Drugs *and* Hugs: Stimulating Moral Dispositions as a Method of Moral Enhancement

MICHAŁ KLINCEWICZ,[1] LILY EVA FRANK, AND MARTA SOKÓLSKA[2]

Abstract

Advocates of moral enhancement through pharmacological, genetic, or other direct interventions sometimes explicitly argue, or assume without argument, that traditional moral education and development is insufficient to bring about moral enhancement. Traditional moral education grounded in a Kohlbergian theory of moral development is indeed unsuitable for that task; however, the psychology of moral development and education has come a long way since then. Recent studies support the view that moral cognition is a higher-order process, unified at a functional level, and that a specific moral faculty does not exist. It is more likely that moral cognition involves a number of different mechanisms, each connected to other cognitive and affective processes. Taking this evidence into account, we propose a novel, empirically informed approach to moral development and education, in children and adults, which is based on a cognitive-affective approach to moral dispositions. This is an interpretative approach that derives from the cognitive-affective personality system (Mischel and Shoda, 1995). This conception individuates moral dispositions by reference to the cognitive and affective processes that realise them. Conceived of in this way, moral dispositions influence an agent's behaviour when they interact with situational factors, such as mood or social context. Understanding moral dispositions in this way lays the groundwork for proposing a range of indirect methods of moral enhancement, techniques that promise similar results as direct interventions whilst posing fewer risks.

1. Direct vs Indirect Methods of Moral Enhancement

The viability of the project of enhancing human morality by manipulating our brains with pharmaceuticals or direct brain stimulation (hereafter, "moral bioenhancement") is controversial among

©The Royal Institute of Philosophy and the contributors 2018

philosophers, neuroscientists, and psychologists. However, the motivations for widespread moral enhancement offered by its proponents remain persuasive. Our species may not be morally prepared to deal with rapid technological advance and a changing environment.[3] Traditional non-invasive forms of moral enhancement, such as moral education, but also more structural political and institutional changes, have not been given much attention in the moral enhancement debate, with some exceptions.[4] This is at least in part because the advocates of moral bioenhancement convincingly argue that moral education and other non-invasive means are not potent enough, do not act quickly enough, and have failed to work in the past. But pessimism about the efficacy of noninvasive means of moral enhancement is premature. We critically review promising methods and suggest new ones that take into account developmental psychology, neuroscience, and some themes from current debates in moral psychology.

In section 2 we review evidence that moral cognition[5] is multi-faceted and sketch some of the characteristics of that complexity. In section 3 we outline the cognitive-affective model of moral dispositions, recently articulated by Nancy Snow and Daniel C. Russell, among others. This model fits well with the complex picture of moral cognition discussed in section 2. If this model is plausible, we can then create non-invasive methods of moral enhancement that target moral dispositions of the individual. Section 4 presents recent insights into child development and gives examples of childrearing practices that can positively influence

---

[3]   I. Persson and J. Savulescu, 'Unfit for the Future? Human Nature, Scientific Progress, and the Need for Moral Enhancement', in J. Savulescu, R. ter Meulen, and G. Kahane (eds), *Enhancing Human Capacities* (Oxford: Wiley-Blackwell, 2011), 486–500.

[4]   See for example: O. Flanagan, *The Geography of Morals: Varieties of Moral Possibility* (Oxford: Oxford University Press, 2016); B. Fröding, *Virtue Ethics and Human Enhancement* (Cham: Springer Science & Business Media, 2012); B. Fröding and W. Osika, *Neuroenhancement: How Mental Training and Meditation Can Promote Epistemic Virtue* (Cham: Springer International Publishing, 2015); R. Powell and A. Buchanan, 'The Evolution of Moral Enhancement', in T. Coady, S. Sanyal, and A. Giubilini (eds), *The Ethics of Human Enhancement: Understanding the Debate* (Oxford: Oxford University Press, 2015).

[5]   We use the term "moral cognition" as a shorthand for all the psychological mechanisms that underlie moral behaviour. These include the capacity to have beliefs, emotions, and other mental states or processes, which may but need not be connected to morality. It is important to note that we use these terms without making any commitment to a particular view of moral psychology as inherently cognitive as opposed to affective.

330

children's moral development. Section 5 gives examples of non-invasive practices for adults that can take advantage of the cognitive-affective model of moral dispositions. Section 6.1 and 6.2 consider objections to our view and our replies.

2. Moral Cognition is Multi-Faceted

There are, broadly speaking, two main contemporary theoretical positions on the nature of moral cognition: the domain-specific and the domain-general approach.[6] Simplifying greatly, the domain-specific approach assumes that there is a psychological mechanism dedicated to moral functioning. The domain-general approach assumes that a dedicated mechanism does not exist and that moral functioning is realised by a number of other mechanisms.

According to the domain-specific approach, the moral mechanism (or module) receives external inputs from, for example, the self, intentions, or predictions about consequences of actions. On the basis of these inputs, the moral mechanism generates a particular moral output: a judgement, a reaction, a behaviour, etc.[7,8] Arguably, the inspiration for this approach is the notion of universal grammar in linguistics and the module that implements it. According to that theory, people acquire and develop their linguistic abilities because they have innate linguistic competencies, which interact with their linguistic environment. The domain-specific approach predicts that there is a similar set of competencies or dispositions to acquire morality.

John Rawls first brought attention to the possibility of applying the sort of arguments one finds supporting the universal grammar approach in linguistics to moral functioning.[9] Mikhail and Hauser take up this idea and argue that an adequate theory of morality will include an account of a computational mechanisms responsible for

[6]   L. Young and J. Dungan, 'Where in the Brain is Morality? Everywhere and Maybe Nowhere', *Social Neuroscience* **7**:1 (2012), 1–10.

[7]   M. Hauser, F. Cushman, L. Young, R. Kang-Xing Jin, and J. Mikhail, 'A Dissociation Between Moral Judgments and Justifications', *Mind & Language* **22**:1 (2007), 1–21.

[8]   J. Mikhail, 'Universal Moral Grammar: Theory, Evidence and the Future', *Trends in Cognitive Sciences* **11**:4 (2007), 143–152.

[9]   J. Rawls, *A Theory of Justice* (Cambridge, MA: Harvard University, 1971), 46–47.

bridging the gap between one's perception of an act and one's emotional/cognitive reaction to it and eventual judgement about it.

However, numerous studies that concern moral cognition show that no specific neural structures or neural functions for morality exist.[10] This leads us to the alternative domain-general approach, which is the view that no specific mechanism or module for moral cognition exists. Instead, a number of other mechanisms, such as the mechanisms of reasoning or affect, realise what we characterise as moral cognition.[11,12] This approach is general in the sense that it assumes that moral cognition, moral development, and the process of making moral judgements are unified at a functional level, but not on a biological level.

On the domain-general approach, human morality is likely to be realised by a higher-order cognitive function. If so, then understanding the mechanisms of moral cognition demands an analysis of the relationship between lower-order mechanisms and how they interact to create moral judgements, dispositions, and behaviours at a higher level of psychological organisation.[13,14] Moral cognition on this view is multi-faceted because it is a higher-order domain-general function involving several lower-level functions.

The other relevant feature of moral cognition is that it is affected by many factors, some outside of the moral domain, and often without the individual's awareness.[15] One's social context can impact on one's hierarchy of values, as well as the way in which one orders the

[10]   For a review see: J. D. Greene, 'The Cognitive Neuroscience of Moral Judgment and Decision Making', in M. S. Gazzaniga (ed.), *The Cognitive Neurosciences*, 5th edn (Cambridge, MA: MIT Press, 2014), 1013–1023. See also J. D. Greene, 'The Rise of Moral Cognition', *Cognition* **135**, 39–42.

[11]   J. Moll, R. Zahn, R. de Oliveira-Souza, F. Krueger, and J. Grafman, 'The Neural Basis of Human Moral Cognition', *Nature Reviews Neuroscience* **6**:10 (2005), 799–809.

[12]   J. Moll, D. Oliveira-Souza, and R. Zahn, 'The Neural Basis of Moral Cognition', *Annals of the New York Academy of Sciences* **1124**:1 (2008), 161–180.

[13]   W. D. Casebeer and P. S. Churchland, 'The Neural Mechanisms of Moral Cognition: A Multiple-Aspect Approach to Moral Judgment and Decision-Making', *Biology and Philosophy* **18**:1 (2003), 169–194.

[14]   J. Prinz, 'Is Morality Innate?', in W. Sinnott-Armstrong (ed.), *Moral Psychology,* Vol. 1 (Cambridge, MA: MIT Press, 2008), 367–406.

[15]   J. J. Van Bavel, O. Feldman Hall, and P. Mende-Siedlecki, 'The Neuroscience of Moral Cognition: From Dual Processes to Dynamic Systems', *Current Opinion in Psychology* **6** (2015), 167–172.

needs of others and one's own needs.[16] Other experiments demonstrate that moral decisions preceded by conversations that are morally oriented are more likely to result in participants telling the truth than when the conversations are self-oriented.[17]

The third relevant feature of moral cognition is that it varies between individuals[18] and within individuals over time.[19,20] Furthermore, different cognitive mechanisms are crucial in the exercise of moral abilities at different times during human development. While affective structures are the most relevant in children, in adults structures connected to executive control and other higher order processes play a bigger role by integrating information about one's affective reactions with inferences about the mental states of others.[21]

This brief sketch of current empirical research and theoretical approaches to moral cognition and moral development strongly suggests that they involve many disparate mechanisms. Their complexity casts doubt on the idea that a specific pharmacological, neural, or genetic intervention will consistently yield moral enhancement on a wide scale.[22] This is because we cannot confidently state that each

[16]   E. Sharma, N. Mazar, A. L. Alter, and D. Ariely, 'Financial Deprivation Selectively Shifts Moral Standards and Compromises Moral Decisions', *Organizational Behavior and Human Decision Processes* **123**:2 (2014), 90–100.
[17]   B. C. Gunia, L. Wang, L. I. Huang, J. Wang, and J. K. Murnighan, 'Contemplation and Conversation: Subtle Influences on Moral Decision Making', *Academy of Management Journal* **55**:1 (2012), 13–33.
[18]   K. Prehn, I. Wartenburger, K. Mériau, C. Scheibe, O. R. Goodenough, A. Villringer, and H. R. Heekeren, 'Individual Differences in Moral Judgment Competence Influence Neural Correlates of Socio-Normative Judgments', *Social Cognitive and Affective Neuroscience* **3**:1 (2008), 33–46.
[19]   W. Edelstein, M. Keller, and E. Schröder, 'Child Development and Social Structure: A Longitudinal Study of Individual Differences', in Paul B. Baltes, David L. Featherman, Richard M. Lerner (eds), *Life-Span Development and Behavior*, Vol. 10 (London: Psychology Press, 2016), 151–185.
[20]   R. Svensson, L. J. Pauwels, F. M. Weerman, and G. J. Bruinsma, 'Explaining Individual Changes in Moral Values and Moral Emotions Among Adolescent Boys and Girls: A Fixed-Effects Analysis', *European Journal of Criminology* **14**:3 (2017), 290–308.
[21]   J. Decety and L. H. Howard, 'The Role of Affect in the Neurodevelopment of Morality', *Child Development Perspectives* **7**:1 (2013), 49–54, 49.
[22]   H. Wiseman, *The Myth of the Moral Brain: The Limits of Moral Enhancement* (Cambridge, MA: MIT Press, 2016).

individual's cognitive and affective systems will be identically impacted by an intervention. The system that underlies moral cognition is unlike the serotonin reuptake system which can be predictably modulated by an appropriate intervention.

3. The Cognitive-Affective Conception of Moral Dispositions

If moral cognition is a domain-general, higher-order, dynamic faculty, influenced by a variety of non-moral factors, then it is also difficult to accept the existence of permanent moral traits. People are not morally good, morally bad, honest, or cowardly, and so on, except in particular situations. If that is true, then moral enhancement should not aim at improving moral traits, but other, more general faculties or at improving the environment itself. Limiting situational factors that contribute to immoral behaviour may be the best way towards moral enhancement.

This line of thinking is advanced in the situationist challenge to the existence of moral virtues.[23,24] Situationists rely on evidence from cognitive and social psychology, citing experiments that show our judgements and behaviours come under the influence of unconscious processes and situational factors.[25] Mark Alfano created a taxonomy of the kinds of factors that interfere with moral judgement and behaviour: 'bad reasons, situational non-reasons, and non-moral individual differences' and identifies situational non-reasons as the most troubling for the existence of moral character traits.[26] These are factors like ambient smells, sounds, and mood.[27] This evidence undermines the vision of a thoughtful moral reasoner weighing her options and also seems to undermine the classic virtue ethics picture of the individual's virtues responsible for consistent behaviour across a range of situations.

On the other hand, there is also a growing body of research critiquing situationism and the situationist interpretation of the relevant

[23]   G. Harman, 'Skepticism About Character Traits', *The Journal of Ethics* **13**:2/3 (2009), 235–242.
[24]   J. M. Doris, *Lack of Character: Personality and Moral Behaviour* (Cambridge: Cambridge University Press, 2002).
[25]   C. A. Anderson, W. E. Deuser, and K. M. DeNeve, 'Hot Temperatures, Hostile Affect, Hostile Cognition, and Arousal: Tests of a General Model of Affective Aggression', *Personality and Social Psychology Bulletin* **21**:5 (1995), 434–448.
[26]   M. Alfano, *Character as Moral Fiction* (Cambridge: Cambridge University Press, 2013), 40.
[27]   Alfano, *Character as Moral Fiction*, 44–46.

334

experiments.[28] The upshot of this research can be simplified as follows. Firstly, some people, even under external pressure or conditions which tend to stimulate immoral acts, choose the action considered moral nonetheless. Secondly, moral dispositions should not be understood merely in behavioural terms, that is, as behaviour that is recurrent in a specific type of situation. Instead, moral dispositions should be understood as abilities that allow the individual to interpret and to adapt to situations in appropriate ways.

Drawing similar conclusions, Snow,[29] Russell,[30] and Natasza Szutta[31] advocate a theory of moral traits as a subset of personality traits in the cognitive-affective personality system (CAPS).[32] CAPS characterises personality traits as dispositions that are the effect of longitudinal interactions between endogenous factors (mood, beliefs, etc.) and exogenous factors (situation, context, etc.). On this view, dispositions are constitutive parts of one's personality and mediate interaction with exogenous factors. One's personality traits are expressed in the constant interaction of individual and situational factors, where that expression exhibits relatively stable patterns of behaviour across situations.

Szutta (2015) argues that CAPS accords with the observation that most people are relatively consistent in their behaviour, if we take into consideration their distinctive cognitive-affective system and the way

[28]   S. Arjoon, 'Ethical Decision-Making: A Case for the Triple Font Theory', *Journal of Business Ethics* **71**:4 (2007), 395–410; E. J. Wielenberg, 'Saving Character', *Ethical Theory and Moral Practice* **9**:4 (2006), 461–491; J. J. Kupperman, 'The Indispensability of Character', *Philosophy* **76**:2 (2001), 239–250; R. Kamtekar, 'Situationism and Virtue Ethics on the Content of Our Character', *Ethics* **114**:3 (2004), 458–491; E. S. Radcliffe, 'Moral Naturalism and the Possibility of Making Ourselves Better', in B. K. Wilburn (ed.), *Moral Cultivation: Essays on the Development of Character and Virtue* (London: Rowman & Littlefield, 2010), 101–24; J. Sabini and M. Silver, 'Lack of Character? Situationism Critiqued', *Ethics* **115**:3 (2005), 535–562; G. Sreenivasan, 'The Situationist Critique of Virtue Ethics', in D. C. Russell (ed.), *The Cambridge Companion to Virtue Ethics* (Cambridge: Cambridge University Press, 2013), 290–314.

[29]   N. E. Snow, *Virtue as Social Intelligence: An Empirically Grounded Theory* (Abingdon: Routledge, 2010).

[30]   Daniel C. Russell, *Practical Intelligence and the Virtues* (Oxford: Oxford University Press, 2009).

[31]   N. Szutta and A. Szutta (eds), *W poszukiwaniu moralnego charakteru*, Vol.1 (Lublin: Wydawnictwo Academicon, 2015).

[32]   W. Mischel and Y. Shoda, 'A Cognitive-Affective System Theory of Personality: Reconceptualizing Situations, Dispositions, Dynamics, and Invariance in Personality Structure', *Psychological Review* **102**:2 (1995), 246–268.

they construe their own situation. For example, consider the case of imposter syndrome.[33] Kate, a hypothetical woman suffering from it, is an intelligent person, but also extremely insecure. She was often criticised by her parents as a child and young adult, and her society discriminates against women. From the perspective of an external observer, it appears that Kate does not have much to say. But when we learn about her past and her fears, it becomes difficult to view her behaviour in the same way: now we know she suffers from imposter syndrome, which can be understood as a trait. Attributing this trait to Kate allows us to make better sense of her behaviour and make better predictions about her.

Generalising, personality traits can be characterised by "if… then…" profiles of behaviour, which are relatively stable, but situationally dependent. Just as we can say solubility in water is a property of salt, we can call the relatively stable patterns of human moral behaviour traits, without having to say much about what underlies them. When human behaviour is inconsistent (as it often is), the trait remains. This consistent with Snow who argues that 'social-cognitivists stress the importance that the meanings of situations have for people, and claim that evidence of personality coherence can be found by paying attention to those meanings'.[34]

Returning to Kate, consider the following generalisation about her: when she is around people who she thinks will criticise her and threaten her sense of security, her characteristic way of reacting in these situations will be activated, and she will stay silent. However, when she feels comfortable, for example, among her close friends, she is not afraid to give her opinion, because there is nothing that triggers her insecurity. Nonetheless, her imposter syndrome remains.

So much is suggested by the one of the original Mischel and Shoda experiments that grounds CAPS. The experiment, conducted at a residential summer camp for children, investigated, among other things, the situations in which specific behaviours occur. Their aim was to capture psychological features, such as the subject's attitude toward specific circumstances, that triggered the behaviour. They observed the children's behaviours 'on selected dimensions' (e.g., verbal aggression, withdrawal, friendly, prosocial behaviour) and then made predictions about patterns in children's behaviour,

---

[33]   John Kolligian, Jr and Robert J. Sternberg, 'Perceived Fraudulence in Young Adults: Is There an "Imposter Syndrome"?', *Journal of Personality Assessment* **56**:2 (1991), 308–326.
[34]   Snow, *Virtue as Social Intelligence: An Empirically Grounded Theory*, 38.

336

based on an individual 'profile of situation-behaviour relationships'.[35] These profiles were accurate predictors of the ways children reacted in various situations. Their behavioural responses depended on the child's interpretation of their situation and were consistent with their "if…then…" patterns of reacting.

These sorts of results demonstrate that making generalisations about patterns of behaviour by attributing traits has significant predictive power. In the same way, before we knew the chemical composition of salt, we could still predict how it would "behave" by attributing to it a certain trait. If salt is put into water and it does not dissolve, this does not mean that salt lacks the relevant trait. The situation may have changed and something may be preventing the trait from expressing itself. If we know enough about the situation we may also be able to predict that salt will not dissolve, though we cannot explain why it will not in terms of the relevant chemical reaction.

In similar vein, Snow argues that CAPS captures local personality traits which can be characterised as "virtues" and "flaws". Further, she argues that personality traits can become more global and can be developed through individual conscious reflection. For example, when Kate realises that she feels very sorry for impoverished children in her town, she can ask herself why she does not have the same feelings toward children in other regions, who are significantly worse off. Resulting from this reflection, she decides to give more attention to the suffering of distant children. She may then notice the common, morally relevant features shared by all suffering children.

According to Snow, desirable traits can also be developed by limiting related vices. This is corroborated by Patricia Devine and Margo Monteith's research on methods for decreasing the influence of negative stereotypes in the individual.[36,37] Their research on the Behavioural Inhibition System makes sense of cases in which it is possible to change stereotypical thinking through inhibition of intermediary processes, such as negative thoughts and emotions related to the stereotype. The goal is to stop stereotypical behaviour and let one act with more self-awareness in line with one's updated beliefs.

[35] Mischel and Shoda, 'A Cognitive-Affective System Theory of Personality', 249.

[36] M. J. Monteith, J. W. Sherman, and P. G. Devine, 'Suppression as a Stereotype Control Strategy', *Personality and Social Psychology Review* **2**:1 (1998), 63–82.

[37] P. G. Devine and M. J. Monteith, 'Automaticity and Control in Stereotyping', in Shelly Chaiken and Yaacov Trope (eds), *Dual-Process Theories in Social Psychology* (New York: Guildford Press, 1999), 339–360.

337

It is easy to generate examples of this practice. For instance, someone who was raised in an environment where one ethnic group was discriminated against will likely end up with the disposition to behaviourally express negative stereotypes about this group. However, if Snow is right, she can change her attitude, if she consciously aims to rid herself of stereotypical reactions.

Snow refers to these practices as cultivation of moral virtue. Bracketing the notion of virtue, understood as a moral quality of a person, we can treat cultivation as a method of enhancement of traits. Traits can be developed and improved over time with the sort of deliberate practices that Snow suggests, even if we completely abandon the idea of virtues as they feature in context of virtue ethics.

Snow also points out that it is possible to develop traits intentionally and automate moral behaviour though automatic goal-dependent actions; these actions are voluntarily initiated and lead to a definite goal. While the goal activation is unconscious, the reactions themselves are both intelligent and situationally sensitive. Snow connects this research to the concept of virtue, but it seems equally justified to connect it to the notion of moral dispositions, as we have used it here. Returning to to Kate, she feels compassion toward poor children and consciously decides to help them. Helping becomes her goal and over time whenever she sees an impoverished child she thinks about how best to help. In other words, she intentionally developed a context-sensitive disposition that eventually becomes automatic.

All of these presented ways are good examples of how to make local CAPS traits global, in the sense of expanding the contexts in which they manifest. This process will involve, if this research is right, at least some conscious reflection or deliberate choice of a goal. This is why Snow's proposal, for example, is likely to work best for adults in whom higher cognitive functions such as reasoning or cognitive control are more developed, but not for young children, whose abstract thinking ability is limited.

Snow and Russell use CAPS as a rebuttal to the situationist challenge to virtue ethics and as an example of a new and improved way of talking about moral personality traits, which is consistent with the empirical evidence marshalled by situationists. Their approach is a significant improvement on the classic virtue ethics approach at least in part because it has the benefit of making sense of people acting morally in relatively consistent ways across a variety of situations, while at the same capturing the insight that these traits are highly context sensitive.

Their view is also an improvement over the approaches to moral enhancement open to situationists. A situationist would argue that the way to improve morally would be to change the situations in which

338

people find themselves. In other words, the focus should be on moral technologies that improve the environment and not technologies and interventions that target moral agents and their dispositions or traits. This approach intentionally ignores the idea of stable traits.

If CAPS is an improvement on the traditional conceptualisation of traits, we can take advantage of it to find strategies for improvement of the agent. But we take a more cautious approach than the virtue ethicists just mentioned and remain as neutral as possible in their debate with situationists. For us, CAPS is a good guide to the *mechanisms* that are potential candidates for targeted non-invasive methods of moral enhancement.

The key to our model is identification of traits (A, B, and C in Figure 1). It should be emphasised that on our view traits are not qualities of the person or virtues in the classical sense. Traits are characterisations of context-sensitive dispositions, which, in turn, are realised in particular individuals in particular ways. This means that, to some extent, they are useful fictions.

However, unlike mere fictions, understanding their complex relationships can tell us a lot about potential non-invasive strategies for individual moral enhancement. On our view, the best non-invasive strategies of enhancement are going to focus on components of moral cognition that are relevant to the realisation of dispositions. Interventions should start early, with child development. This is the first set of strategies we consider.

4. Child Development and Moral Cognition

There are significant differences in the processing of morally-laden stimuli in children and adults. Jean Decety and Lauren Howard's line of research sheds light on the relationship between particular emotions, motivations, and other types of neural activation in people exposed to morally laden stimuli.[38] Their results suggest that affect is key in early development and that higher order cognitive mechanisms, such as cognitive control and response inhibition, have greater influence in adults. This yields an important insight: higher-order cognitive abilities have a larger role in adults, whose patterns of processing emotions are different.

Supporting this view, Decety and Howard point to research on what may be described as precursors of moral dispositions in

[38] J. Decety and L. H. Howard, 'The Role of Affect in the Neurodevelopment of Morality', *Child Development Perspectives* **7**:1 (2013), 49–54.

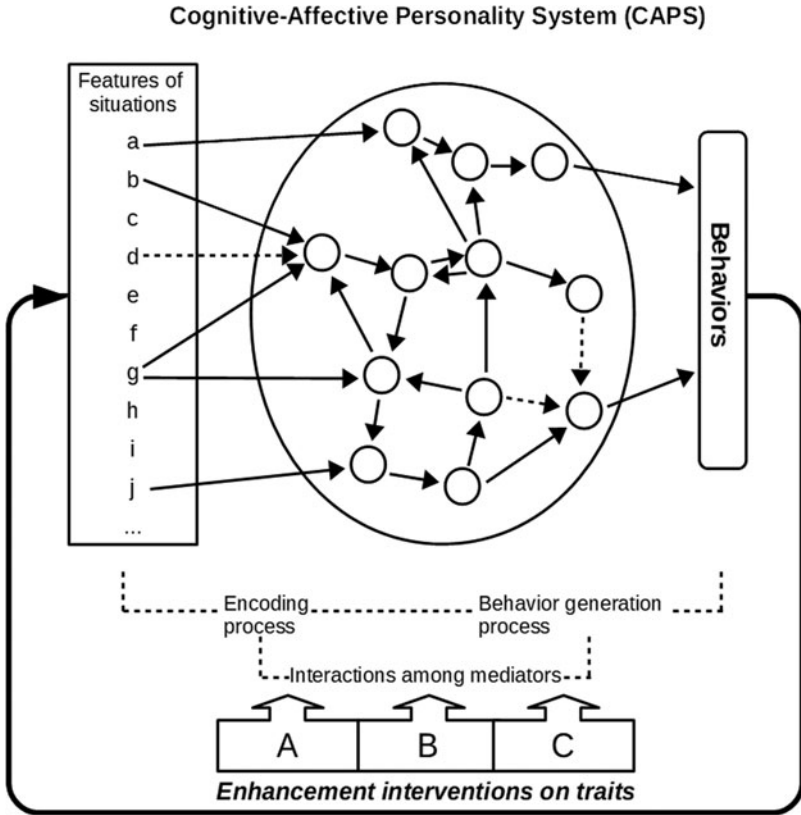**Cognitive-Affective Personality System (CAPS)**

Figure 1. (Adopted from Mischel and Shoda, 1995.) Simplified illustration of types of cognitive-affective mediating processes that generate an individual's distinctive behaviour patterns. Non-invasive enhancement interventions on traits A, B, and C aim at strengthening or inhibiting interactions among mediating units, which together comprise a stable network of relations that characterise an individual.

infants. In early development affective mechanisms play a key role with relatively little involvement of cognitive mechanisms, which at that time, are under-developed. Bandstra, et al. demonstrate that young children display a much higher level of personal distress when witnessing someone's sadness than older children. Older children display higher levels of sympathy than personal distress.[39]

[39]   N. F. Bandstra, C. T. Chambers, P. J. McGrath and C. Moore, 'The Behavioural Expression of Empathy to Others' Pain Versus Others' Sadness in Young Children', *Pain* **152**:5 (2011), 1074–1082.

340

This is one example, among many, of the ways that development of aspects of childrens' social cognition is connected to the rate of maturation of the prefrontal cortex. Again, it reinforces that moral cognition is not based on any morally-specific neural structures, but is rather an integrative cognitive function. It further shows that stimulating moral cognition in children should focus on mediating factors that are anatomically and functionally different from those that underlie moral abilities in adults.

Furthermore, in young children (4–8 years old) there is stronger (than in adults) functional connectivity between ventromedial prefrontal cortex (vmPFC) and the brainstem, whereas in 18–25 year-olds, functional connectivity is stronger between vmPFC and posterior superior temporal sulcus/amygdala. The idea that emotions play a key role in moral development in the early life is supported by evidence of greater activation of parts of brain regions responsible for affect at younger ages. This is corroborated by lesion studies, which show that early damage to the vmPFC and amygdala cause much more significant problems in processing morally-laden stimuli than similar damage later in life.[40,41] In sum, this research confirms the platitude that emotional development early in life is crucial for moral functioning later in life.

Caregivers and children's relation to them play also an important role in moral development.[42,43] For example, in families with a higher level of emotional warmth, children are more altruistic.[44] Inconsistent child-rearing strategies leads to depression, anxiety, lowered social abilities, or aggression.[45] Attitudes of the parents

---

[40] P. Shaw, E. J. Lawrence, C. Radbourne, J. Bramham, C. E. Polkey, and A. S. David, 'The Impact of Early and Late Damage to the Human Amygdala on "Theory of Mind" Reasoning', *Brain* **127**:7 (2004), 1535–1548.

[41] S. W. Anderson, A. Bechara, H. Damasio, D. Tranel, and A. R. Damasio, 'Impairment of Social and Moral Behaviour Related to Early Damage in Human Prefrontal Cortex', *Nature Neuroscience* **2**:11 (1999), 1032–1037.

[42] M. Christen and D. Narvaez, 'Moral Development in Early Childhood is Key for Moral Enhancement', *AJOB Neuroscience* **3**:4 (2012), 25–26.

[43] J. M. Cowell and J. Decety, 'Precursors to Morality in Development as a Complex Interplay Between Neural, Socioenvironmental, and Behavioural Facets', *Proceedings of the National Academy of Sciences* **112**: 41 (2015), 12657–12662.

[44] G. H. Brody and D. R. Shaffer, 'Contributions of Parents and Peers to Children's Moral Socialisation', *Developmental Review* **2**:1 (1982), 31–75.

[45] A. Schore, 'Bowlby's "Environment of Evolutionary Adaptedness": Recent Studies on the Interpersonal Neurobiology of Attachment and

and their engagement and reactivity to their child's behaviour, help develop a sense of conscience, and result in a higher likelihood of the child developing proper moral attitudes.

There are also practices that may help in the development of particular traits: breastfeeding for empathy, immediate response to the child's needs for conscience, touch for control of one's behaviour, letting the mother stay with the child right after birth for self-regulation, play for lowering aggression, and parental support for resistance to stress.[46] Additionally parental attitudes and values play an important role because of the attitudes they convey to the child. Overall, value transmission from parent to child seems to be the result of multifarious interactions.

None of the processes that have an effect on a child's personality development are likely to occur in isolation from others, and are probably best understood as biology and environment dynamically interacting. On the one hand, innate emotional and cognitive predispositions are important in some models of interaction between the parent and the child. On the other hand, the environment is equally important. Relevant work done on moral technologies, technologies for behaviour change, nudges, etc. is useful. In sum, unlike direct, biomedical methods of moral enhancement, those grounded in education and the early-life shaping of moral attitudes and their character allow for a more nuanced and individualised approach.

5. Non-Invasive Methods of Moral Enhancement in Adults and Older Children

Targeted moral improvement in adults by non-invasive means can take many forms.[47] One particularly promising method focusses on affective perspective taking, sometimes considered to be constitutive of empathy.[48] One strategy for improving affective perspective taking is reading fiction, which increases the ability to understand others'

Emotional Development', in D. Narvaez, J. Panksepp, A. Schore, and T. Gleason (eds), *Evolution, Early Experience and Human Development* (New York: Oxford University Press, 2013) 31–67.

[46]   Christen and Narvaez, 'Moral Development in Early Childhood', 25.

[47]   For example: D. Narvaez and T. Bock, 'Developing Ethical Expertise and Moral Personalities', *Handbook of Moral and Character Education* (Abingdon: Routledge, 2014), 140–158.

[48]   J. Decety and J. M. Cowell, 'The Complex Relation Between Morality and Empathy', *Trends in Cognitive Sciences* **18**:7 (2014), 337–339.

342

mental states and increases empathic concern.[49] The mediating mechanism that underlies development of these traits is likely to be theory of mind (ToM).[50] Reading is an effective part of programmes aimed at promoting social skills, treating personality disorders, and increasing empathy among groups ranging from prisoners to medical students and physicians.[51] In prisons the role of books in improving ToM skills can have a significant impact on inmates' rehabilitation. At the same time other art-forms, such as films and theatre, may similarly facilitate empathy and identification with the experiences of others.[52,53]

Another possible non-invasive method of moral enhancement is Forum Theatre – a version of the Theatre of the Oppressed.[54] In Forum Theatre the play is always presented twice – the first time to show "anti-model", which embodies the morally, politically, or socially problematic and oppressive state of affairs and the second time to give the audience an opportunity to react and intervene. Audience members can stop the performance and present their own solution to the conflict presented on stage.

What is crucial in Forum Theatre is that the division between the audience and the performers is partially erased. The spectators can at

[49] D. Carr, 'On the Contribution of Literature and the Arts to the Educational Cultivation of Moral Virtue, Feeling and Emotion', *Journal of Moral Education* **34**:2 (2005), 137–151; D. R. Johnson, G. K. Cushman, L. A. Borden, and M. S. McCune, 'Potentiating Empathic Growth: Generating Imagery While Reading Fiction Increases Empathy and Prosocial Behaviour', *Psychology of Aesthetics, Creativity, and the Arts* **7**:3 (2013), 306–312.

[50] D. C. Kidd and E. Castano, 'Reading Literary Fiction Improves Theory of Mind', *Science* **342**:6156 (2013), 377–380.

[51] J. Billington, '"Reading for Life": Prison Reading Groups in Practice and Theory', *Critical Survey* **23**:3 (2011), 67–85; J. Billington, E. Longden, and J. Robinson, 'A Literature-Based Intervention for Women Prisoners: Preliminary Findings', *International Journal of Prisoner Health* **12**:4 (2016), 230–243; J. F. Shapiro, 'Using Literature and the Arts to Develop Empathy in Medical Students', in F. D. T. Farrow and P. Woodruff (eds), *Empathy in Mental Illness* (Cambridge: Cambridge University Press, 2007), 473–494.

[52] J. M. Reilly, J. Trial, D. E. Piver, and P. B. Schaff, 'Using Theater to Increase Empathy Training in Medical Students', *Journal for Learning Through the Arts: A Research Journal on Arts Integration in Schools and Communities* **8**:1 (2012): https://escholarship.org/uc/item/68x7949t.

[53] A. Eisenberg, S. Rosenthal, and Y. R. Schlussel, 'Medicine as a Performing Art: What We Can Learn About Empathic Communication from Theater Arts', *Academic Medicine* **90**:3 (2015), 272–276.

[54] Augusto Boal, 'Theatre of the Oppressed', trans. by Charles A. McBride and Maria-Odilia Leal McBride (New York: Theatre Communication Group, 1985), 144–45.

any point become the actors, replacing an actor currently on stage.[55] The audience is invited to engage in affective perspective-taking. They are asked to broaden their point of view and consequently their concern for other, sometimes very different, people.

Forum Theatre has been shown to be very effective in developing the relevant attitudes towards victims of sexual assault in an empirical study on 561 students, which

> showed that the participants exposed to [a Forum Theatre] performance, in comparison to a sexual assault lecture and a control group, reported greater perceived self-efficacy in: perspective taking, emotional contagion, empathic concern, and comforting behavior toward potential sexually assault survivors.[56]

Extensive evaluation studies in other domains are not available, but it seems reasonable to expect that being involved in a scene, and becoming an oppressed character on stage may trigger the relevant mechanism of perspective taking in general. This rate of success likely generalises to other domains from issues of sexual assault.

The other capacity that can be influenced through Forum Theatre is an ability to notice morally relevant problems in daily life and to react in the best possible ways. Again, it is important to emphasise that this method was not evaluated scientifically; however, its form suggests that it could be an appropriate tool for enhancement of particular components of moral cognition. It is also safe to assume that this method can be modified to focus on specific moral problems, such as environmental degradation or global poverty.

There is some reason to think that stimulating empathy inhibits aggressive behaviours and elicits altruism.[57] This idea has been successfully applied in childhood education[58] and in prison rehabilitation programmes.[59] In some classrooms, for instance, children care for

---

[55]   The Forum Project: http://theforumproject.org/whatistoforumtheatre/.

[56]   J. I. Rodríguez, et al. 'Assessing the Impact of Augusto Boal's "Proactive Performance": An Embodied Approach for Cultivating Prosocial Responses to Sexual Assault', *Text and Performance Quarterly* **26**:3 (2006), 229–252, 245.

[57]   C. Zahn-Waxler and M. Radke-Yarrow, 'The Origins of Empathic Concern', *Motivation and Emotion* **14**:2 (1990), 107–130.

[58]   B. Daly and S. Suggs, 'Teachers' Experiences with Humane Education and Animals in the Elementary Classroom: Implications for Empathy Development', *Journal of Moral Education* **39**:1 (2010), 101–112.

[59]   R. Kohl and A. Wenner, *State Comparison of Life Sentenced Inmates* (Milford, MA: Massachusetts Department of Correction, Office of Strategic Planning and Research, 2012).

344

and interact with a pet, encouraging them to develop a sense of responsibility for the creature's well-being. When pets are also subjects of discussion empathy for the animals is stimulated.[60]

In a growing number of prison programmes inmates have an opportunity to care for an animal: for example, service dog training.[61] In the Wild Horse Inmate Program,[62] supervised prisoners train and gently break wild horses. The bond this creates with a horse increases the inmates' empathic concern and sense of responsibility, in addition to giving them practical skills.[63]

One reason to think that animals are useful in moral development programmes is evidence that many overlapping brain-regions are activated while looking pictures of humans *and* animals in distress.[64] Creating and maintaining a relationship with an animal positively affects the development of empathy in children.[65] This confirms the link between aggressive behaviours toward animals in childhood and similar behaviours toward other people later in life.[66] Aggression levels are negatively correlated with empathy levels, so that a person with a higher level of empathy will be less inclined to hurt others.[67]

---

[60] Daly and Suggs, 'Teachers' Experiences with Humane Education', 105–106.

[61] C. Deaton, 'Humanizing Prisons with Animals: A Closer Look at "Cell Dogs" and Horse Programs in Correctional Institutions', *Journal of Correctional Education*, 46–62.

[62] Colorado Correctional Industries, 'Wild Horse Inmate Program': https://www.coloradoci.com/serviceproviders/whip/.

[63] Arizona Correctional Industries, 'Wild Horse Program': https://aci.az.gov/wild-horse-program/.

[64] R. G. Franklin, Jr, A. J. Nelson, M. Baker, J. E. Beeney, T. K. Vescio, A. Lenz-Watson, and R. B. Adams, Jr, 'Neural Responses to Perceiving Suffering in Humans and Animals', *Social Neuroscience* **8**:3 (2013), 217–227.

[65] Thompson, Gullone, Daly, and Suggs refer to the project of *humane education* which focusses on issues such as social justice and citizenship and which is meant to encourage people to respect and be compassionate toward all beings, human and nonhuman, and to take into consideration the environment and its welfare. However, the benefits of using pets, presented by the mentioned authors, make it an interesting method also for moral education, which are not directly connected to *humane education* in its original presentation. See notes 68 and 69 for full citations.

[66] M. Robin and R. T. Bensel, 'Pets and the Socialization of Children', *Marriage & Family Review* **8**:3/4 (1985), 63–78.

[67] P. A. Miller and N. Eisenberg, 'The Relation of Empathy to Aggressive and Externalizing/Antisocial Behaviour', *Psychological Bulletin* **103**:3 (1988), 324–344.

There is also evidence of the impact that interaction with an animal has on emotional development and empathic concern.[68] Most teachers who had classes incorporating animals observed higher levels of empathy and faster socio-emotional development.[69] However, there is no systematic research on the effectiveness of animal use in rehabilitation programmes.[70] What we have is anecdotal evidence from prison staff and inmates, who are overwhelmingly positive about using animals as components of rehabilitation.

6.1. Objections and Replies: Cognitive Demands

One should note that Russell, Szutta, and Snow's application of the CAPS system to moral traits has been subject to various criticisms,[71] to which virtue ethicists have responded.[72] The debate is by no means over. Since we do not adopt an interpretation of traits as moral virtues, but merely treat traits as guides to an individual's components of moral cognition, we will not address these objections here, save one concerning self-control and integrity, raised by Laura Papish.[73]

Papish points out that a central ingredient in Snow's account of the development of virtue is the claim that humans desire internal mental and emotional consistency in moral and non-moral matters,

[68]   K. L. Thompson and E. Gullone, 'Promotion of Empathy and Prosocial Behaviour in Children Through Humane Education', *Australian Psychologist* **38**:3 (2003), 175–182.

[69]   B. Daly and S. Suggs, 'Teachers' Experiences with Humane Education and Animals in the Elementary Classroom: Implications for Empathy Development', *Journal of Moral Education* **39**:1 (2010), 101–112.

[70]   R. Kohl and A. Wenner, *State Comparison of Life Sentenced Inmates* (Milford, MA: Massachusetts Department of Correction, Office of Strategic Planning and Research, 2012).

[71]   M. Alfano and D. Russell, 'Review of "Practical Intelligence and the Virtues"', *Ethical Theory and Moral Practice* **16**:3, 671–673; C. Miller, 'Does the CAPS Model Improve Our Understanding of Personality and Character?', in J. Webber and A. Masala (eds), *From Personality to Virtue* (Oxford: Oxford University Press, 2016), 155–185.

[72]   R. West, 'Virtue Ethics is Empirically Adequate: A Defense of the CAPS Response to Situationism', *Pacific Philosophical Quarterly* **99**:S1 (2018), 79–111.

[73]   L. Papish 'CAPS Psychology and the Empirical Adequacy of Aristotelian Virtue Ethics', *Ethical Theory and Moral Practice* **20**:3 (2017), 537–549.

yet it is not obvious that people actually have such a desire. Our ability to act in ways that are inconsistent with our values is evident even in participants of the Milgram experiment who report being very disturbed and distressed by their own behaviour as they shocked the learners. Our everyday experiences of inner moral conflict also suggest that we are actually fairly comfortable maintaining moral conflicts. If that is the case, then any theory that assumes that we aim at coherent moral integration in our moral lives depends on a questionable assumption about moral psychology.

Furthermore, suppose that we do value integration and that we are capable of achieving it, contrary to evidence that we do not and that we often cannot. If so, then we would move towards a morally praiseworthy personality trait (or virtue) only if we revise and integrate it in the right way. For example, we choose to try to rid ourselves of negative stereotypes rather than the egalitarian values with which they conflict. This also means, as Papish points out, that we have to count on people holding some set of morally right beliefs in the first place. Unfortunately, this may be an unreasonable assumption about anyone and difficult to verify, even if true.

This line of objection is relevant to our use of the CAPS model to inform indirect moral enhancement methods in the following ways: 1) if we do not strive for integrated moral personalities, then some of the methods suggested may not be effective, and 2) if we do strive for integrated moral personalities, then moral enhancement will depend on individuals holding at least some morally right beliefs to begin with and working towards an integration with those, which we cannot depend on happening.

Our first response is to concede to point (1). Conscious self-reflection and a motivation to have a consistent moral personality are prerequisites for strategies initiated by the individual, as in the case of Kate who holds stereotypes. These are unlikely to work without conscious self-reflection and the relevant motivation. However, we further point out that these methods explicitly endorsed by Snow and others are not a part of all the indirect methods we suggest.

Early childhood interventions, participatory theatre, reading fiction, and animal therapy do not demand conscious self-reflection on one's moral beliefs or a motivation to have an integrated moral personality. Instead, these methods focus on the mediating mechanisms that underlie moral dispositions. For example, on developing the capacity to take the perspective of others through enactment in Forum Theatre or on developing empathic concern by gently

347

breaking horses. Engaging in these practices morally enhances an individual indirectly, by improving capacities that underlie moral cognition in general.

This still leaves point (2) which suggests some limits to our proposal that Papish's objections bring into relief. Indirect methods are unlikely to be effective in changing moral dispositions of people with personalities constituted mostly by traits that dispose them to act immorally. This is perhaps also why it is good to be cautious about the standalone potential of indirect methods. The most successful moral enhancement programme would include the widest possible range of techniques, including indirect methods, but also potentially, pharmacological, neural, and genetic interventions.

6.2. Objections and Replies: Identifying Mediating Mechanisms

On our view, names of moral traits are understood to be useful shorthand labels for underlying moral dispositions realised by a complex interplay of cognitive and affective mechanisms. Indirect moral enhancement targets psychological and biological mechanisms that underlie those dispositions. This creates an epistemic problem for our view: it is not clear how to determine when and why particular mechanisms should be targeted.

The glib answer to this objection is: it depends. The serious answer is twofold. First, there are ways to categorise groups that can help determine the appropriate target for an indirect intervention. Second, there are also ways to account for relevant individual differences in moral cognition.

Some group selection criteria were already mentioned in this chapter. For example, children below a certain age will be ideal targets for interventions that focus on affect and processing of affective stimuli. Adults may be better suited for strategies that rely on conscious self-reflection. People with particular personalities are potential target groups.[74] Particularly useful here may be research into personality traits that compose the so-called Dark Triad, which underlies much of what is typically considered immoral

---

[74]   Avshalom Caspi, et al., 'Personality Differences Predict Health-Risk Behaviours in Young Adulthood', *Journal of Personality and Social Psychology* **73**:5 (1997), 1052–1063.

348

behaviour.[75] Individuals can also be distinguished from one another based on aspects of their personality, elements of their cognition, or even features of their learning styles.[76]

This goes some way toward answering the objection, but nowhere near fully. Group differences and individual differences in context of moral cognition, with an eye to biological and cognitive mechanisms, are not typically studied in the context of possible interventions. However, this information is not impossible to get in principle and nothing stands in the way of this research being carried out.

7. Conclusion

Advocates of moral bioenhancement via invasive means aim to find the best interventions, but as they are well aware, existing methods are on their own inadequate and offer only directions for future research. There is, for example, some evidence of selective serotonin reuptake inhibitors (SSRIs) reducing aggressive attitudes and increasing willingness to co-operate for common goals, and also evidence of oxytocin increasing in-group trust and competence in ToM tasks.[77] Such methods carry with them side-effects, which differ across individuals. SSRIs, for example, may weaken episodic and long-term memory performance and oxytocin may lower the performance of the amygdala and lower fear-responses to potentially dangerous stimuli.

It could turn out that indirect methods, such as those we discuss here, are more effective than moral bioenhancement. But if we are serious about moral enhancement, we should consider all possible interventions, even the direct ones. Whatever problems moral bioenhancement may have, it may become more effective if supplemented with indirect methods. Looking forward, the idea that we can morally improve and enhance individuals by indirect

---

[75] Bianca Bertl, et al., 'More or Less than the Sum of its Parts? Mapping the Dark Triad of Personality Onto a Single Dark Core', *Personality and Individual Differences* **114** (2017), 140–144.

[76] David F. Bjorklund and Kayla B. Causey, *Children's Thinking: Cognitive Development and Individual Differences* (Thousand Oaks, CA: SAGE Publications, 2017).

[77] R. De Jongh, I. Bolt, M. Schermer, and B. Olivier, 'Botox for the Brain: Enhancement of Cognition, Mood and Pro-Social Behaviour and Blunting of Unwanted Memories', *Neuroscience & Biobehavioral Reviews* **32**:4 (2008), 760–776.

methods with an eye to biological factors leaves much room for exploration and experimentation.

Institute of Philosophy, Cognitive Science, Jagiellonian University
michal.klincewicz@uj.edu.pl
Ethics and Philosophy, Technical University, Eindhoven
l.e.frank@tue.nl
Institute of Philosophy, Cognitive Science, Jagiellonian University
sokolska.marta@gmail.com