

1 **On defining moral enhancement: a clarificatory taxonomy**

2

3 **ABSTRACT**

4 Recently there has been some discussion concerning a particular type of enhancement, namely ‘moral
5 enhancement’. However, there is no consensus on what precisely constitutes moral enhancement, and
6 as a result the concept is used and defined in a wide variety of ways. In this article, we develop a
7 clarificatory taxonomy of these definitions and we identify the criteria that are used to delineate the
8 concept. We think that the current definitions can be distinguished from each other by the criteria used
9 for determining whether an intervention is indeed moral enhancement. For example, some definitions
10 are broad and include moral enhancement by any means, while other definitions focus only on moral
11 enhancement by means of specific types of intervention (e.g. biomedical or genetic interventions).
12 Moreover, for some definitions it suffices for an intervention to be *aimed* or *intended* to morally
13 enhance a person, while other definitions only refer to ‘moral enhancement’ in relation to interventions
14 that are actually *effective*. For all these differences in definitions we discuss some of their (more
15 normative) implications. This shows that definitions are significantly less descriptive and more
16 normative than they are regularly portrayed to be. We therefore hope that the taxonomy developed in
17 this paper and the comments on the implications for the normative debate of the variety of definitions
18 will provide conceptual clarity in a complex and highly interesting debate.

19

20 **INTRODUCTION**

21 The debate about enhancing human traits has been raging for some decades, and more recently there
22 has been some discussion concerning one particular category of enhancement, namely ‘moral
23 enhancement’. What is at issue in moral enhancement is not the improvement of physical and/or
24 cognitive capacities, but improvement in the way in which we act or reflect morally. Concerns have
25 been voiced that tinkering with our beliefs of what is right and wrong, or our motivation to act rightly
26 or wrongly, might be to open a Pandora’s box that could lead to unforeseen and potentially disastrous
27 consequences [1-3]. Others praise moral enhancement as an essential step in guaranteeing even the

28 very survival of the human race as the potential for doing great harm (e.g. with biological or other
29 weapons of mass destruction) continues to increase [4-7].

30

31 Although the debate is of recent date, ‘moral enhancement’ or ‘moral bioenhancement’ has already
32 become an established concept. Nevertheless, it is far from clear what precisely constitutes moral
33 enhancement. Different authors use different definitions or meanings of term. Existing definitions can
34 differ to such a degree that a particular intervention would constitute or result in moral enhancement
35 according to one definition, but not according to another. However, even though many different
36 definitions of the term exist, this is not always acknowledged in the debate. In this respect, John Shook
37 has stated:

38 Too many discussions are proceeding as if both the meaning and the possibility of moral
39 enhancement were already widely understood and agreed upon. (...) Asking such questions,
40 and offering answers, depend on assigning some sense or another to “moral enhancement.”
41 However, clear and precise definitions of “moral enhancement” are not to be found; what has
42 been called “moral” enhancement ranges from feeling empathic concern to increasing personal
43 responsibility all the way to heightening respect for global fairness [8, p.3]

44

45 And also:

46 anyone using the term ‘moral enhancement’ as if everyone knows what is meant must either
47 be simplifying matters to the point of negligence, or trying to speak only to those already in
48 local moral consensus. [8, p.4]

49

50 We therefore believe that it is important to chart the different types of definition and so develop a
51 taxonomy of existing definitions or uses of the concept of moral enhancement. We focus on a number
52 of criteria for what counts as a moral enhancement that are included in some definitions but are not
53 included or are included, but in a different form, in others. Whenever possible, we use definitions
54 given by authors in their articles and stay as close as possible to the author’s wording. However, not

55 every author gives an explicit definition, and in some cases we must therefore focus on how these
56 authors use the concept of moral enhancement and in what context.

57

58 Of course, creating a descriptive taxonomy can only be a first step in the debate on moral
59 enhancement, as is clear from the quotes from Shook given above. Another issue that is rarely
60 acknowledged in the debate on moral enhancement, is that behind the seeming neutrality of defining
61 the concept, there often lie philosophical battles as to what constitutes morality and what it means to
62 act morally. In this paper we will therefore also discuss the (normative) implications of using certain
63 types of definition and of including or excluding certain elements from the definition. Our aim is to go
64 beyond the simplification and local moral consensus described by Shook above, to chart the
65 complexity of the concept and its implications for the normative debate on the permissibility or
66 desirability of specific interventions aimed at moral enhancement.

67

68 We believe a descriptive taxonomy, combined with a discussion of some of the main implications of
69 using certain types of definition, are important tools for anyone wishing to conduct a normative
70 analysis of the ethical desirability of moral enhancement. We will distinguish different definitions of
71 moral enhancement based on the criteria they use for determining whether a certain intervention is
72 indeed a moral enhancement. We each time focus on a single criterion whereby we discuss (1) how
73 definitions can be distinguished based on their inclusion or exclusion of this criterion, and (2) what the
74 possible (normative) implications are of including or excluding it. For example, some definitions are
75 broad and include moral enhancement by any means, while other definitions focus only on moral
76 enhancement by means of specific types of intervention (e.g. biomedical or genetic interventions).
77 Another element is that some definitions consider as moral enhancement those interventions that
78 change a person's moral behaviour while for other definitions an intervention can only be seen as a
79 moral enhancement when it targets a person's moral capacities. However, in this paper we will take no
80 position on the definitions or criteria we prefer or disagree with, and we will not formulate definitions
81 of our own.

82

83 **FOCUS ON THE INTERVENTION OR ON THE INDIVIDUAL**

84 When comparing different definitions or uses of the concept of ‘moral enhancement’, one has to make
85 sure one is not comparing apples and oranges. Many publications touch on the question: what is moral
86 enhancement? However, this question seems to be understood in two different ways. Some authors
87 formulate the question as: when can a certain *intervention* be considered a moral enhancement?. In
88 doing so, they focus on the criteria that need to be met by the interventional process. David DeGrazia,
89 for example, seems to understand moral enhancements as:

90 interventions that are intended to improve our moral capacities such as our capacities for
91 sympathy and fairness’. [9, p.1]

92

93 Other commentators focus less on the intervention, and seem to understand the question of what is
94 moral enhancement as: when can an individual be seen to have been morally enhanced or what does it
95 mean for an individual to be morally enhanced?. James Hughes, for example, argues for a conception
96 of moral enhancement that focuses heavily on the individual rather than the intervention:

97 Moral enhancement is not just the jacking up of virtue with neurochemicals. It is more broadly
98 taking conscious control of our lives to build the kind of character we want to have. [10, p.4]

99

100 It will be clear that focusing on the intervention rather than on the individual allows consideration of
101 the efficacy of the intervention in particular cases to be deferred, and even that the nature of the moral
102 improvement becomes less central.

103

104 **BROAD VERSUS MORE SPECIFIC INTERVENTIONAL MEANS**

105 Another difference is that some authors [8; 11-13] use moral enhancement as a broad concept that
106 covers any practice that causes or is intended to cause a change in the functioning of moral capacities,
107 with types of interventions ranging from non-invasive (e.g. moral education) to highly invasive (e.g.
108 deep brain stimulation or brain surgery). Others [4, 14] are more specific and focus on moral
109 *bioenhancement* which they seem to understand as ‘moral enhancement by biomedical and genetic
110 means’ [4, p.162]. Hence, these authors make explicit that they only focus on certain (invasive) means

111 for enhancement, thereby leaving out such interventions as moral education or talk therapy. A
112 particular example of a specific focus can be found in an article by Mark Walker which discusses
113 ‘enhancing genetic virtue’, a specific way of morally enhancing individuals:

114 Engineering genetic virtue (...) would mean promoting genes that influence the acquisition of
115 the virtues. [15, p.26]

116

117 It is clear that certain (more invasive) interventions are more controversial than other (less invasive)
118 ones such as talk therapy and moral education. When one limits moral enhancement to more
119 controversial cases, it is clear that when it comes to arguing for the ethical desirability of moral
120 enhancement, a stronger justification will be needed.

121

122 Moreover, using a broad conception of moral enhancement seems to imply that interventions such as
123 moral education and neurological interventions do not differ in principle (as they are both captured
124 under the heading of ‘moral enhancement’), but solely in terms of their practical implications/effects,
125 for example, invasiveness or effectiveness. Using the concept of moral enhancement for *all*
126 interventions or only for those interventions that are (most) invasive, can serve to bias the normative
127 debate towards or away from a conclusion of permissibility or desirability.

128

129 **ENHANCING INDIVIDUALS VERSUS ENHANCING HUMANITY**

130 Definitions can differ in what they consider to be the *target* of moral enhancement. Most definitions
131 stipulate that ‘moral enhancement’ refers to interventions that are used on individual persons. For
132 example, Tom Douglas [16,17] defines moral enhancement as:

133 interventions that will expectably leave an individual with more moral (*viz.* morally better)
134 motives or behaviour than she would otherwise have had. [17, p.3]

135

136 Some commentators, however, use a definition that also seems to include interventions that operate on
137 groups of persons, or even on society and/or humanity in general. This is the case, for example, for
138 Ingmar Persson and Julian Savulescu who discuss enhancing ‘the moral character of humanity’ [4].

139

140 To illustrate the difference between morally enhancing individuals and morally enhancing society,
141 consider the following imaginary case:

142 A country's government passes a law that requires that for every case of In Vitro Fertilisation,
143 only embryos that do not possess a certain set of genes associated with a higher risk of
144 developing anti-social personality disorder are suitable candidates for implantation.

145

146 In this case one could claim – following well-known arguments made by Derek Parfit in his *Reasons*
147 *and Persons* [18] – that no individual would end up with better motives or behaviour than she would
148 otherwise have had (as she either would not have been born or would have been chosen anyway), yet
149 the society she lives in might end up with more people behaving morally. The theoretical possibility of
150 such moral enhancement through genetic selection has been discussed by Walker [15] and Halley
151 Faust [19]. According to definitions such as the one mentioned above by Douglas, such a program of
152 pre-implantation genetic testing and selection would simply not be considered to be a moral
153 enhancement as, following arguments voiced by Parfit, no individual could claim to be morally
154 improved and so would fall outside the normative debate in this field. As with the previous section,
155 excluding the 'difficult' and including the 'easy' both have implications for the debate.

156

157 These definitions, covering the individual alone or covering society as well, thus also have relevant
158 implications. For justifying moral enhancement applied solely to individuals, justifications related to
159 autonomy and personal benefit might suffice. However, in justifying moral enhancement on a societal
160 level one would need to turn to other justifications such as issues of justice or of achieving a common
161 good. Moreover, concerns about altering human nature would then come into view as well

162

163 **MORAL TREATMENT VERSUS MORAL ENHANCEMENT**

164 Many definitions understand moral enhancement as any form of moral improvement, regardless of
165 whether it involves an improvement towards average functioning of moral capacities or one towards
166 above average functioning. However, there are exceptions. Nicholas Agar, for example, has explicitly

167 criticised the definition used by DeGrazia (quoted above) for not differentiating between *moral*
168 *enhancement as improvement* (something Agar considers to be synonymous with moral therapy), and
169 *moral enhancement beyond human norms*.¹ For Agar, an example of moral *therapy* would be
170 ‘endowing the likes of John Wayne Gacy or Ted Bundy with a normal sensitivity to suffering’ [21,
171 p.73]. Moral *enhancement*, on the other hand, ‘has the purpose of boosting responsiveness to ethical or
172 moral reasons to levels beyond that considered normal for human beings’ [21, p.73]. For Agar, moral
173 enhancement refers exclusively to interventions that raise people to a higher level of functioning of
174 moral capacities that is infrequently observed among humans or even to levels beyond those ever
175 observed. Agar considers moral therapy and moral enhancement to be sufficiently different not to be
176 lumped together under one term.

177

178 Likewise, Dorothee Horstkötter et al. argue that bringing people to average levels of moral reflection
179 or behaviour is medical treatment and not enhancement. They claim that:

180 if there is a health problem, medical treatment is the reasonable reaction, while enhancement,
181 either moral or otherwise, does not arise [22, p.27]

182

183 In order to be able to distinguish moral treatment from moral enhancement in this way, one needs to
184 determine what constitutes an average or normal level of functioning of moral capacities or behaviour.
185 Every intervention that brings a person to this average level would then be treatment, while
186 enhancement would consist of improving beyond this average level. To put this somewhat more
187 analytically, those who maintain a difference between treating and enhancing, have to be able, at least
188 theoretically, to distinguish:

189 1) An intervention (X), used on a person or group with a below average functioning of moral
190 capacities and/or behaviour (M^{BA}) that is intended to raise or succeeds at raising this person’s
191 or group’s functioning of moral capacities and/or behaviour to an average level (M^A)

¹ The debate on a possible distinction between ‘treatment’ and ‘enhancement’ is, of course, by no means exclusive to the moral enhancement debate. Many commentators have already discussed the relevance or irrelevance of such a distinction for the general debate on enhancement [20].

192

193

2) An intervention (Y), used on a person or group with an average functioning of moral

194

capacities and/or behaviour (M^A) that is intended to raise or succeeds at raising this person's

195

or group's functioning of moral capacities and/or behaviour to an above average level (M^{AA})

196

197

An intervention X could then be considered 'treatment', while interventions of type Y could then be

198

labelled as 'enhancement'. Naturally, there is also a third possibility which is:

199

3) An intervention (Z), used on a person or group with a below average functioning of moral

200

capacities and/or behaviour (M^{BA}) that is intended to raise or succeeds at raising this person's

201

or group's functioning of moral capacities and/or behaviour to an above average level (M^{AA})

202

203

Classifying interventions of type Z is more difficult as this intervention has both a treatment and an

204

enhancement aspect to it.

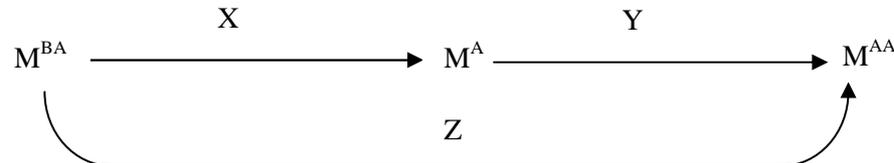
205

206

Put in a schematic way this becomes:

207

208



209

210

Scheme 1: Three types of interventions

211

212

However, spelling out the cut-off point between moral therapy and moral enhancement may often be

213

next to impossible. M^{BA} , M^A , and M^{AA} do not represent single and generally agreed upon levels of

214

moral capacities. There is a continuum that ranges from below average functioning of moral capacities

215

through to above average functioning. On this continuum, different cut-off points could be made as to

216

what falls within average functioning (and hence what falls outside), depending on how many standard

217

deviations from average one allows for an individual to still be considered as having normally

218

functioning moral capacities. As there is no objective way of determining what falls within the range

219

of 'normal' moral behaviour or functioning of moral capacities, every choice of cut-off point is,

220 necessarily, a normative one. Even the question as to what constitutes ‘moral capacities’ has no
221 straightforward answer.² Our proposed scheme only purports to create conceptual clarity, hence we
222 take no position as to where we believe the cut-off point should lie. We limit ourselves to pointing out
223 that different cut-off points *can* be made and that our scheme *can* be applied regardless of which cut-
224 off point one adopts.

225

226 Moreover, distinguishing moral treatment from moral enhancement raises other issues as it brings
227 morality within the medical domain. What seems to be implied by distinguishing moral enhancement
228 from moral treatment is not only that some people who lack in moral behaviour or capacities are
229 suffering from a disease or health problem, but also that they can sometimes be treated or cured. This
230 raises the question of whether, and if so under which conditions, certain forms of immorality should be
231 medicalized.

232

233 When one distinguishes moral treatment from moral enhancement, it also becomes impossible to
234 determine *whether a particular intervention, in itself, is a moral enhancement or not*. It is possible that
235 exactly the same intervention could be used in one case to raise a person to an average functioning of
236 moral capacities (intervention X), and in another case to raise a person to an above average
237 functioning of moral capacity (intervention Y). A single intervention could thus be a moral therapy or
238 a moral enhancement depending on the situation in which it is used.

239

240 Making a distinction between moral treatment and moral enhancement requires taking a normative
241 stance on what constitutes average or normal moral behaviour or average or normal moral capacities.

242 Even if one does not wish to make this distinction, however, claims about what constitutes an

² One might understand ‘moral capacities’ as being those capacities we actually use when making moral decisions (which capacities these are, is researched in moral psychology and neuroscience). [For example, a recent review on the neurobiology of morality argues that it is most plausible to depict moral processes as requiring the engagement of both emotional and cognitive neural networks] [29].] However, ‘moral capacities’ might also be used to refer to those capacities one believes one *should* use when making moral decisions, or capacities that, when used more or better, would lead to better moral decisions. These capacities might, for example, constitute the capacity for sympathy and fairness [9] or cognitive capacities [1].

243 *improvement* of moral capacities or behaviour – which is just as much up for debate – are inevitable. If
244 one leaves out standards on what it means to improve moral behaviour or capacities, the concept
245 becomes indistinguishable from mere mental modification, as has been noted by Filippo Santoni de
246 Sio et al. [28].

247
248 Again, as noted in earlier sections, limiting or extending the definition of moral enhancement by
249 excluding the controversial or by including the uncontroversial clearly has implications for the
250 normative debate.

251

252 **INTENDED VERSUS EFFECTIVE INTERVENTIONS**

253 Another element of difference between existing definitions concerns the question whether moral
254 enhancement only refers to interventions that are *successful* at improving a subject's moral capacity, or
255 whether it also includes interventions that are merely *intended* or *expected* to improve moral capacity
256 (regardless of whether they actually do).

257

258 The abovementioned definitions by DeGrazia [9] and Douglas [17] include cases of *intended or*
259 *expected but failed* interventions. For them, what is important in determining whether an intervention
260 is to be considered a moral enhancement is the intention with which the intervention is used.

261

262 John Harris, however, has reacted to this and has stated:

263 I do not define enhancements in terms of the intention or the motivation of those who produce
264 them but rather in terms of their effect. [23, p.1]

265

266 Another example is Shook [8], for whom an intervention can be considered a moral enhancement if it
267 affects an individual's moral beliefs, moral motives and/or moral behaviour. Thus, moral enhancement
268 should do one or more of five things:

269 1) Enhance a person's sensitivity to moral features of situations – resulting in heightened
270 moral *appreciation*

- 271 2) Enhance a person's thoughtfulness about doing the right thing – resulting in stronger moral
272 *decisions*
- 273 3) Enhance a person's moral judgments that get the right moral answer – resulting in more
274 correct moral *judgments*
- 275 4) Enhance a person's motivated choice to do what moral judgment indicates – resulting in
276 improved moral *intentions*
- 277 5) Enhance a person's volitional power to act upon a moral intention – resulting in more *will*
278 *power* [8, p.6; our italics]

279

280 Hence, for Harris and Shook an intervention that is intended to morally enhance a person, but fails to
281 affect that person's moral reflection process or behaviour, is not a moral enhancement. This difference
282 between moral enhancement as an intervention intended to enhance versus an intervention successful
283 in achieving enhancement becomes relevant when it comes to moral justification. If, as Shook's
284 definition seems to imply, an intervention can only be labelled a moral enhancement if it has a positive
285 effect then one important reason to oppose moral enhancement is eliminated. Indeed, all problematic
286 cases where no enhancement is reached or a person is left *less* moral than before the intervention,
287 would quite simply not be considered cases of moral enhancement at all. This way of justifying moral
288 enhancement would avoid all debate concerning potential risks, side-effects, etc.

289

290 **CAPACITIES-ORIENTED VERSUS BEHAVIOUR-ORIENTED INTERVENTIONS**

291 Some commentators in the moral enhancement debate label a certain intervention a moral
292 enhancement depending on its (real or intended) effect on a person's *behaviour*. Others see moral
293 enhancements as interventions that target or are intended to target a person's *capacities of moral*
294 *reflection*. We shall refer to this difference as the difference between a behaviour-oriented and a
295 capacities-oriented intervention.

296

297 Keeping in mind that an intervention of moral enhancement can target a person's behaviour or their
298 capacities of moral reflection, any intervention can have one of four results. It can result in a person:

Reflecting the same	Reflecting the same	299
Acting the same	Acting differently	300
		301
Reflecting differently	Reflecting differently	302
Acting the same	Acting differently	303
		304

305 *Scheme 2: four possible results*

306

307 Currently existing definitions differ as to the question in which of the quadrants in the scheme given
 308 above one can find cases of moral enhancement. To clarify this difference, imagine the following
 309 extreme case:

310 Jack is a man with paedophilic urges who is currently incarcerated for having sexually
 311 molested a child. Despite a large amount of therapy, Jack fails to see what is wrong with him
 312 interacting with children in a sexual way. It is therefore decided to sedate Jack against his will
 313 and bring him to a surgery room. Neurosurgeons implant a chip (call this intervention X) that
 314 will stop Jack from molesting children.

315

316 For this case, we need not deal with the question whether this intervention is ethically justified; we
 317 will merely focus on the question whether we should consider this to be ‘moral enhancement’. This is
 318 a case similar to one suggested by DeGrazia [9] who considers this an extreme form of ‘moral
 319 bioenhancement’. In identifying which interventions can be labelled moral enhancements, some
 320 commentators rely heavily or solely on the intervention’s achieved (or intended) effect on a person’s
 321 behaviour [e.g. 17]. For them, intervention X, from the imaginary case above, changes Jack’s
 322 behaviour for the better and hence would be considered moral enhancement, regardless of whether the
 323 intervention also affects Jack’s moral capacities. Therefore, for those commentators defining moral
 324 enhancement in a behaviour-oriented way, interventions of moral enhancement can be found only in
 325 the right half of scheme 2 above (reflecting the same and acting differently & reflecting differently and
 326 acting differently).

327

328 It is also clear that the requirement for successful behaviour change would not suffice to label an
329 intervention a moral enhancement by certain other authors, but instead would be categorised by them
330 as a form of behaviour control (e.g. Harris) or moral therapy (e.g. Agar). Such an intervention could
331 either fall on the side of mere behaviour control, or on the side of moral therapy, depending on the
332 individual undergoing the intervention (e.g. an individual with or without a sexual disorder) and the
333 goal in question (e.g. as a means of rehabilitation or mere crime reduction by the criminal justice
334 system). Such differences in definition have implications for the acceptability of certain moral
335 enhancement interventions, and may render certain interventions more acceptable if labelled as a
336 moral enhancement rather than as a form of behaviour control. The latter, if misused, may result in
337 ethically dubious practices (cf. the chemical castration laws of certain US States) (see Focquaert [30]).
338 Hence, depending on one's definition, the term moral enhancement could be knowingly or
339 unknowingly misused to justify practices that would otherwise be deemed immoral.

340

341 Indeed, others understand moral enhancement in a different way, and argue that it does matter whether
342 or not an intervention affects the functioning of the subject's moral capacities. True, they might say, in
343 the imaginary case Jack will no longer molest children, but if the implant is put in place against Jack's
344 will and his subsequent actions do not stem from a moral judgment on Jack's behalf concerning what
345 is morally right and wrong, this is not a *moral* enhancement.³ In their view, making someone 'more
346 moral' involves more than merely altering a person's behaviour. Or, as Harris puts it:

347 I take moral enhancement to involve enhancing our ability to think ethically (...), not
348 manipulating the probability of some reacting in ways that *others* deem ethical. [23, p.3; italics
349 in original]

350

³ Of course, again, we should not mistake the debate on what to call these interventions with debate on the ethical validity of such interventions. Authors such as Harris can consistently claim: (1) that an intervention such as the one performed on Jack is not a *moral* enhancement, and (2) that there might be some cases where such an intervention is morally justified.

351 Fabrice Jotterand has also criticised the one-sidedness of focussing on behaviour. He argues that most
352 moral neuroenhancement is unlikely to morally enhance people in the true sense of the term and notes
353 that:

354 While the manipulation of moral emotions might change the behavior of an individual, it does
355 not provide any content, for example, norms or values to guide one's behavioral response. [24,
356 p.6]

357

358 In a similar vein, William Simkulet argues that:

359 (i) forcing agents to act rightly, (ii) preventing agents from acting wrongly, and (iii) making it
360 harder for moral agents to act wrongly fail to constitute genuine moral enhancement. [25,
361 p.17]

362

363 For these commentators, as well as for those who use similar lines of argumentation [e.g. 13, 26], the
364 criterion for whether an intervention constitutes a moral enhancement is not behaviour, but whether
365 the intervention affects or improves an individual's capacities for moral reflection. They situate moral
366 enhancements in the lower half of scheme 2 (reflecting differently and acting the same & reflecting
367 differently and acting differently).

368

369 Defining moral enhancement in a capacities-oriented way has some important consequences. If moral
370 enhancement indeed merely refers to interventions that improve a subject's capacities for moral
371 reflection, then it is possible to morally enhance an individual without this resulting into improved
372 moral behaviour. Moreover, it is not unlikely that, for example, when one improves an individual's
373 moral beliefs, this does not result in improved moral behaviour, as it is well known that knowing the
374 good does not automatically lead to doing the right thing. This also means that if one uses a
375 *capacities-oriented* definition of moral enhancement, actually measuring whether and to what degree
376 an individual is morally enhanced is difficult, since it would require an assessment of a person's
377 reflection processes.

378

379 As such, the debate on whether moral enhancement can be better understood in a capacities-oriented
380 or a behaviour-oriented way is often only the surface of more fundamental debates, for example on
381 how important freedom is for moral action. Can behaviour be called moral if it is not free? Indeed,
382 only accepting changes in behaviour as a criterion for determining what constitutes moral
383 enhancement leads one to claiming that certain interventions that cause an individual to display moral
384 behaviour in an automatic manner, are examples of moral enhancements. A relevant thought
385 experiment in this respect is that of the ‘God Machine’, formulated by Savulescu and Persson [30], a
386 machine that monitors everyone’s desires and intentions, and which intervenes every time a person
387 forms an intention to perform a great moral evil (e.g. murder or rape) by simply changing that person’s
388 intention and thus her behaviour. In this scenario, people are still able to choose to do the right thing
389 (i.e. not to murder or rape), but unable to chose or perform moral evil. Those authors defending a
390 capacities-oriented approach might object that a person who initially intends to murder or rape but has
391 his mind changed by the God Machine is not at all morally enhanced, as his subsequent decision to act
392 or not act on that intention would not be free or autonomous.⁴ For authors such as Harris, an
393 intervention can only be a moral enhancement if it leaves the freedom to fall, i.e. to do the wrong
394 thing. Of course, focussing solely on the (intended or achieved) effect of interventions on a person’s
395 capacities for moral reflection, may commit one to calling certain interventions moral enhancements
396 even though they may in no way change a person’s behaviour.

397

398 One way to address this would be to combine the capacities-oriented and the behaviour-oriented
399 approaches. For example, one might say that an intervention is a moral enhancement if it changes
400 behaviour *or* if it changes one’s capacities for moral reflection. This broadens the field of what is
401 moral enhancement (as it would only exclude interventions in the upper left corner of scheme 2 - i.e.
402 reflecting the same and acting the same), but, of course, it also combines the problems mentioned
403 earlier.

⁴ In this respect, Savulescu and Persson do mention at the end of their paper that ‘[s]uch interventions and such control are not plausibly moral enhancements of that person’ [30, p.417]. They mainly argue for the God Machine on the grounds that there would be many positive effects and as such their view perhaps does not differ that much from authors such as Harris.

404

405 Another possible response would be to say that an intervention is only a moral enhancement if it
406 changes behaviour *and* it changes one's capacities for moral reflection. This allows one to avoid the
407 problems mentioned earlier, but of course it also narrows the scope of moral enhancement. For,
408 according to this way of delineating the concept, interventions of moral enhancement can only be
409 found in the lower right corner of scheme 2 (i.e reflecting differently and acting differently). In some
410 of his statements, Harris seems to be defending this position, for example when he says:

411 It seems to me that moral enhancement, properly so called, must *not only* make the doing of
412 good or right actions more probable and the doing of bad ones less likely, but must *also*
413 include the understanding of what constitutes right and wrong action. [23, p.172; our italics]

414

415 **ACTIVE INVOLVEMENT VERSUS PASSIVE RECEIVING**

416 At the heart of the capacities-oriented versus behaviour-oriented debate mentioned above lies the
417 question of what is or should be *targeted* with moral enhancements - moral capacities, moral
418 behaviour, either one of them, or both. This is not the same as another distinction that is made in the
419 moral enhancement debate, namely that between enhancement by means of a process requiring *active*
420 *involvement* on behalf of the individual it is used on, and enhancement in which the subject of the
421 intervention is a *passive recipient*. This distinction does not concern the *target* of a moral
422 enhancement, but rather *the way in which the enhancement is achieved*. A moral enhancement by way
423 of active involvement would then be an enhancement requiring conscious mental processes in the
424 subject as a means to achieve its result. Moral education would be a classical example. In contrast,
425 moral enhancement involving passive receiving would either be enhancement in an immediate way
426 where no active involvement was possible (e.g. a pill with immediate effect) or enhancement by way
427 of a process that required no deliberate involvement of the recipient (e.g. classical conditioning).

428

429 This distinction is relevant since, for some authors, the goal of moral enhancement is for individuals to
430 become more virtuous, and often these authors follow Aristotle in claiming that becoming virtuous is
431 always a conscious and deliberate process, where the way to becoming virtuous is just as essential as

432 the result (being virtuous). Chris Zarpantine talks about ‘the thorny and arduous path of moral
433 progress’ [27, p.141], while Jotterand states:

434 Virtue is a behavioral habit under the supervision of reason that can be taught and learned. The
435 control and manipulation of moral emotions by technological means reduce the human mind to
436 neurochemical processes and threaten the very essence of moral agency, that is, autonomy.
437 [24, p.7]

438

439 This way of thinking about moral enhancement may lead some commentators to reject passive ways of
440 enhancement as examples of genuine moral enhancement.

441

442 Interestingly, the distinction between moral enhancement by means of an active process and passive
443 moral enhancement can sometimes be used to complement the capacities-oriented versus behaviour-
444 oriented debate. Among the interventions that target the functioning of a person’s moral *capacities*,
445 one could distinguish those interventions that do so using an *active process* (e.g. moral education)
446 from those that do so in a *passive way* (e.g. a pill to clear up one’s moral reflection). However,
447 interventions that solely target a person’s *behaviour*, will generally do so in a *passive way* (e.g.
448 implants to make a person exhibit a certain behaviour), as moral enhancement by means of *active*
449 *involvement* requires conscious mental processes and, unless the subject is simply trying to learn how
450 to pass himself off as a moral person, will thus always also affect the person’s capacities for moral
451 reflection.

452

453 Whether one thinks of moral enhancement as a process requiring active involvement or as (also)
454 covering more passive ways of changing capacities and/or behaviour, has important implications. It
455 matters with regard to the relation between the person (or group) doing the enhancing and the person
456 (or group) being enhanced and the voluntariness of the enhancement. In moral interventions requiring
457 active involvement, the person being enhanced is at least aware of the process, is given an important
458 role and, most likely, is free to stop the process as her cooperation is essential. For more passive
459 interventions, it is possible to enhance a person against their will or even without them knowing they

460 are being morally enhanced, which is clearly morally problematic. For these reasons, interventions of
461 moral enhancement that do not require active involvement of the person being enhanced will most
462 likely be more controversial.

463

464 **CONCLUDING REMARKS**

465 In this paper we have tried to show that the term ‘moral enhancement’ is used in ways that cover a
466 large variety of different practices. First we examined and made explicit the different ways in which
467 the concept is used and defined, and we identified the criteria that are used to delineate the concept.
468 Next we showed that, despite claims of being descriptive, it is often well nigh impossible to separate a
469 definition’s descriptive content from its normative content. While some differences, have relatively
470 limited implications, perhaps simply in skewing the normative debate, other differences have more
471 far-reaching implications and are revealing of underlying normative theories or positions.

472

473 Particular implications that result from the choices made when faced with the alternative definitions
474 identified in the preceding sections can be summarised as follows:

475 (1) Broad versus specific means of moral enhancement carry implications with regard to the
476 invasiveness and practical effects of the intervention in question. More invasive interventions clearly
477 need a stronger justification. Moreover, definitions including broad means of moral enhancement may
478 obscure the often-voiced concern that biomedical means of enhancement pose a greater threat to
479 concepts of authenticity and identity compared to non-biomedical means.

480 (2) Conceptualizing moral enhancement on an individual level versus a level which includes society
481 has implications for its justification. For example, potential instances of invasive societal moral
482 enhancement will need greater justification compared to individual level and non-invasive societal
483 enhancements. Both specific safeguards pertaining to the rights and needs of the individual and a
484 justification pertaining to the rights and needs of society are likely to be required in the first case.

485 Certain less invasive interventions, such as moral education, may on the other hand be decided
486 through a democratic mandate and may not necessarily require consent from each individual in
487 question.

488 (3) Whether one includes or excludes moral treatment from the scope of moral enhancement has
489 implications regarding the medicalization of immoral behaviour and the desirability or undesirability
490 of labelling (some forms of) immoral behaviour as a disease. Distinguishing moral treatment from
491 moral enhancement has another implication. Anyone who accepts a distinction between moral
492 treatment and moral enhancement has to, at least in theory, acknowledge a level of ‘average’
493 functioning or behaviour in order to distinguish treatment from enhancement. This brings with it
494 significant difficulties as determining what is average functioning or behaviour inherently involves a
495 normative decision.

496 (4) Restricting moral enhancement to covering only effective interventions has implications for its
497 moral justification. Classifying potentially non-effective interventions as moral enhancement needs
498 additional justification compared to effective interventions since the beneficial effect can be regarded
499 as an important justificatory reason for pursuing specific enhancement interventions.

500 (5) Focusing on capacities versus behaviour as the target of moral enhancement may have important
501 implications for one’s implicit or intuitive acceptance of the interventions in question. Certain
502 interventions that can be considered immoral under specific circumstances (e.g. physical or chemical
503 castration by the criminal justice system in the Czech Republic), may appear more acceptable if
504 framed under the label of ‘moral enhancement’. This could lead to certain ethically dubious practices
505 being more widely accepted and more easily institutionalized. Moreover, the capacities/behaviour
506 distinction also has important implications for the value of freedom in morality and whether or not
507 morality requires ‘the freedom to fall’.

508 (6) Passive enhancement interventions have the potential to circumvent an individual’s approval and
509 may therefore more easily be considered controversial, especially if one considers society-wide
510 passive enhancement (e.g. the addition of some kind of neurochemical to drinking water). Greater
511 caution may thus be required when implementing passive interventions. Moreover, some normative
512 ethical positions may not regard passive interventions as moral enhancements at all, and may
513 therefore, justifiably or unjustifiably, exclude consideration of all passive means when moral
514 enhancement is considered.

515

516 Moreover, the particular approach one takes to normative ethical questions may have important
517 implications. For some utilitarian thinkers, if an intervention achieves behavioral control, this may be
518 sufficient for it to be labeled a moral enhancement, whereas this is not likely to be the case for, for
519 example, a virtue ethicist. For a rights-based ethicist, the ‘freedom to fall’ will typically be considered
520 to be an essential part of morality. For a virtue ethicist, passive interventions do not amount to moral
521 enhancements, whereas, provided certain safeguards are met, such interventions may count as
522 enhancements for utilitarian ethicists and rights-based ethicists. These underlying normative views are
523 important and may have a huge impact on one’s preferred definition of moral enhancement. The
524 capacities-oriented versus behavior-oriented approaches show, for example, how Harris would label
525 some interventions that Douglas considers ‘moral enhancement’ to be examples of mere behavior
526 control because in those examples the subjects lack the freedom to fall. Similarly, virtue ethicists may
527 not consider certain interventions proposed by Douglas and/or Persson and Savulescu to be moral
528 enhancements since some of these do not rely on the active involvement of the subject. If definitions a
529 priori rule out certain forms of moral enhancement or a priori include interventions that would not be
530 labeled moral enhancements by others, then this needs to be mentioned and explicitly acknowledged
531 in order for the debate to be able to move forward in a constructive and open manner.

532

533 This paper should not be read as a plea for one single and universally agreed upon definition. There
534 are many different types of interventions for which the concept can be used. We believe there need be
535 no problem with leaving ‘moral enhancement’ simpliciter as an umbrella term that may be used for
536 many kinds of interventions, as long as one makes it sufficiently clear just how one is using the
537 concept or what one is having it refer to and is aware of how one’s underlying normative position may
538 influence one’s understanding of the concept. Our paper is also not a plea for a descriptive definition
539 for, as we hope to have made clear throughout the paper, expelling all normative elements from such a
540 definition is impossible. What we wish to emphasise is that most authors fail to identify the impact of
541 their normative positions on their proposed definition of moral enhancement, and may thus, knowingly
542 or unknowingly, portray their definition as neutral. This may lead one to believe that a given definition
543 frames all the different aspects of the debate and all the different interpretations of what should and

544 should not be labeled a moral enhancement. Explicitly acknowledging one’s normative stance and
545 how this might impact one’s views will likely ameliorate the current normative debate and provide for
546 a more constructive approach to the question of the ethical desirability of specific interventions.”

547

548 We hope that this paper will enable progress in the debate on moral enhancement, by providing a
549 taxonomy of the many different definitions and uses of the term ‘moral enhancement’, which had been
550 lacking to date, and by discussing several implications of particular definitions for the normative
551 debate on moral enhancement.

552

553 **REFERENCES**

554 1. Harris, John. 2011. Moral Enhancement and Freedom. *Bioethics* 25(2): 102–111.

555

556 2. Harris, John. 2012. What It’s Like to Be Good. *Cambridge Quarterly of Healthcare Ethics* 21(3):
557 293–305.

558

559 3. Wasserman, David. 2013. When bad people do good things: will moral enhancement make the
560 world a better place. *Journal of Medical Ethics*. doi:10.1136/medethics-2012-101094

561

562 4. Persson, Ingmar and Julian Savulescu. 2008. The Perils of Cognitive Enhancement and the Urgent
563 Imperative to Enhance the Moral Character of Humanity. *Journal of Applied Philosophy* 25(3): 162–
564 177.

565

566 5. Persson, Ingmar and Julian Savulescu. 2011. The turn for ultimate harm: a reply to Fenton. *Journal*
567 *of Medical Ethics* 37: 441–444.

568

569 6. Persson, Ingmar and Julian Savulescu. 2012. *Unfit for the Future: The Need for Moral*
570 *Enhancement*. Oxford: Oxford University Press.

571

- 572 7. Fenton, Elizabeth. 2010. The perils of failing to enhance: a response to Persson and Savulescu.
573 *Journal of Medical Ethics* 36: 148–151.
574
- 575 8. Shook, J.R. 2012. Neuroethics and the Possible Types of Moral Enhancement. *AJOB Neuroscience*,
576 3(4), pp.3–14.
577
- 578 9. DeGrazia, David. 2013 Moral enhancement, freedom, and what we (should) value in moral
579 behaviour. *Journal of Medical Ethics*. doi:10.1136/medethics-2012-101157
580
- 581 10. Hughes, James. 2011. After Happiness, Cyborg Virtue. *Free Inquiry* 32(1): 1–7.
582
- 583 11. Harris, John. 2013. “Ethics is for bad guys!” Putting the “moral” into moral enhancement.
584 *Bioethics* 27(3): 169–173.
585
- 586 12. Harris, John. 2013. Moral Progress and Moral Enhancement. *Bioethics* 27(5): 285–290.
587
- 588 13. Lev, Ori. 2012. Enhancing the Capacity for Moral Agency. *AJOB Neuroscience* 3(4): 20–22.
589
- 590 14. Agar, Nicholas. 2013. A question about defining moral bioenhancement. *Journal of Medical*
591 *Ethics*. doi:10.1136/medethics-2012-101153
592
- 593 15. Walker, Mark. 2009. Enhancing genetic virtue. *Politics and the Life Sciences* 28(2): 27–47.
594
- 595 16. Douglas, Thomas. 2008. Moral Enhancement. *Journal of Applied Philosophy* 25(3): 228–245.
596
- 597 17. Douglas, Thomas. 2011. Moral Enhancement Via Direct Emotion Modulation: A Reply to John
598 Harris. *Bioethics* 27(3): 160–168.
599

- 600 18. Parfit, Derek. 1989. *Reasons and Persons*. Oxford: Clarendon Press.
- 601
- 602 19. Faust, Halley S. 2008. Should we select for genetic moral enhancement? A thought experiment
603 using the MoralKinder (MK+) haplotype. *Theoretical Medicine and Bioethics* 29: 397–416.
- 604
- 605 20. Daniels, Norman. 2000. Normal Functioning and the Treatment-Enhancement Distinction.
606 *Cambridge Quarterly of Healthcare Ethics* 9: 309-322.
- 607
- 608 21. Agar, Nicholas. 2010. Enhancing genetic virtue? *Politics and the Life Sciences* 29(1): 73–75.
- 609
- 610 22. Horstkötter, Dorothee, Ron Berghmans, and Guido de Wert. 2010. Moral Enhancement for
611 Antisocial Behavior? An Uneasy Relationship. *AJOB Neuroscience* 3(4): 26–28.
- 612
- 613 23. Harris, John. 2013. Taking liberties with free fall. *Journal of Medical Ethics*.
614 doi:10.1136/medethics-2012-101092.
- 615
- 616 24. Jotterand, Fabrice. 2011 “Virtue Engineering” and Moral Agency: Will Post- Humans Still Need
617 the Virtues? *AJOB Neuroscience* 2(4): 3–9.
- 618
- 619 25. Simkulet, William. 2012. On Moral Enhancement. *AJOB Neuroscience* 3(4): 17–18.
- 620
- 621 26. Baertschi, Bernard. 2012. Neuromodulation in the Service of Moral Enhancement. *Brain*
622 *Topography*. doi:10.1007/s10548-012-0273-7.
- 623
- 624 27. Zarpentine, Chris. 2013. “The Thorny and Arduous Path of Moral Progress”: Moral Psychology
625 and Moral Enhancement. *Neuroethics* 6: 141–153.
- 626

- 627 28. Santoni de Sio, Filippo, Hannah Maslen, and Nadira Faulmüller. 2012. The Necessity of Objective
628 Standards for Moral Enhancement. *AJOB Neuroscience* 3(4): 15–16.
- 629
- 630 29. Pascual, Leo, Paulo Rodrigues, David Gallardo-Pujol. 2013. How does morality work in the brain?
631 A functional and structural perspective of moral behavior. *Frontiers in Integrative Neuroscience* 7(65).
632 doi: 10.3389/fnint.2013.00065.
- 633
- 634 30. Savulescu, Julian, Ingmar Persson. 2012. Moral Enhancement, Freedom, and the God Machine.
635 *The Monist* 95(3): 399–421.
- 636