation to do the right thing *de dicto* and to a motivation to do the right thing *de re* (understood as above), albeit in very different ways. And interestingly, the different ways in which different kinds of moral motivation are produced are related to the different kinds of meditation distinguished in this article’s introduction. As mentioned there, meditation practices are often classified as belonging to one of two main kinds: *samatha* and *vipassanā* meditation. The aim and purpose of *samatha* or tranquility meditation, which is associated with mindfulness, is to achieve states of concentration in which all attention is focused on its object. The aim and purpose of *vipassanā* or insight meditation is to gain and deepen understanding of impermanence, no-self, and suffering.

The meditation on death as described in Buddhaghosa’s *Visuddhimagga* does not fit well in this scheme, because it can be understood as being of the first as well as of the second kind. If the goal of meditation on death is mindfulness of death, then it is a kind of tranquility meditation; if its goal is to gain or deepen insight into death as impermanence and a cause of suffering, then it is a kind of insight meditation. Perhaps then, it is both.

²⁵Parfit (*Matters* 565-569) argues for this claim on the grounds that suffering is never deserved. From that premise, and the premise that undeserved suffering is bad, it follows that all suffering is bad.

Buddhaghosa writes that the meditation on death should lead to mindfulness of death and *saṃvega* (VIII.5-6), suggesting that these are distinct goals, and therefore, that *saṃvega* is not mindfulness of death, but is related to some kind of insight (if the tranquility/insight classification is exhaustive). That *saṃvega* is not mindfulness of death is also implied by the suggestion that the latter should lead to the former in the last sentence of III.58. Conversely, one would expect that *saṃvega* leads to some kind of mindfulness of death. Regardless of how exactly *saṃvega* and mindfulness of death are causally related, these (and other) passages in Buddhaghosa's *Visuddhimagga* strongly suggest that the two come together (or should come together, at least). Hence, *saṃvega* may not be mindfulness of death, but is inseparable from it, and is simultaneously related to a relevant kind of insight. Terror Management Theory explains how mindfulness of death produces *de dicto* moral motivation. *Saṃvega*-related insight, on the other hand, produces *de re* moral motivation.

Terror Management Theory (TMT) is a psychological theory based on Ernest Becker's writings about the denial of death (see above).²⁶ According to Becker, much of civilization (religion especially) is a defense mechanism against the fear (or "terror") of death. Civilization is an "immortality project"—it allows me to become part of and contribute to something that not just survives my biological death, but that seems to offer the promise of eternal life. TMT took many of Becker's ideas, turned them into testable hypotheses, and then put them to the test. According to TMT, "the awareness of death gives rise to potentially debilitating terror that humans manage by perceiving themselves to be significant contributors to an ongoing cultural drama," and "reminders of

²⁶For short introductions into TMT, see Greenberg and Arndt or Kesebir and Pyszczynski. For a non-academic, book-length introduction, see Solomon, Greenberg, and Pyszczynski.

death increase devotion to one's cultural scheme of things" (Solomon, Greenberg, and Pyszczynsky 211). Hence, much of what we (humans) do and believe is driven by "terror management," controlling the fear of death, and "effective terror management is faith in a meaning providing cultural worldview and the belief that one is a valuable contributor to that meaningful world" (Greenberg and Arndt 403). In other words, reminding people of their mortality—or increasing "mortality salience"—leads them to bolster both their worldviews and their beliefs that they are valuable contributors to the world according to that worldview. This hypothesis is usually called the "Mortality Salience Hypothesis" (MSH), and is the most extensively tested of the TMT hypotheses. A meta-analysis covering 164 articles on 277 experiments concluded that MSH "is robust and produces moderate to large effects" (Burke, Martens, and Faucher 187).

Mindfulness (*sati/smṛti*) literally means memory or consciousness of something.²⁷ Mindfulness of death is the conscious awareness that one will eventually and inevitably die. Hence, mindfulness of death is mortality salience (although perhaps an especially intense variety thereof), and achieving mindfulness of death is increasing mortality salience. Mindfulness of death is religiously motivating because—according to MSH—increasing mortality salience leads among others to religious worldview defense. Awareness of death raises religious (and other cultural) identification and strengthens belief in (or consent to) religious doctrine, but also increases negativity (and even hostility) towards other religions (and cultures). *De dicto* moral motivation is produced in the same way.²⁸ For example, if tolerance is a key value in the subject's worldview, then

²⁷See Lopez (92-99) on the concept of "mindfulness" and its recent evolution.

²⁸Religious motivation is always *de dicto*.

an increase in mortality salience leads to an increase in tolerance (rather than in hostility; Greenberg et al.).

Recall that Buddhaghosa recommended the meditation on death as one of two general meditation subjects (III.57-59). The other kind of exercise is aimed at gradually extending “loving-kindness” (*mettā*), from (at first) just the community of monks to all of mankind (III.58) and ultimately the whole universe (chapter IX). Although “*mettā*” is usually translated as “loving-kindness” or “benevolence,” these terms do not cover the same semantic field (as is the case with many other translations of Buddhist terms). *Mettā* is comparable to the love one feels for one’s friends—it is a genuine care for the other’s well-being similar to what Daniel Batson (“Empathy”; *Altruism*) calls “empathic concern”—but to be applied universally rather than to specific targets. In the combination of Buddhaghosa’s two general meditation subjects, it is this second meditation that provides the content to the *de dicto* moral motivation: mindfulness of death or mortality salience produces a motivation to be benevolent.²⁹ Furthermore, Batson (*Altruism*) has shown that empathic concern produces altruistic motivation, and consequently, if *mettā* is empathic concern, then a *de dicto* moral motivation to be benevolent is also a motivation to do the right thing *de re*.³⁰

While mindfulness of death may be inseparable from *saṃvega*, the two are not identical, and the TMT explanation of how mindfulness of death produces *de dicto* moral motivation does not explain how *saṃvegic shock* motivates. The state of (increased) mortality salience and the fear

²⁹This, however, is not how Buddhaghosa argues for the “usefulness” of the meditation on loving-kindness in III.58. Rather, his argument, which reminds of Aristotelian virtue ethics, is that with loving-kindness to all of mankind, the monk will be trusted and not disliked by everyone, and therefore, can go wherever he wants.

³⁰But this is rightness *de re*, not motivation *de re*.

of death figuring in TMT are not *saṃvega*; they are not states of shock or panic such as those described by James Baillie and others. And the MSH explanation of how mindfulness is or can be morally motivating does not depend on *saṃvega* either. By implication, if it is the shocking aspect of *saṃvega* that is morally motivating, then it does not motivate (just) through mindfulness. As already mentioned above, however, the fact that *saṃvega* is the desired outcome of a meditation exercise suggests another way in which *saṃvega* could be morally motivating: if the meditation on death is both tranquility meditation and insight meditation in a single exercise, and *saṃvega* plays no *necessary* role in the first (that is, in the kind of meditation associated with mindfulness), then it must be associated with insight; then *saṃvega* must be morally motivating through an improved understanding of impermanence, no-self, and/or suffering.

If it is assumed that *saṃvega* results in insight(s), then what kind of insights could those be, and how do they lead to moral motivation in turn? Unfortunately, while TMT provides an empirically tested framework explaining how and when mindfulness of death can be morally motivating, there is no similar empirically supported theoretical framework for explaining *saṃvega* as insight, and for that reason, the answer to these questions in the remainder of this section is somewhat speculative.

There are at least three obvious (to me) ways in which *saṃvega* could be morally motivating through insight. Firstly, for the lucky among us who never experience(d) life-threatening suffering or violence, *saṃvega* is the closest to that experience one can get. *Samvega* is facing death from a safe distance, but it is facing death nevertheless, and it is the only way to experience what victims of extreme suffering or violence go through, aside from undergoing such suffering or violence oneself. In this way, *saṃvega* produces insight into suffering.

Secondly, the confrontation with the threat of death in *saṃvega* counters another, misleading encounter with that threat that many of us are probably more familiar with: the adrenaline rush resulting from doing something stupid enough to almost get oneself killed. Someone who never experienced *saṃvega* but has experienced that adrenaline rush can only (sub-consciously) associate life-threatening situations with that adrenaline rush. *Saṃvega* hurts; the adrenaline rush, on the other hand, is experienced by many as pleasurable (and some people actively seek it). The consequence should be obvious: having experienced *saṃvega* changes one's subjective perspective on life-threatening situations: its experience drives out any positive or pleasurable association. And because the threat of death is one of the most elemental kinds of suffering, this change in perspective produces further insight in suffering.

Thirdly, whatever the mind consciously or unconsciously infers from the new or changed insight in its own mortality depends to at least some extent on what else it believes. If one does not (firmly) believe in an afterlife or some other kind of survival after death, then in or after *saṃvega*, it doesn't just become painfully obvious that the existence of the subjective self is limited, but also what this implies: that there are no second chances or do-overs. It is the insight that the brevity of life implies that there is a limit to what one can do, and that the irreversibility of time implies that nothing of what one ends up doing can be undone. Of course, one may already be disposed to affirm all of this (or have system 2 beliefs with these contents), but during and after *saṃvega* it is system 1 that may come to realize all this (and while system 1 can ignore system 2, system 1 cannot be ignored). This insight (or these insights) then, results (result) in a sense of urgency, which suggests that the common translation of "*saṃvega*" as "sense of urgency"—while not a literal translation—does capture a key aspect of this state and/or its effects.

If these sketchy remarks are right, then *saṃvega* produces and deepens insights in the nature of suffering and the brevity and irretrievability of an individual's life. Presumably, in most individuals these insights together strengthen the motivation to avoid and/or decrease suffering, but that is not a moral motivation (*de re*, by the standard set above) if it is just a motivation to avoid and/or decrease one's own suffering. However, if these insights make suffering intolerable, then they will make *all* suffering (rather than just my own) intolerable if I do not strictly separate myself from others, and that would result in moral motivation (*de re*, by that same standard).

This argument may seem similar to Śāntideva's controversial argument for altruism in *Bodhicaryāvatāra* 8:101-103 (see Harris for an analysis of interpretations of Śāntideva's argument), but there are significant differences. Śāntideva's argument seems to be something like the following: Suffering must be prevented (103). No suffering belongs to anyone (101-102). Therefore, all suffering must be prevented (103).³¹ In contrast, my argument is that because all suffering *does* belong to someone, a reduction of the (subjective) distance between self and others makes all suffering intolerable. Furthermore, this "reduction of distance" does not require a complete dissolution of the (sense of) self, but

³¹While this superficially looks like a syllogism, it turns out not to be one upon closer inspection. The two premises can be formalized as $\exists x[Sx \wedge Px]$ (or as $Sc \wedge Pc$) and $\forall x[Sx \rightarrow Ax]$ (in which A means "does not belong to anyone"), respectively, and the conclusion as $\forall x[Sx \rightarrow Px]$. Obviously, that conclusion does not follow from the premises, but it does if a premise is added stating that "for any x , if x does not belong to anyone and there is something that does not belong to anyone and that must be prevented, then x must be prevented"; formally: $\forall x[(Ax \wedge \exists y[Ay \wedge Py]) \rightarrow Px]$. This additional premise is implausible, however, and therefore, if this interpretation is correct, then Śāntideva's argument fails. Considering his intended audience—monks who were already convinced of the truth of the premises *and* the conclusion—the alleged "argument" may not have been intended as a formal argument, however (see also Harris).

merely a well-developed sense of empathic concern (i.e., genuine care for others' well-being), and developing such a sense of empathic concern is the aim of loving-kindness meditation. By implication, the combination of the meditation on death with loving-kindness meditation is as essential for (*de re*) moral motivation through insight as it is for (*de dicto*) motivation through mindfulness.

While *saṃvega* is not mindfulness, this non-identity should not be mistaken for non-relation. *Saṃvega* as death-related shock is likely to increase mortality salience and mindfulness of death, and conversely, mindfulness of death may lead to *saṃvega* (as Buddhaghosa suggests in III.58). In all likelihood, *saṃvega* and mindfulness of death come together, which makes the separation of their effects on moral motivation a rather scholastic exercise, especially when it is taken into consideration that in either case something like loving-kindness meditation is a necessary catalyst.

Concluding Remarks

Saṃvega is a religiously and morally motivating state of shock that according to Buddhaghosa can and should be evoked by meditating on death. What kind of state *saṃvega* is, and how it produces moral motivation remains unclear in his (and other) writings, however. Addressing this omission, in this article I proposed a theory of what *saṃvega* is, and a hypothesis of how it works.

According to dual process theories there are two kinds of thought-like mental processes, organized (loosely) in two systems: the experiential, automatic system 1, and the rational, controlled system 2. In normal circumstances, system 1—more or less—believes it is immortal because its beliefs are fed by experience and any experience presupposes

and reconfirms the subject's existence. *Saṃvega* is a state of shock caused by the occurrence of a system 1 belief in its own mortality when system 2 is already disposed to affirm this.

This state of shock is morally motivating in (at least) two ways, but in either way, Buddhaghosa's second general meditation exercise—that on loving-kindness (*mettā*)—is an essential ingredient. Meditation on death is aimed at both *saṃvega* and mindfulness of death, but those two are related in such a way that they most likely come together (or in close succession). Mindfulness of death increases mortality salience, which together with an increase in loving-kindness as a core value (through loving-kindness meditation) produces a motivation to reduce suffering. *Saṃvega* itself produces insight in suffering, which (again) with an increase in loving-kindness (etc.) produces a motivation to reduce suffering. If it is assumed that suffering is morally bad—and that seems to be a plausible assumption—then this motivation to reduce suffering is a moral motivation. While this may explain how *saṃvega* is morally motivating, it must be emphasized that parts of this explanation are somewhat speculative and in need of further empirical research. Hence, the choice of the term “hypothesis” two paragraphs back.

As mentioned, Buddhaghosa recommends the meditation on death as one of two general meditation subjects that are beneficial to anyone. The other is loving-kindness meditation. As argued above, the combination is essential. What has not been addressed yet, however, is the recommendation itself. Should everyone strive to experience *saṃvega*? Although this question is outside the scope of this article, it is an important question, and in closing, I want to comment briefly on what is recommended and the reasons for recommending it.

There are very many different forms and techniques of meditation in Buddhism, and many of those seem to have little in common. This suggests that “meditation” is a functional rather than a substantive cat-

egory—that is, it is not defined by some substantive properties that all activities called “meditation” have in common, but by those activities’ function, by what meditation is intended to establish. If this is right, then meditation is any mental exercise that results in a certain, specified goal. Hence, meditation on death is any mental exercise that results in *saṃvega* and mindfulness of death (i.e., awareness of one’s mortality); and loving-kindness meditation is any mental exercise that results in a significant increase in *mettā*, loving-kindness, but perhaps here better translated as “empathic concern.” Taking this into account, Buddhaghosa’s recommendation can be read as a recommendation to experience *saṃvega*, increase mortality awareness (or mortality salience), and increase (untargeted, universal) empathic concern.³²

Buddhaghosa recommends the two general meditation subjects because they are “very beneficial” (*bahūpakāra*; III.59) to the meditating individual, and they probably are if the underlying karmic worldview is accepted, but without assuming *kamma*, rebirth, and liberation (*nibbāna*) this eudaimonistic argument is questionable. Arguably, in order to flourish a human being needs neither *saṃvegic* shock³³ nor loving-kindness beyond a minimum required to be socially accepted. However, this judgment changes radically if one takes a social (rather than individual, eudaimonistic) point of view. If suffering is bad, and people with *saṃvegic* experience(s) and increased loving-kindness/empathic concern are motivated to reduce suffering, then indeed mental exercises to evoke *saṃvega* and to increase mortality awareness and empathic concern are socially desirable.

³²Perhaps increasing both empathic concern and familiarity with *saṃvega* results in the (more frequent) experience of empathic *saṃvega* (see above) as well, and probably, empathic *saṃvega* further increases both insight in suffering and empathic concern.

³³On the contrary, for human flourishing, the death-denying (and non-dissociative) *D-state* would be preferable even to the death-affirming *A-state*.

References

Albahari, Miri. "Alief or Belief? A Contextual Approach to Belief Ascription." *Philosophical Studies* 167 (2013), 701-720.

Āṅguttaranikāya [Gradual Sayings / Numerical Discourses].

Baillie, James. "The Expectation of Nothingness." *Philosophical Studies* 166.S (2013), S185-S203.

Batson, C. Daniel. "The Things Called Empathy: Eight Related but Distinct Phenomena." In *The Social Neuroscience of Empathy*, edited by J. Decety and W. Ickes, 3-15. Cambridge, MA: MIT Press, 2009.

_____. *Altruism in Humans*. New York: Oxford University Press, 2011.

Becker, Ernest. *The Denial of Death*. New York: Simon & Schuster, 1973.

Brekke, Torkel. "The Role of Fear in Indian Religious Thought with Special Reference to Buddhism." *Journal of Indian Philosophy* 27 (1999), 439-467.

Brons, Lajos. "The Incoherence of Denying My Death." *Journal of Philosophy of Life* 4.2 (2014), 68-89.

Buddhaghosa. *Visuddhimagga*. Translation: Bhikkhu Nyanamoli. Onalaska: BPS-Pariyatti, 1999.

Burke, Brian L., Andy Martens, and Erik H. Faucher. "Two Decades of Terror Management Theory: a Meta-Analysis of Mortality Salience Research." *Personality and Social Psychology Review* 14.2 (2010), 155-195.

Carruthers, Peter. *The Opacity of Mind*. Oxford: Oxford University Press, 2011.

_____. "On Knowing Your Own Beliefs." In *New Essays on Belief*, edited by N. Nottelman, 145-165. London: Palgrave MacMillan, 2013.

_____. *The Centered Mind: What the Science of Working Memory Shows Us About the Nature of Human Thought*. Oxford: Oxford University Press, 2015.

Chagnon, Napoleon A. *Yanomamö: The Fierce People*. New York: Holt, Rinehart, and Winston, 1968.

Coomaraswamy, Ananda K. "Saṁvega, 'aesthetic shock.'" *Harvard Journal of Asiatic Studies* 7.3 (1943), 174-179.

Davidson, Donald. "The Myth of the Subjective." In *Subjective, Intersubjective, Objective*, 39-52. Oxford: Oxford University Press, 2001.

Dennett, Daniel C. "The Self as a Center of Narrative Gravity." In *Self and Consciousness: Multiple Perspectives*, edited by F. S. Kessel, P. M. Cole and D. L. Johnson, 103-115. Lawrence Erlbaum: Mahwah, NJ, 1992.

Emavardhana, Tipawadee and Christopher D. Tori. "Changes in Self-Concept, Ego Defense Mechanisms, and Religiosity Following Seven-Day Vipassana Meditation Retreats." *Journal for the Scientific Study of Religion* 36.2 (1997), 194-206.

Epictetus. *Ἐπικτήτου διατριβαί* [The Discourses].

_____. *Ἐγχειρίδιον Ἐπικτήτου* [The Manual].

Epstein, Seymour. "The Self-Concept Revisited, or a Theory of a Theory." *American Psychologist* 28 (1973): 404-416.

_____. "Cognitive-Experiential Self-Theory: an Integrative Theory of Personality." In *The Relational Self: Theoretical Convergences in Psychoanalysis and Social Psychology*, edited by R. C. Curtis, 111-137. New York: Guilford, 1991.

_____. "Integration of the Cognitive and the Psychodynamic Unconscious." *American Psychologist* 49 (1994), 709-724.

Epstein, Seymour and Rosemary Pacini. "Some Basic Issues Regarding Dual-Process Theories from the Perspective of Cognitive-Experiential Self-Theory." In *Dual-Process Theories in Social Psychology*, edited by S. Chaiken & Y. Trope, 462-482. New York: Guilford, 1999.

Farias, Miguel and Catherine Wikholm. *The Buddha Pill: Can Meditation Change You?* London: Watkins, 2015.

Fisher, Mark and Jeremy Gilbert. "Capitalist Realism and Neoliberal Hegemony: a Dialogue." *New Formations* 80-81 (2013), 89-101.

Gawronski, Bertram and Laura A. Creighton. "Dual Process Theories." In *The Oxford Handbook of Social Cognition*, edited by D. E. Carlston, 282-312. New York: Oxford University Press, 2013.

Gendler, Tamar Szabó. "Alief and Belief." *The Journal of Philosophy* 105, no. 10 (2008), 634-663.

_____. "Alief in Action (and Reaction)." *Mind & Language* 23, no. 5 (2008), 552-585.

_____. "On the Epistemic Cost of Implicit Bias." *Philosophical Studies* 156 (2011), 33-63.

Greenberg, Jeff and Jamie Arndt. "Terror Management Theory." In *Handbook of Theories of Social Psychology*, Volume one, edited by P. A. M. van Lange, A. W. Kruglanski, and E. T. Higgins, 398-415. London: Sage, 2012.

Greenberg, Jeff, Linda Simon, Tom Pyszczynski, Sheldon Solomon, and Dan Chatel. "Terror Management and Tolerance: Does Mortality Salience Always Intensify Negative Reactions to Others Who Threaten One's Worldview?" *Journal of Personality and Social Psychology* 63, no. 2 (1992), 212-220.

Harris, Stephen. "Does Anātman Rationally Entail Altruism? On *Bodhicaryāvatāra* 8:101-103." *Journal of Buddhist Ethics* 18 (2011), 93-123.

Heim, Maria. "The Aesthetics of Excess." *Journal of the American Academy of Religion* 71, no. 3 (2003), 531-554.

Hume, David. *A Treatise on Human Nature*.

Johnston, Mark. *Surviving Death*. Princeton: Princeton University Press, 2010.

Kahneman, Daniel. *Thinking, Fast and Slow*. London: Penguin, 2011.

_____. "A Perspective on Judgment and Choice: Mapping Bounded Rationality." *American Psychologist* 58 (2003), 697-720.

Kesebir, Pelin and Tom Pyszczynski. "The Role of Death in Life: Existential Aspects of Human Motivation." In *The Oxford Handbook of Human Motivation*, edited by R. M. Ryan, 43-64. New York: Oxford University Press, 2012.

Kierkegaard, S. (as Anti-Climacus). *Sygdommen til Døden* [The Sickness unto Death].

Kierkegaard, S. (as Vigilius Haufniensis). *Begrebet Angest* [The Concept of Anxiety].

Kriegel, Uriah. "Moral Motivation, Moral Phenomenology, and the Alief/Belief Distinction." *Australasian Journal of Philosophy* 90, no. 3 (2011), 469–486.

Locke, John. *An Essay Concerning Human Understanding*.

Lopez, Donald S. *The Scientific Buddha: His Short and Happy Life*. New Haven: Yale University Press, 2012.

Lindholm, Charles. "The Rise of Expressive Authenticity." *Anthropological Quarterly* 86, no. 2 (2013), 361-396.

Mandelbaum, Eric. "Against Alief." *Philosophical Studies* 165 (2012), 197-211.

Metzinger, Thomas. *Being No-one: the Self-Model Theory of Subjectivity*. Cambridge, MA: MIT Press, 2003.

Nagel, Jennifer. "Gendler on Alief." *Analysis Reviews* 72, no. 4 (2012): 774-788.

Nagel, Thomas. *The View from Nowhere*. Oxford: Oxford University Press, 1986.

Parfit, Derek. *Reasons and Persons*. Oxford: Clarendon, 1984.

_____. *On What Matters*, Volume 2. New York: Oxford University Press, 2011.

Śāntideva. *Bodhicaryāvatāra*. Translation: K. Crosby and A. Skilton. Oxford: Oxford University Press, 2011.

Schwitzgebel, Eric. "Acting Contrary to Our Professed Beliefs or the Gulf between Occurrent Judgment and Dispositional Belief." *Pacific Philosophical Quarterly* 91 (2010), 531-553.

Smith, Michael. *The Moral Problem*. Oxford: Basil Blackwell, 1994.

Solomon, Sheldon, Jeff Greenberg, and Tom Pyszczynski. *The Worm at the Core: on the Role of Death in Life*. New York: Random House, 2015.

Stanovich, Keith E. and Richard F. West. "Individual Differences in Reasoning: Implications for the Rationality Debate?" *Behavioral and Brain Sciences* 23 (2000), 645-726.

Williams, Paul. "General Introduction: Śāntideva and His World." In Śāntideva's *Bodhicaryāvatāra*, vii-xxvi.