Oxford Studies in Ancient Philosophy, Volume 57

Author(s) Caston, Victor

Imprint Oxford University Press, 2020

ISBN 9780191885709, 9780198850847, 0198850840

Permalink https://books.scholarsportal.info/uri/ebooks/ebooks6/oso6/2020-02-24/1/oso-9780198850847-Caston

Pages 273 to

Downloaded from Scholars Portal Books on 2026-11-05
Téléchargé de Scholars Portal Books sur 2026-11-05

BECOMING BAD: ARISTOTLE ON VICE AND MORAL HABITUATION

RACHEL BARNEY

ARISTOTLE has remarkably little to say about moral badness or vice or *kakia*.[1] What he does say is both fragmentary and puzzling: Aristotle specifies various features of badness in different contexts, but never explains how the pieces fit together.[2] Nor do we ever get a psychologically vivid picture of the *kakos*, to place alongside his portraits of the akratic, the *megalopsuchos*, and the *phronimos*. As a result, students of Aristotle have traditionally ignored the subject as marginal.[3] Yet badness is not only an ethically urgent topic in its

© Rachel Barney 2019

For criticisms, suggestions, and unanswerable questions on this paper I would like to thank colleagues and audiences at the University of Chicago, Colorado, University of Oklahoma, Oxford, São Paulo (USP), Toronto, and UCLA; and, especially, Hugh Benson, Sarah Broadie, Tom Hurka, Sean Kelsey, Mitzi Lee, Jozef Müller, Karen Nielsen, Martha Nussbaum, Bryan Reece, and Marco Zingano. I am also endebted to Victor Caston for his careful and perceptive editorial scrutiny. I also appreciate the opportunity to acknowledge here my very deep debt (and not only on questions of Aristotelian ethics, or even ancient philosophy in general) to John Cooper, from whose guidance I have acquired only good habits.

[1] I opt for 'badness' to translate *kakia*, sometimes specifying 'moral'; as the most generic term of strong deprecation in English, it is the closest counterpart to Aristotle's term. But I will also use 'vice' indifferently: it has the merit of neatly matching 'virtue', and helpfully has a plural. I will also speak of the *kakos* as the 'fully' or 'truly' bad person, 'bad in the strict sense', etc., reflecting Aristotle's contrast between *kakia* and mere *akrasia*.

[2] I will draw on both versions of Aristotle's *Ethics*, giving references for the common books using Kenny's notation (that is, *AE* A–C = *EE* 4–6 = *NE* 5–7) in A. Kenny, *The Aristotelian Ethics*, 2nd ed. (Oxford, 2016). I have not tried to identify any differences in view between the *Nicomachean Ethics* and the *Eudemian Ethics* on this topic; the question seems worth investigating. Greek texts are taken from the most recent OCTs. Translations of the *Nicomachean Ethics* are from Ross, with revisions; of the *Eudemian Ethics*, from Kenny (likewise); of other texts, from the Revised Oxford Translations.

[3] General studies with interesting reflections on Aristotelian badness include N. Sherman, *The Fabric of Character: Aristotle's Theory of Virtue* [*Fabric*] (Oxford, 1989), 108–17; S. Broadie, *Ethics with Aristotle* (Oxford, 1991), 99–103, 160–78; H. Curzer, *Aristotle and the Virtues* [*Virtues*] (Oxford, 2012), Chapter 17 (and see index for the particular vices).

Oxford Studies in Ancient Philosophy, Volume 57 First edition. Victor Caston (ed.)
This chapter © Rachel Barney 2019. First published 2019 by Oxford University Press

own right; it's hard to see how we could hope to understand Aristotelian virtue apart from it, given that knowledge of contraries belongs to the same science (*Top.* 1. 14, 105^{b}5–6, *Metaph.* B. 2, 996^{a}20). In recent years, interpreters have given Aristotelian badness some overdue attention:[4] but the result has been a range of dramatically different portraits of the Aristotelian bad person.[5] This might reinforce the suspicion that Aristotle simply does not give us enough to go on.[6]

This paper tries to show that, on the contrary, Aristotle has a clear and rich understanding of what moral badness is. I will proceed by working through four of the basic theses which structure his view. First, vice is symmetrical with virtue in a number of central respects (though asymmetrical in others), as the contrary *hexis* of the non-rational soul. I'll discuss this *Symmetry* principle in Section 1. A symmetry of especial importance is that vice, like virtue, is acquired through repeated action (*Habituation*, Section 2). Third, what distinguishes the bad person from the mere akratic is that his reason endorses his actions (*Endorsement*, Section 3).

[4] These include (in alphabetical order) T. P. S. Angier, 'Aristotle', in id. (ed.), *The History of Evil in Antiquity: 2000 BCE–450 CE* (London, 2019), 145–62; J. Annas, 'Virtue, Skill and Vice' ['Virtue'], *Etica & Politica*, 17 (2015), 94–106; T. Brickhouse, 'Does Aristotle Have a Consistent Account of Vice?' ['Account'], *Review of Metaphysics*, 57 (2003), 3–23; A. Fermani, '*To kakon pollachōs legetai:* The Plurivocity of the Notion of Evil in Aristotelian Ethics', in C. Baracchi (ed.), *The Bloomsbury Companion to Aristotle* (London and New York, 2014), 241–59; T. Irwin, 'Vice and Reason' ['Reason'], *Journal of Ethics*, 5 (2001), 73–97; P. Kontos, 'Non-virtuous Intellectual States in Aristotle's Ethics' ['Non-Virtuous'], *Oxford Studies in Ancient Philosophy*, 47 (2014), 205–43 and 'Radical Evil in Aristotle's Ethics and Politics' ['Radical'], in id. (ed.), *Evil in Aristotle* [*Evil*] (Cambridge, 2018), 75–97; J. J. Mulhern, '*Kakia* in Aristotle', *Mnemosyne*, suppl. 307 (2008), 233–54; J. Müller, 'Aristotle on Vice' ['Vice'], *British Journal for the History of Philosophy*, 23 (2015), 459–77, cf. J. R. Elliott, 'Reply to Müller: Aristotle on Vicious Choice' ['Reply'], *British Journal for the History of Philosophy*, 24 (2016), 1193–203; K. Nielsen, 'Vice in the *Nicomachean Ethics*' ['Vice'], *Phronesis*, 62 (2017), 1–25; C. D. C. Reeve, 'Good and Bad in Aristotle' ['Good'], in Kontos, *Evil*, 17–31; and D. Roochnik, 'Aristotle's Account of the Vicious: A Forgivable Inconsistency' ['Inconsistency'], *History of Philosophy Quarterly*, 24 (2007), 207–20.

[5] For a particularly stark contrast, see Nielsen, 'Vice' vs. Müller, 'Vice'.

[6] My interest here will be in the picture of vice we get from Aristotle's two *Ethics*. The *Politics* has much more to offer interpreters interested in human defectiveness: see e.g. B. Yack, *The Problems of a Political Animal: Community, Justice, and Conflict in Aristotelian Political Thought* (Berkeley, 1993), 98–108, and R. Kraut, 'The Political *kakon*: The Lowest Forms of Constitutions', in Kontos, *Evil*, 170–88, and cf. Reeve, 'Good', 23. But whether the *Politics* is consistent with the *Ethics* (or even internally) on the topic is a question I avoid here.

Fourth, this endorsement involves a kind of corrupt activity on the part of reason (*Corruption*, Section 4). I will try to show that these principles suffice to give us an account of vice with some major theoretical advantages. Aristotle's account is rightly open to the many diverse ways there are of being bad—so much so that a number of recent interpretations go wrong, in my view, by mistaking a particular kind of bad person for the whole genus. Yet it also makes a plausible and elegantly simple proposal as to what all truly deplorable people have in common.

In working out Aristotle's view, it may help to keep in mind some of its rivals. We have a rich cultural gallery of competing candidates for the titles *bad*, *vicious*, *evil*, *worst*. There is the pursuer of disvalue as such, like Hannibal Lecter or Milton's Satan; the wanton or brutish slave to low desires; the Dostoevskeian outlaw, committer of some unforgiveable crime; and the amoral egoist or sociopath who greets all moral considerations with a shrug. There is also a more complex figure we might call the *misguided enkratic*:[7] the superficially rational and self-controlled agent in thrall to some mistaken principle, embodied by Hollywood in the honourable Nazi officer and some of the more philosophical James Bond villains. We will see that Aristotle rejects some of these types as psychologically impossible, while others turn out to be something less than bad strictly speaking. His own understanding of the vicious person is, I will try to show, quite different from any of them.[8]

The account presented here will differ from other recent discussions in a number of ways. One is that I will have little to say about an admittedly important fifth thesis, Aristotle's claim that every vice is a form of excess or deficiency. The doctrine of the mean presents complicated problems of its own; and it seems to be primarily designed to diagnose the emotions (*πάθη*) characteristic of good and bad agents, whereas my focus will be on deliberation,

[7] Not an Aristotelian enkratic strictly speaking, for the *enkratēs* without qualification acts correctly (cf. *AE* C. 9), but in the broader sense of someone who does as his practical reason bids. Müller, 'Vice', calls this figure the *principled vicious person* (PVP), arguing instead for a reading of Aristotle's *kakos* as the *conflicted vicious person* (CVP). I take Sherman, *Fabric*; Irwin, 'Reason'; and Nielsen, 'Vice', to spell out in different ways a conception of the *kakos* as a misguided enkratic.

[8] See Section 3 with reference to Sherman, *Fabric*; Irwin, 'Reason'; Annas, 'Virtue'; Müller, 'Vice'; and Nielsen, 'Vice'.

action, and the cognitive aspects of vice. Another difference is that I set aside a text which has traditionally been a stumbling block to interpretation. This is the extended discussion at *Nicomachean Ethics* 9. 4 in which Aristotle depicts the wicked doer of terrible deeds as prey to regret, self-loathing, and motivational conflict (1166^b2–29). This passage is, in my view, simply impossible to square with what Aristotle says about the bad person elsewhere, despite the heroic attempts of recent scholars.[9] My inference is that it is not about the bad person at all; and it is a striking feature of the passage that nowhere does the word *kakos* or any cognate occur.[10] Rather the discussion weaves back and forth between two other groups: the *phauloi* (a group evidently including *all* the non-virtuous 'many', and explicitly including akratics, 1166^b8) and the *mochthēroi* (apparently a subset of the *phauloi* who have committed wicked and criminal acts).[11] How either group relates to the *kakoi* properly speaking is an interesting and non-obvious question—one which Aristotle, in avoiding his technical term, is apparently seeking to duck.[12] So I will avoid reliance on *Nicomachean Ethics* 9. 4 here:

[9] Above all Müller, 'Vice'. Cf. Nielsen, 'Vice'; Elliott, 'Reply'; Irwin, 'Reason'; Brickhouse, 'Account'; and (in a way) Roochnik, 'Inconsistency'.

[10] The parallel discussion of *EE* 7. 6, 1240^b11–18 also avoids *kakos*, in favour of *mochthēros* and *ponēros*.

[11] Nothing in Aristotle's moral psychology entails that *only* the vicious person can do something seriously wrong; the vignette of the *mochthēroi* here might be intended precisely to depict the *non*-vicious doers of wicked deeds. Moreover, since vice is 'sectoral' (see Section 1), an agent might be thoroughly vicious in one respect while retaining decent impulses and judgements in another. This opens up huge scope for conflict and regret—not experienced by the bad person *qua* bad, however, but from the standpoint of the virtue which he might otherwise attain. An almost-generous person will reproach himself if he cannot help others because of self-indulgence, and a mostly friendly person will feel bitter regret if his bad temper spoils his friendships. Thus Aristotle could be right to say both that the vicious person *as such* has no regrets and that a more complex mixed type, the *mochthēros*, will do so.

[12] Admittedly these terms (particularly *kakia* and *mochthēria*) often *are* used interchangeably; but Aristotle sometimes uses them for quite distinct purposes. As *Nicomachean Ethics* 9. 4 itself brings out, *phaulos* is much weaker and broader than the other two. It is often used, as at 1166^b3, to deprecate the ordinary non-virtuous masses (cf. *NE* 2. 3, 1104^b21; 4. 3, 1123^b35; 4. 9, 1128^b25; *EE* 7. 2, 1238^a33), just as *ta phaula* are any base actions (*NE* 2. 3, 1104^b10; 2. 6, 1107^a13; *AE* A. 11, 1138^a28; C. 1, 1145^b 10–14). So what Aristotle says about the *phauloi* cannot be expected to fit all types of non-virtuous people equally well; if it is a particularly poor fit for the *kakoi* as such, we should infer that they are not here in view. (This might signal that they are a comparatively small minority among the *phauloi*: so far as I can see, Aristotle never indicates whether he thinks of vice strictly speaking as common or as rare.) *Mochthēros*

this will allow us to give full weight to Aristotle's insight that the bad person *as such* is unrepentant.

1. Symmetry (and its limits)

For Aristotle, badness is the disposition contrary to virtue (ἕξις ἐναντία) of the non-rational soul, and like it an acquired determination of a distinctively human capacity (δύναμις).[13] Aristotle emphasises that virtue and vice are neither mere capacities nor episodic experiences like emotions; nor are they simply a matter of having good or bad desires. Rather they are permanent structural features of the soul, which determine how it is disposed vis-à-vis the episodic *pathē*.[14] Vice like virtue is a *hexis prohairetikē*, a state concerned with *prohairesis* or choice (*EE* 3. 1, 1228^{a}23–5; 3. 7, 1234^{a}23–5): it is expressed in an agent's conception of the good and so in their rational wish (βούλησις), as worked out in deliberation (*NE* 3. 4–5; *AE* B. 2, 1139^{a}22–3).

So virtue and vice are symmetrical as the best and worst ethical conditions: the fully bad person consistently does the wrong thing and experiences the wrong emotional reactions on the basis of his character (*NE* 2. 4–5; 5. 10; *EE* 2. 2). Vice also mirrors virtue in being a condition of psychic harmony. In the good person, Aristotle says, reason and desire are harmonious, or speak with one voice (ὁμοφωνεῖ): the non-rational part of the soul 'listens to' (κατήκοον), 'is persuaded by' (πείθεται), and 'obeys' (πειθαρχικόν) the correct *logos* (*NE* 1. 13, 1102^{b}26–1103^{a}3).[15] In the bad person too, non-rational

and cognates are more powerful terms, often associated with criminal acts (*AE* A. 2, 1130^{b}24; *EE* 2. 3, 1221^{b}21; cf. *Rhet*. 1. 13, 1374^{a}11), etc. But this association with particular acts means that, despite its strongly deprecatory flavour, in the right context *mochthēros* can be *broader* than *kakos*: so, for instance, 'incontinence seems to be a wickedness [*mochthēria*]' (ἡ δ' ἀκρασία μοχθηρία δοκεῖ εἶναι, *EE* 2. 7, 1223^{a}36–7; cf. *Rhet*. 1. 10, 1368^{b}13–14). All this should be sufficient to warn us against simply assuming that the *kakoi* must be in view in *NE* 9. 4.

[13] *NE* 2. 1; 2. 5; 2. 6 *passim*; 2. 8, 1108^{b}11–19; *AE* A.1–2, 1129^{a}14–26; *EE* 3. 1, 1228^{a}23–5; 3. 2, 1231^{a}36–7. On the metaphysics of *hexeis*, see *Physics* 7. 3; *Cat*. 10; *Metaph*. *Δ*. 20.

[14] On the puzzling question whether Aristotle thinks vice is curable, cf. Curzer, *Virtues*, 367–73; Kontos, 'Non-Virtuous', 234–9; and G. Di Muzio, 'Aristotle on Improving One's Character', *Phronesis*, 45 (2000), 205–19.

[15] That is why Aristotle assures us that this part of the soul must itself have some specifically receptive kind of rationality (*NE* 1. 13, 1103^{a}1–3).

desire and reason are in agreement, as Aristotle stresses in differentiating badness from *akrasia*.[16]

Thus vice is not merely a negative condition for Aristotle, as it is for the Stoics, but a *positive* state.[17] Vice is also symmetrical with virtue in being a *generic* or 'sectoral' concept. There are many vices: in general, one of excess and another of deficiency in relation to each one of the virtues (*NE* 2. 6–9). To be *kakos* is thus to be vicious in some *specific* way (cf. *AE* A. 11, 1138^{a}14–18): we need to hear *kakos* as a placeholder term, standing in for a predication of a specific vice such as 'cowardly', 'stingy', or 'unjust'. This is an important respect in which 'evil' fails to match *kakos* semantically. Aristotle's technical term picks out the person who has a moral disposition which reliably causes wrong action within a particular sphere—not, or not necessarily, the wrongdoer as such, the criminal, or the person who is depraved overall.

At this point, we can begin to see where the symmetries between virtue and vice break down. Whereas the person who has one ethical virtue must have them all, and so can be described as 'virtuous' *simpliciter*, there is no comparable 'unity of the vices'. To be bad *perfectly* and without qualification we would presumably have to collect the whole set; but Aristotle never discusses this ideal type. This may be in part because he has no sense of humour, but also because such a state is not fully realizable: at any rate, vices of excess and deficiency in respect of the same *pathos* will tend to exclude each other.[18] In fact, there is no reason to suppose that the vices cluster even in a general sort of way: one could easily be enkratic in respect of fears and dangers, but bad-tempered or unjust.

A closely related asymmetry is that while the intellectual virtue of *phronēsis*, practical wisdom, is mutually entailing in relation to the ethical virtues, there's no obvious cognitive counterpart in the

[16] One wants here to say that instead of reason coming to rule irrational desire, irrational desire comes to rule reason; and Aristotle does say that 'passion rules' (*κρατεῖ τὸ πάθος*) the bad person in his crucial discussion at *AE* C.8, 1151^{a}22–4. On the whole, though, we get less *Republic*-style talk of 'rule' by one part of the soul over another than we might expect. This is perhaps because it would suggest a stronger commitment to the division of the soul than Aristotle wants (note his striking tentativeness at *NE* 1. 13, 1102^{a}27–^{b}2).

[17] According to the *Categories*, a privation of a *hexis* is *not* its contrary (10, 12^{b}26–13^{a}36).

[18] Only 'tend to' because rashness and cowardice, for instance, can actually be combined (*NE* 3. 7, 1115b29–33).

case of vice. Aristotle does contrast *phronēsis* with another intellectual state, cleverness [δεινότης], which *is* available to the vicious person (*AE* B. 12–13). Cleverness is a kind of unmoralized dexterity in means-ends reasoning: it is what *would* be practical wisdom if it were synched to the ethical virtues. In the bad person it becomes *panourgia*, cunning (*AE* B. 12, 1144^{a}26–8). But this does not make cleverness as such a marker of badness, or even a necessary ingredient.[19]

A third asymmetry lies in the descriptions under which good and bad agents choose their actions. It is part of being a good person to choose the right action 'for the sake of the fine', as Aristotle puts it, or 'for its own sake'. But no parallel requirements hold in the vicious case. Aristotle's bad person does not pursue the bad under that description, let alone choose the cowardly act *qua* cowardly or the stingy act *qua* stingy. On the contrary, like anybody else, he experiences *boulēsis*, the rational species of desire, for what appears good to him (*NE* 3. 4). (Thus Aristotle flatly excludes as psychologically impossible one of the traditional gallery of evil types: the 'Satanic' agent who pursues the bad as such.)[20] The truly vicious person, then, is one whose pursuit of the good in a particular 'sector' *fails* in some reliable way (e.g. *AE* B. 5, 1140^{b}17–20, B. 12, 1144^{a}34–6; see Section 4).

2. Habituation

The second basic building block of Aristotle's account is a claim about how badness is produced. This is by the same mechanism as virtue: both are acquired through habituation [ἐθισμός]. As Aristotle says in a key passage, the person who becomes cowardly or unjust

[19] Annas, 'Virtue', gives an illuminating account of the ways in which badness is learned as something like a skill (cf. Section 3 below); but she seems to me to overemphasise the connection between badness and cleverness [δεινότης] (101, 103–6). One important type of bad person, the immoralist ventriloquized by Glaucon in *Republic* 2, certainly *thinks* of himself as defined by a kind of cleverness which 'good' people lack; but Aristotle is no more likely than Plato to accept that self-conception.

[20] For argument against this position cf. M. Stocker, 'Desiring the Bad: An Essay in Moral Psychology', *Journal of Philosophy*, 76 (1979), 738–53 and J. D. Velleman, 'The Guise of the Good', *Noûs*, 26 (1992), 3–26.

is like the person who, through repeated action, becomes a bad builder or bad lyre-player:

πράττοντες γὰρ τὰ ἐν τοῖς συναλλάγμασι τοῖς πρὸς τοὺς ἀνθρώπους γινόμεθα οἱ μὲν δίκαιοι οἱ δὲ ἄδικοι, πράττοντες δὲ τὰ ἐν τοῖς δεινοῖς καὶ ἐθιζόμενοι φοβεῖσθαι ἢ θαρρεῖν οἱ μὲν ἀνδρεῖοι οἱ δὲ δειλοί. (*NE* 2. 1, 1103^{b}14–17)

For it is by acting as we do in our dealings with other people that some of us become just, others unjust; and by acting as we do in the face of danger, and by becoming habituated to feeling fear or confidence, that some become courageous and others cowardly.

I will discuss this claim more fully in Section 3, when I turn to the role of reason in the vicious person's psyche.

Interpreters have generally had little to say about how this habituation works in the vicious case. Many accounts of moral development in Aristotle slide without comment from discussions of habituation in general to the virtuous case in particular.[21] Some even erase the bad case by a kind of terminological fiat, saying things like 'Ethical habituation means repeatedly engaging in virtuous activity'.[22] This blind spot is, I think, due to a preoccupation with the mysteries of *virtuous* habituation—with the puzzle of how mere repeated action could produce something as comprehensive and cognitively sophisticated as full virtue. The usual solution has been to present an *enriched* picture of what habituation involves: the person developing the virtues does not merely repeat right actions by rote, but acquires relevant perceptual sensitivities and affective responses, an ever-more-refined and powerful sense of shame, a strengthened commitment to the guidance of reason, and so forth.[23] But for all their philosophical sophistication, such

[21] Interpreters who make this move *might* mean to claim that there is simply no such thing as habituation into badness: the bad person is someone in whom no habituation has taken place. But this would imply that all the non-virtuous are vicious, as they are for the Stoics, which is clearly not Aristotle's view. Habituation is a neutral concept in Aristotle, not one reserved for the normative case.

[22] J. Moss, *Aristotle on the Apparent Good: Perception,* Phantasia*, Thought, and Desire* (Oxford, 2012), 207. Moss' talk of the 'character-shaping pleasures of habituation' (206) sounds neutral between virtue or vice; but by a page later habituation is being characterised in terms of 'one type of passion which attends all virtuous activity' (207).

[23] Cf. e.g. in addition to the general accounts already cited, M. F. Burnyeat, 'Aristotle on Learning to be Good', in A. Rorty (ed.), *Essays on Aristotle's Ethics* (Berkeley, 1980), 69–92; J. McDowell, 'Deliberation and Moral Development in Aristotle's Ethics', in S. Engstrom and J. Whiting (eds.), *Aristotle, Kant, and the*

accounts just leave us with a deeper mystery. For people who are becoming bad will have access to few of these enriched mechanisms—apart perhaps from shame, in the bad person being produced by a bad society—and yet habit is also producing a reliable disposition in them. So we need to start by understanding the mechanics of *brute* habituation, the generic process at work in *every* case where a moral disposition is produced. Without one, how could we hope to know what enrichments to habituation are necessary or possible in the virtuous case?

Brute habituation is explained by Aristotle most clearly in a famous passage of the *Eudemian Ethics*:

ὅτι μὲν τοίνυν ἡ ἠθικὴ ἀρετὴ περὶ ἡδέα καὶ λυπηρά ἐστι, δῆλον· ἐπεὶ δ' {ἐστὶ} τὸ ἦθος, ὥσπερ καὶ τὸ ὄνομα σημαίνει, {ὅτι} ἀπὸ ἔθους ἔχει τὴν ἐπίδοσιν, ἐθίζεται δὲ τὸ ὑπ' ἀγωγῆς μὴ ἐμφύτου τῷ πολλάκις κινεῖσθαι πώς, οὕτως ἥδη {τὸ} ἐνεργητικόν (ὃ ἐν τοῖς ἀψύχοις οὐχ ὁρῶμεν· οὐδὲ γὰρ ἂν μυριάκις ῥίψῃς ἄνω τὸν λίθον, οὐδέποτε ποιήσει τοῦτο μὴ βίᾳ)—διὸ ἔστω <τὸ> ἦθος †τοῦτο†, ψυχῆς κατὰ ἐπιτακτικὸν <λόγον τοῦ ἀλόγου μέν,> δυναμένου δ' ἀκολουθεῖν τῷ λόγῳ ποιότης. (*EE* 2. 2, 1220^{a}38–^{b}7)

That moral [*ἠθική*] virtue, then, is concerned with the pleasant and the painful is clear. Now character [*ἦθος*], as its name indicates, is something that develops from habit [*ἔθος*]; and for something to be habituated is for it to be moved repeatedly in a certain way, by a guidance [*ὑπ' ἀγωγῆς*][24] which is not innate, so that it eventually is active [*ἐνεργητικόν*] in the same way. (In inanimate things we do not see this: even if you throw a stone upwards ten thousand times, it will never go upwards except by force.) So let character be considered to be a quality in accordance with governing reason, belonging to the part of the soul which, though non-rational, is yet able to obey reason.[25]

Stoics: Rethinking Happiness and Duty (Cambridge, 1996), 19–35; R. Hursthouse, 'Moral Habituation: A Review of Troels Engberg-Pedersen, *Aristotle's Theory of Moral Insight*', *Oxford Studies in Ancient Philosophy*, 6 (1988), 201–19; M. Jimenez, 'Aristotle on "Steering the Young by Pleasure and Pain"', *Journal of Speculative Philosophy*, 29 (2015), 137–64; G. Lawrence, 'Acquiring Character: Becoming Grown-up', in M. Pakaluk and G. Pearson (eds.), *Moral Psychology and Human Action in Aristotle* (Oxford, 2011), 233–84.

[24] *ἀγωγή* can mean many things, but 'leading' is at the root of all of them (from *ἄγειν*, to lead; cf. *LSJ* s.v.); so Kenny's translation 'impulse' and Inwood and Woolf's 'pattern of conduct' are both misleading. Aristotle's point, made explicit by the 'not innate' and by the parallel with the stone, is that habituation is the internalization of *external* guidance—a distinctively human and somewhat mysterious process.

[25] The text at 1220^{b}1–3 is messy, though the general sense is clear, and my translation is a bit loose.

The stone (dis)analogy is repeated in *Nicomachean Ethics* 2. 1 (1103[a]20–3) and brings out that habituation is a distinctively human process, on a par with nature and reason in its powers (*Pol.* 7. 12, 1332[a]40). In fact a fully formed habit becomes a 'second nature', almost as hard to dislodge as the first (cf. *AE* C. 10, 1152[a]30–4; C. 14).[26] The contrast case of the stone indicates that the initial actions that lay down a habit are given impetus and guidance from outside. So it seems that to be habituated into anything is simply a matter of being guided to repeat the relevant actions until they become like natural ones—that is, effortless and (thus) reliable. Since habituation occurs first and foremost during youth, as part of being educated into one's place in the community, this guidance presumably comes from the modelling, exhortations, and rewards and punishments offered to young people by parents, teachers, role models, and society as a whole. For a type of action to become habitual, then, is for less and less of this kind of external push to be needed, until it comes to be performed reliably with full autonomy. 'Habituation' is thus a success term for Aristotle, but a morally neutral one.

But *why* does this process work? In his ethical works Aristotle seems to take it as just obvious that habit habituates—that is, that the mere repetition of some behaviour, however strongly guided from outside, tends to induce increasingly autonomous repetition in future.[27] (This obviousness is itself part of why moral character is something we can be held responsible for: everyone knows, or ought to know, that their actions are habit-forming, *NE* 3. 5, 1114[a]6 ff.) But this is rather puzzling. After all, repeated experience *can* result in apathy or even aversion, rather than a desire for more of the same. (Violin and piano lessons seem to be the canonical examples here.) If for instance a child is guided to engage in athletic efforts—praised and given prizes for racing and swimming, and shamed for avoiding them—why should the adult continue to run or swim

[26] As Klaus Corcilius notes, Aristotle sometimes even speaks of habit as creating nature: cf. *AE* C. 14, 1154[a]33; *Mem.* 2, 452[a]27–30. K. Corcilius, 'Aristotle's Definition of Non-Rational Pleasure and Pain and Desire', in J. Miller (ed.), *Aristotle's* Nicomachean Ethics: *A Critical Guide* (Cambridge, 2011), 117–43.

[27] As Broadie notes, 'it is remarkable that he has almost nothing to say about *how* or *why* by acting in a certain way we acquire the corresponding moral disposition. That skills, too, are acquired only through practice makes it no less remarkable' (*Ethics with Aristotle*, 104).

once those motivating supports are removed? Many don't. And yet we too use the terms 'habit' and 'habituation' to pick out a kind of repetition which *does* become self-reinforcing: we seem to agree with Aristotle that there is such a thing. To see how it works, we need to turn to the *Rhetoric*, and a passage which has not received the attention it deserves in the ethical context:

ὑποκείσθω δὴ ἡμῖν εἶναι τὴν ἡδονὴν κίνησίν τινα τῆς ψυχῆς καὶ κατάστασιν ἁθρόαν καὶ αἰσθητὴν εἰς τὴν ὑπάρχουσαν φύσιν, λύπην δὲ τοὐναντίον. εἰ δ' ἐστὶν ἡδονὴ τὸ τοιοῦτον, δῆλον ὅτι καὶ ἡδύ ἐστι τὸ ποιητικὸν τῆς εἰρημένης διαθέσεως, τὸ δὲ φθαρτικὸν ἢ τῆς ἐναντίας καταστάσεως ποιητικὸν λυπηρόν. ἀνάγκη οὖν ἡδὺ εἶναι τό τε εἰς τὸ κατὰ φύσιν ἰέναι ὡς ἐπὶ τὸ πολύ, καὶ μάλιστα ὅταν ἀπειληφότα ᾖ τὴν ἑαυτῶν φύσιν τὰ κατ' αὐτὴν γιγνόμενα, καὶ τὰ ἔθη (καὶ γὰρ τὸ εἰθισμένον ὥσπερ πεφυκὸς ἤδη γίγνεται· ὅμοιον γάρ τι τὸ ἔθος τῇ φύσει· ἐγγὺς γὰρ καὶ τὸ πολλάκις τῷ ἀεί, ἔστιν δ' ἡ μὲν φύσις τοῦ ἀεί, τὸ δὲ ἔθος τοῦ πολλάκις), καὶ τὸ μὴ βίαιον (παρὰ φύσιν γὰρ ἡ βία, διὸ τὸ ἀναγκαῖον λυπηρόν, καὶ ὀρθῶς εἴρηται 'πᾶν γὰρ ἀναγκαῖον πρᾶγμ' ἀνιαρὸν ἔφυ'), τὰς δ' ἐπιμελείας καὶ τὰς σπουδὰς καὶ τὰς συντονίας λυπηράς· ἀναγκαῖα γὰρ καὶ βίαια ταῦτα, ἐὰν μὴ ἐθισθῶσιν· οὕτω δὲ τὸ ἔθος ποιεῖ ἡδύ. (*Rhetoric* 1. 11, 1369^{b}33–1370^{a}13)

We may lay it down that pleasure is a movement, a movement by which the soul as a whole is consciously brought into its normal state of being; and that pain is the opposite. If this is what pleasure is, it is clear that the pleasant is what tends to produce this condition, while that which tends to destroy it, or to cause the soul to be brought into the opposite state, is painful. It must therefore be pleasant for the most part to move towards a natural state of being, particularly when a natural process has achieved the complete recovery of that natural state. Also habits [are necessarily pleasant]. (For what is habitual becomes just like what's natural. Habit is something like nature: what happens often is akin to what happens always, natural events happening always, habitual events often.) Also what is not forced [is necessarily pleasant]. (For force is unnatural, and that is why what is compulsory is painful, and it has been rightly said, 'All that is done under compulsion is bitterness unto the soul.') But acts of concentration, effort, and strain [are necessarily] painful: they all involve compulsion and force, unless they are habitual, in which case it is habit that makes them pleasant. (trans. Roberts, revised)[28]

[28] This conception of pleasure as a motion is unlike Aristotle's analyses in *Aristotelian Ethics* C and *Nicomachean Ethics* 10, and one might dismiss it as an immature view not to be relied on here. I infer instead that the *Rhetoric* is, here as on many points, an endoxic work, designed to be insofar as possible independent of proprietary Aristotelian views. That makes it an excellent place to look if our question is: what might Aristotle in the *Ethics* be taking as *just obvious* about how habituation works?

This is how a habitual action comes to be 'second nature': it becomes easy and pleasant to us in just the same way that physically natural movements are. A habitual action is thus one which restores us to our default setting, or what has come to feel like it. If like a Spartan or Celtic child I am forced to wear only minimal clothes in winter, I will find it much easier than others to do so as an adult because it will feel natural to me—I will take a certain comfort in the sheer familiarity of the experience. This is not to say that the habitual action is always pleasant *overall* (the cold will still be intrinsically unpleasant, like playing the violin), or always the *most* attractive option available. But it does put a thumb on the deliberative scale. And these *pleasures of habit as such* will be hard to recognise, we might even say non-transparent. For an action incentivized in this way will not *feel* like one motivated by a desire for pleasure; we will just be doing what comes naturally. Often we will be aware of the affective pull only negatively: as Aristotle says at an important moment of *Nicomachean Ethics* 10. 9, what has become customary 'won't be painful' ($1179^{b}35$).[29]

This almost invisible hedonic incentive explains Aristotle's odd and remarkable (but surely true) claim that a person may sometimes pursue pleasure self-indulgently even when they have little or no real appetite for it (*AE* C. 4, $1148^{a}17$–22; C. 7, $1150^{a}25$–31).[30] Why should anyone pursue an excessive pleasure—finish the bag of potato chips, say—absent any real craving? Such actions are perplexing but (*ahem*) not wholly unfamiliar, and we sometimes describe them as mechanical or automatic. This is a giveaway that they are habit-driven; and the 'force of habit' is mysterious precisely because it seems to float free of any gain or pleasure inherent in the action performed. Aristotle can explain that since a habitual action is one which has become effortless, and thus feels to us like a return to what is right and natural, it is pleasant or at least pain-reducing in itself *regardless* of the nature of the habitual act. Thus

[29] By the same token, habit must intensify certain pains: the self-indulgent person is tormented if he must pass up some pleasure. It is also by habituation that the corrupted person comes to find the *wrong* things pleasant (cf. *NE* 10. 5, $1176^{a}19$–26)—and also good (*NE* 3. 4–5), presumably because his conception of the good is largely shaped by his experiences of pleasure and pain (*NE* 3. 4, $1113^{a}31$–b2).

[30] Stewart even presents this as essential to the vicious-akratic distinction (J. A. Stewart, *Notes on the Nicomachean Ethics of Aristotle*, 2 vols. (Oxford, 1892), ii. 199).

as Aristotle says, habit can lead us to take pleasure in things which are not inherently pleasant in the first place—as, he remarkably observes, in the case of those who have been sexually abused since childhood (*AE* C. 5, 1148^{b}28–31). Through brute habitation, then, something not pleasant by nature may easily become pleasant *for me*. This enables us to see that the self-indulgent person is *not* best understood as the person who is subject to particularly powerful appetites and gives in to them. Rather she's the person who defers to her appetites in a routinized way *regardless* of their strength, because anything else feels difficult, unnatural, and exhausting.[31]

These *pleasures of habit as such* need to be carefully distinguished from the other pleasures found together with them. In particular, they are not to be confused with the associative pleasures used in the initial 'guidance', as when the child is rewarded for athletic efforts with a cookie. Such associative pleasures are very useful to the educator in incentivizing and reinforcing the 'guided' behaviour. But if the habituated person continues to run and swim as an adult, after the cookies have gone away, it is not necessarily because of these associations alone. In fact, motivation by past associative pleasures is likely to be at odds with the agent's ongoing experience in the present (*where's my cookie?*), and so a source of cognitive dissonance—hardly the best basis for a stable disposition. The pleasures of habit are different. If habituation has worked, the adult will genuinely enjoy running: she will find that it now comes naturally, feel a sense that something is wrong when a run is skipped, and so on. (To enjoy running because of an insight into its value as an

[31] Cf. the remarkable depiction of vicious habituation in C. S. Lewis' *Screwtape Letters*. The demon Screwtape is here giving advice to a colleague about his human 'patient': 'as habit renders the pleasures of vanity and excitement and flippancy at once less pleasant and harder to forgo (for that is what habit fortunately does to a pleasure) you will find that anything or nothing is sufficient to attract his wandering attention. You no longer need a good book, which he really likes, to keep him from his prayers or his work or his sleep; a column of advertisements in yesterday's paper will do. You can make him waste his time not only in conversation he enjoys with people whom he likes, but in conversations with those he cares nothing about on subjects that bore him. You can make him do nothing at all for long periods. You can keep him up late at night, not roistering, but staring at a dead fire in a cold room. All the healthy and outgoing activities which we want him to avoid can be inhibited and *nothing* given in return, so that at last he may say, as one of my own patients said on his arrival down here, "I now see that I spent most of my life in doing *neither* what I ought *nor* what I liked"' (C. S. Lewis, *The Screwtape Letters* (London, 1942; repr. Glasgow, 1977), 63–4 in the 1977 repr.).

admirable exercise of our natural capacities would be another kind of pleasure again, resulting from habituation of the 'enriched' kind.) What is distinctive about the pleasures of habit as such, then, is that they are both reliable and independent of the content of the habit in question. That means that they are equally powerful in reinforcing habits good and bad.

This, then, is Aristotle's picture of brute habituation—the basic structure common to both virtue and vice, to which various enrichments may be added in the former case. Repeated episodes of action come to feel increasingly natural and thus pleasant to the agent, until she is reliably motivated to perform them *without* any external guidance. (Of course continued 'guidance' can't hurt, and a wisely designed society will continue to incentivize good behaviour through social mechanisms of shame and fear (cf. *NE* 10. 9).)[32] If we're tempted to doubt the power that Aristotle attributes to this mechanism, we should bear in mind that habituation will typically be reinforced by yet another source of pleasure. Aristotle holds that human beings are also naturally *imitative* creatures:

> *τό τε γὰρ μιμεῖσθαι σύμφυτον τοῖς ἀνθρώποις ἐκ παίδων ἐστὶ καὶ τούτῳ διαφέρουσι τῶν ἄλλων ζῴων ὅτι μιμητικώτατόν ἐστι καὶ τὰς μαθήσεις ποιεῖται διὰ μιμήσεως τὰς πρώτας, καὶ τὸ χαίρειν τοῖς μιμήμασι πάντας.* (*Poetics* 4, 1448^{b}5–9)
>
> Imitation is ingrained in human beings from childhood, and they differ from the other animals in this—that they are most imitative, and learn at first by imitating—and all delight in imitations.

Like his explanation of the power of habit in the *Rhetoric*, this seems intended as a statement of the obvious (no argument for it is given), and I take it to be operative in his *Ethics* in the same endoxic way. For a lot of ethical habituation is bound to be through imitation: the younger child is encouraged to learn by pretending to be Achilles or Hector, or by copying his older siblings who in turn mimic their parents. That means that certain habitual actions—the ones modelled for us and reinforced by pleasures and pains—are triply easy, with the pleasures of habit, of imitation, and any socially imposed associative pleasures all pointing in the same direction; and every society's system of moral education exploits

[32] As Martha Nussbaum has pointed out to me, bad inculcated habituation can also help itself to *some* of the mechanisms of good, 'enriched' habituation, in particular, shame. The person raised in a morally deformed society will be ashamed to have the *right* non-rational ethical responses—the Huck Finn problem.

this potential. I will refer to this deliberate moulding of habits in this way as *inculcated habituation*. It is, I think, what Aristotle has in mind as the standard case by which we become good or bad.[33]

So for Aristotle to say that moral badness comes from habituation is to say that bad people become so through the self-reinforcing repetition of wrong actions, particularly as inculcated early on by their society. And this picture of bad habituation seems designed to cover two quite different kinds of case. Some people will become vicious as the result of an inculcated habituation which *misfires*: the disposition their educators attempt to instil is a genuinely virtuous one, but something goes wrong. Perhaps the actions the young person performs are not quite the ones intended; or perhaps their reactions of pleasure and pain are a bit off, due to some imbalance in their natural endowment; and so a vicious *hexis* results instead of the corresponding virtue. We might think of the case in which a parent tries to teach a child to be thrifty—shaming expenditure, praising and rewarding him for saving his allowance—but pushes a bit too hard, relative to that child's natural propensities. As a result, the child comes to take excessive pleasure in saving, and ends up a miser instead.

In the other kind of bad habituation, educators might *intend* to produce what Aristotle would diagnose as a vice rather than a virtue. For nothing about Aristotle's conception of vice limits it to the person who is recognised as morally defective in his own society.[34] The vicious person may just as easily—indeed much *more* easily—be the product of successful inculcated habituation by a vicious society. Many societies, after all, get the virtues wrong: a pleasure-driven culture will use all its powers of inculcation to produce self-indulgent people, a greed-obsessed society to produce unjust ones. Consider the case of courage, for instance, which belongs only to the honour-loving citizen-soldier who fears the battlefield to the extent

[33] The imitative drive is of course not restricted to childhood. In adulthood, it will continue simply as what we would now call *conformism*: the drive by which it comes naturally to me to do what I see other people doing. Habits formed by conformism too will always have a double power over us, *qua* habit and *qua* imitation—in addition to any pleasures that might flow from that habit in particular, or from social reinforcement.

[34] Like bad habituation more generally, this possibility is generally ignored by interpreters of the *Ethics*, so that I am not sure how far this is a controversial claim. Müller for one remarks that 'vice is generally shunned and shamed by the society in which one lives' ('Vice', 471), as if only the 'outlier' should count.

appropriate and faces that fear for the right reasons (*NE* 3. 6–8; *EE* 3. 1). Societies in which soldiers are an oppressed class, those which delegate fighting to mercenaries, those committed to pacifism, and, at the other extreme, the maniacal Celts (*NE* 3. 7, 1115^{b}24–8)—all will be incapable of producing genuinely courageous people, even if they inculcate something they *think* of as courage. Admittedly no society inculcates vice *de dicto*: as Aristotle says,

> *οἱ γὰρ νομοθέται τοὺς πολίτας ἐθίζοντες ποιοῦσιν ἀγαθούς, καὶ τὸ μὲν βούλημα παντὸς νομοθέτου τοῦτ' ἐστίν, ὅσοι δὲ μὴ εὖ αὐτὸ ποιοῦσιν ἁμαρτάνουσιν, καὶ διαφέρει τούτῳ πολιτεία πολιτείας ἀγαθὴ φαύλης.* (*NE* 2. 1, 1103^{b}3–6)
>
> Legislators make the citizens good by forming habits in them, and this is the wish of every legislator, and those who do not effect it well miss their mark, and it is in this that a good constitution differs from a bad one.

But, again, there are two ways in which the legislators may miss the mark. They may fail to contrive the institutions needed to instil the virtues correctly conceived; or they may misidentify the virtues, and inculcate vicious states instead.

So it is an important Aristotelian insight, not to be glossed over, that virtuous habituation is only one of two paths which moral development may take. Vicious habituation is an equally viable alternative, and both rely on the same morally neutral central mechanism, brute habituation. The vicious form of this in turn subdivides in two, depending on whether habituation has misfired or worked as intended: the vicious agent may be either an *outlier* within her own society or, so to speak, a *trainee*. To ignore the latter possibility, as scholars generally have done, is to underestimate drastically the generality of Aristotle's account, and to suppress some interesting resources for Aristotelian social critique. Moreover, some aspects of his conception of vice will be easier to make sense of if we keep the latter case clearly in view.

So far, I have been discussing habituation as essentially a process of reinforcement, aimed at instilling a pattern of pleasures and pains which will induce autonomous action. But all this is only one side of the story. For Aristotle makes clear that habituation—including habituation into vice—is also a cognitive process: a person is *learning to do* something in becoming fully bad, just like someone who learns to be a bad builder or lyre-player. To see what the bad person learns and how, we need to turn to Aristotle's third central thesis, *Endorsement*.

3. Endorsement

The third building block of Aristotle's account is the thesis that the bad person's reason *endorses* his wrong actions.[35] This is the dividing line between the bad person (in particular, the self-indulgent or intemperate person or *ἀκόλαστος*) and the person who exhibits mere *akrasia*, as discussed in his most extensive treatment of badness, *Aristotelian Ethics* C. Aristotle's central claim here is that the self-indulgent person is essentially different from even the worst *akratēs* because he acts in accordance with choice (*ἄγεται προαιρούμενος*), which flows from rational wish and deliberation (C. 3, $1146^{b}22$–3; cf. *κατὰ τὴν προαίρεσιν*, C. 8, $1151^{a}6$–7). Thus he is persuaded or convinced of what he does, whereas the akratic acts against choice (*παρὰ προαίρεσιν*), and is not so persuaded (*AE* C. 2, $1146^{a}31$–$^{b}2$; C. 8, $1151^{a}6$–26). (The language of persuasion here recalls Aristotle's initial account of intra-soul relations in *NE* 1. 13, $1102^{b}26$–$1103^{a}3$.) In contrast to the tormentedly self-aware akratic, the vicious person thinks he *ought* to act as he does on each occasion (*AE* C. 3, $1146^{b}22$–3; cf. *δεῖν* at C. 8, $1151^{a}23$), and so is unaware of his vice ('it escapes him', *λανθάνει*, C. 8, $1150^{b}36$). Standing by his choice, he does not regret his actions (C. 8, $1150^{b}29$–30). Thus while an akratic is like a city that does not obey its laws, a self-indulgent person is like a city that abides by bad ones (C. 10, $1152^{a}20$–4). The moral difference between those who are truly bad and those who fail to meet that standard thus does not lie in the object of their desires (both the akratic and the self-indulgent seek pleasure, and in the same things); rather it rests in *how* they desire them, and with it how they think of themselves. And though Aristotle confusingly prefers to use *akrasia* restrictively (*AE* C. 3–4), for the less defective counterpart of self-indulgence, it seems clear that there must be counterparts of *enkrateia* and *akrasia* corresponding to every virtue, so that his contrast here is one which applies to every ethical 'sector'.[36]

[35] Here and in Section 4, I attempt to avoid any commitments on the controversial topic of how Aristotelian deliberation works in general and in the virtuous case. Insofar as possible, the account of bad deliberation given here is meant to be independent of any particular account of Aristotle's division of labour between the rational and non-rational soul. Section 4 does presuppose that reason can, and should, play a critical role in relation to the ends from which deliberation begins: this seems to me the clear implication of the passages I discuss there.

[36] That the contrast must hold for all the vices is clear from his association of vice with a false conception of the good (*NE* 3. 4–5, cf. esp. 3. 4, $1113^{a}31$–3) and so with

Aristotle's contrast between the mere akratic and the truly vicious person seems to me to express an important ethical insight. There is indeed a difference between the person who gives in to occasional temptations to pleasure and the one who insists on satisfying every whim. And there is the same kind of difference between the person who does the cowardly thing out of weakness of will—the situational coward, who panics and flees the battlefield in shame—and the smug coward-by-policy, who prides himself on getting others do his fighting for him (through repeated draft deferments, say). Aristotle seems to me right to hold that only the latter has the vice of cowardice in full. In the case of injustice too, there is a difference between the person who on some occasion gives way to a temptation to benefit himself unjustly, and the one committed to *pleonexia* as a way of life—consistent and shameless in evading his taxes, cheating his business partners, and ripping off his customers. In the case of anger, one person gives way now and then to a bout of rage; another is an unrepentant serial bully. The difference between the two is not, or not just, that the latter is very frequently angry—that might also be true of a merely weak-willed person with an unfortunate temperament. It is that the truly bad-tempered person, like the deeply self-indulgent, cowardly, and unjust one, always manages to feel satisfied about his behaviour. And Aristotle's position is that only this kind of person is well and truly vicious or evil.

The truly bad person is self-satisfied because his behaviour has been habituated: he is doing what he always does, and it feels natural to him. But Aristotle also insists—again, I think, rightly—that (since we are rational beings) this reliable self-satisfaction must depend on rational endorsement. Bad people think that they *ought* (*δεῖν*) to act as they do; thus, in the case of self-indulgence, they 'think that the present pleasure ought always to be pursued' (*νομίζων ἀεὶ δεῖν τὸ παρὸν ἡδὺ διώκειν*, *AE* C. 3, 1146^{b}22–3). But we should be careful, I think, not to take this as a requirement that the bad person have a *theory*.[37] The *aei* or 'always' here might seem to

prohairesis (e.g. *AE* A. 8, 1135^{b}17–25) quite generally. Likewise the 'corruption passages' I discuss in Section 4 include some texts in which self-indulgence is particularly in view ([C], [D], and [E]), and others where it is not ([A], [B], and [F]).

[37] Such a case might be uppermost in Aristotle's mind, since he might well be thinking of Callicles in the *Gorgias* (491 E 8–492 A 3). But there is no reason for him to assume that the bad person is always or even usually as reflective as that.

suggest a hedonist along the lines of Oscar Wilde's Lord Henry Wotton—a version of the type I earlier called the misguided enkratic, a rational agent in the grip of a false conception of the good.[38] But the *aei* can, as ever, be serial as easily as continuous (to be translated, that is, 'in each case' rather than 'always'), and might apply to the thinking rather than being part of the content thought; and the 'ought' (*δεῖν*) need not do more than indicate a vague sense that one's action is in order.[39] So Aristotle's description applies equally well to Homer Simpson—that is, to the wholly unreflective agent who, on every occasion that an opportunity for pleasure presents itself, finds some rationale for signing up. The contrast at hand with the akratic requires only that the vicious person does *not* think, as the akratic does, that he ought *not* to do as he does. Vice requires an absence of psychological conflict, sufficient to ensure reliability: not reflection or theoretical principles.[40]

Still, even this weak kind of self-endorsement expresses some kind of engagement by the vicious person's reason. To see what it consists in, let's start by taking a closer look at Aristotle's claim that, like the person who becomes a bad builder or bad lyre-player, the person habituated into badness is *learning* something:

> *ἔτι ἐκ τῶν αὐτῶν καὶ διὰ τῶν αὐτῶν καὶ γίνεται πᾶσα ἀρετὴ καὶ φθείρεται, ὁμοίως δὲ καὶ τέχνη· ἐκ γὰρ τοῦ κιθαρίζειν καὶ οἱ ἀγαθοὶ καὶ κακοὶ γίνονται κιθαρισταί. ἀνάλογον δὲ καὶ οἰκοδόμοι καὶ οἱ λοιποὶ πάντες· ἐκ μὲν γὰρ τοῦ εὖ οἰκοδομεῖν ἀγαθοὶ οἰκοδόμοι ἔσονται, ἐκ δὲ τοῦ κακῶς κακοί. εἰ γὰρ μὴ οὕτως εἶχεν, οὐδὲν ἂν ἔδει τοῦ διδάξοντος, ἀλλὰ πάντες ἂν ἐγίνοντο ἀγαθοὶ ἢ κακοί. οὕτω δὴ καὶ ἐπὶ τῶν ἀρετῶν ἔχει· πράττοντες γὰρ τὰ ἐν τοῖς συναλλάγμασι τοῖς πρὸς τοὺς ἀνθρώπους γινόμεθα οἱ μὲν δίκαιοι οἱ δὲ ἄδικοι, πράττοντες δὲ τὰ ἐν τοῖς δεινοῖς καὶ ἐθιζόμενοι φοβεῖσθαι ἢ θαρρεῖν οἱ μὲν ἀνδρεῖοι οἱ δὲ δειλοί.* (*NE* 2. 1, 1103^{b}6–17)

Further, it is from and through the same causes that every virtue is both produced and destroyed, and likewise every art: for it is from playing the

[38] On the *aei*, see Irwin, 'Reason', 87–9; Müller, 'Vice', 465–6; and Nielsen, 'Vice', 21–2.

[39] For this vague and weakly rational kind of approbation, we might compare the Stoic view that even the thief considers his action *kathēkon*, 'appropriate'.

[40] The bad person does also think that the good one is criticisable in some way: 'the brave man is called rash by the coward, cowardly by the rash man, and correspondingly in the other cases' (*καὶ καλοῦσι τὸν ἀνδρεῖον ὁ μὲν δειλὸς θρασὺν ὁ δὲ θρασὺς δειλόν, καὶ ἐπὶ τῶν ἄλλων ἀνάλογον*, *NE* 2. 8, 1108^{b}24–6). I take it that this is meant to remind us of the kind of distortion of moral language Thucydides famously reports from the Corcyran civil war (3. 82), and the point is again just that such distortions of language express a kind of cognitive commitment absent from the merely akratic.

lyre that people become good and bad lyre-players. And it is analogous in the case of builders and all the rest: from building well, people will be good builders, and from building badly, bad ones. For if it were not so, there would be no need of teaching, but everyone would be born good or bad. And likewise in the case of the virtues. For it is by acting as we do in our dealings with other people that some of us become just, others unjust; and by acting as we do in the face of danger, and by becoming habituated to feeling fear or confidence, that some become courageous and others cowardly.

Commentators have not paid much attention to this analogy,[41] but it is an odd and striking thing to say (about the crafts, let alone the virtues and vices). In what sense is a bad builder or lyre-player something one *learns to be*, through repeatedly doing these things badly?[42] The most obvious kind of 'bad lyre-player', we might assume, is the person who never manages to learn to play at all. But this is clearly not the kind of case Aristotle has in mind. He does not say 'By building some people become good builders but others never become any good, or even learn enough to count as builders at all'.[43] He is contrasting good habituation not with a failure to be habituated, but rather with habituation into the wrong thing. This is confirmed by his official discussion of *technē* in *AE* B. 4. Here the state opposed to *technē* is *atechnia*; and, as Pavlos Kontos has shown, the person characterized by it is not the one wholly uninstructed in an art, but the one who has learned it in some wrong way, involving a false principle or account (*logos*, 1140^{a}20–3).[44]

Aristotle's analogy implies that vicious habituation can produce very diverse results: for bad lyre-players and bad builders come in many varieties. One kind of bad lyre-player would presumably be like the guitar players in *Spinal Tap*: an adept practitioner who successfully imitates bad models, and enjoys achieving effects he ought to be ashamed of. (Or think of a highly competent painter of kitsch, or an Oscar-winning ham actor.) The bad builder is more easily imagined as the one who knows enough about building to

[41] Except for a pair of insightful recent discussions in Annas, 'Virtue' and Kontos, 'Non-Virtuous'.

[42] A paradoxical idea in Greek as well, but Aristotle might be thinking of the famous line from the pseudo-Homeric comic *Margites*: 'he knew many things, but all of them badly' (*πολλὰ μὲν ἠπίστατο ἔργα, κακῶς δὲ ἠπίστατο πάντα*, [Plato], *Alc. 2*, 147^{b}3–4).

[43] Cf. Kontos, 'Non-Virtuous', 209.

[44] Kontos ('Non-Virtuous', 215) also notes that ignorance (*agnoia*) is likewise for Aristotle a positive condition of error, not a mere cognitive blank.

serve selfish ends which conflict with those of the art, to which he's indifferent—the real estate tycoon who gets rich by passing off bad taste and shoddy construction as luxury. But we would surely also count as a bad builder one who, though well-meaning, persistently makes a certain kind of major technical mistake, misjudging the methods and principles of his art. And no doubt there are other possibilities as well. The only requirement is that the bad craftsperson is guided by a mistaken *logos* at *some* level of specificity or other, so as to reliably adopt the wrong end in action: one way or another, they are consistently wrong about what kind of action the situation requires.

This openness has the happy result that in the moral case too, Aristotle's account is compatible with a wide range of profiles in badness. One kind of vicious agent might not even start with the right question, namely, 'What is the right (or "fine", *kalon*) thing to do?'. Instead he asks an instrumentalizing, strategic question: 'How can I act here so as to gain some advantage for myself?'[45] But other agents can go deeply and permanently wrong despite, at some very general level, aspiring to do the right thing. Of these, some will fail to detect the virtue called for in a given kind of situation; others misidentify the content of that virtue, thinking for instance that courage consists in aggression and violence or that temperance consists in extreme austerity. The upshot is, as it seems to me, that most recent interpretations of Aristotle's theory of badness have identified a perfectly viable subspecies of it, and gone wrong only in excluding errors of other kinds or at other levels of generality. Given the framework of Aristotle's basic theses about badness which we have considered, a vicious person *may* be an instrumentalist about his own actions (Sherman, *Fabric*); an adept

[45] Cf. Sherman's account of the Aristotelian *kakos* as committed to an instrumentalized view of his own actions (*Fabric*, 108–17); Irwin, 'Reason' (from whom I take the term 'strategic'); and Annas, 'Virtue', lean in the same direction. Such readings seem unduly narrow since, as Nielsen ('Vice', 12) points out, Aristotle is explicit that 'each state of character [*ἕξιν*] has its own ideas of the noble and pleasant [*καλὰ καὶ ἡδέα*]; and perhaps the good man differs from the other most by seeing the truth in each class of things' (*καθ' ἑκάστην γὰρ ἕξιν ἴδιά ἐστι καλὰ καὶ ἡδέα, καὶ διαφέρει πλεῖστον ἴσως ὁ σπουδαῖος τῷ τἀληθὲς ἐν ἑκάστοις ὁρᾶν*, *NE* 3. 4, 1113^{a}31–3). This implies that a certain kind of bad person might go wrong despite being motivated by the *kalon*—as she understands it—in just the same way as the virtuous person, by being deeply and culpably wrong about which actions in fact are *kala*. However, the bad person does not need to be motivated in this way; again, Aristotle's account allows for enormous diversity in vice.

immoralist or egoist (Annas, 'Virtue'); a misguided enkratic attached to false principles about the good (Nielsen, 'Vice'); a reflectively committed slave to inclination (Irwin, 'Reason'); or a changeable, quasi-wanton pursuer of present pleasure (Müller, 'Vice'). For all these are kinds of agents whose reason endorses their reliable adoption of wrong ends within some given sphere. Thus Aristotle embraces the idea that vice is in an important sense *negative* or indeterminate. It belongs naturally to the class of the unlimited, for it represents a failure to hit a determinate target; and it is possible to fail in many ways, but to be good only in one—a thought which Aristotle inherits from the Pythagoreans and bequeaths to the Stoics (*NE* 2. 6, 1106^{b}28–35).

But this openness to the diversity of vice just seems to reinforce the mystery at hand. In all these varieties of bad person, the same kind of *rational* defect has been acquired in the course of their being wrongly habituated; can anything at all be said, in a general way, about what it is? The answer, perhaps surprisingly, is yes.

4. Rational corruption

Aristotle's fourth principle, *Corruption*, is that the bad person's reason endorses his actions because it has been 'corrupted' (*diaphtheiresthai*) and 'twisted around' or 'distorted' (*diastrephesthai*)—that is, damaged, perverted, or spoiled. This is explained in several texts in very consistent terms; but these are so brief and (it must be admitted) opaque as to be on the face of it unhelpful; this is perhaps why interpreters of Aristotelian badness have had very little to say about them. We need to begin by looking at these *corruption passages* (as I will call them) in conjunction, and with some care. Two are to be found in Aristotle's discussions of practical wisdom, or *phronēsis*:

[A] *ἔνθεν καὶ τὴν σωφροσύνην τούτῳ προσαγορεύομεν τῷ ὀνόματι, ὡς σῴζουσαν τὴν φρόνησιν. σῴζει δὲ τὴν τοιαύτην ὑπόληψιν. οὐ γὰρ ἅπασαν ὑπόληψιν διαφθείρει οὐδὲ διαστρέφει τὸ ἡδὺ καὶ λυπηρόν, οἷον ὅτι τὸ τρίγωνον δύο ὀρθὰς ἔχει ἢ οὐκ ἔχει, ἀλλὰ τὰς περὶ τὸ πρακτόν. αἱ μὲν γὰρ ἀρχαὶ τῶν πρακτῶν τὸ οὗ ἕνεκα τὰ πρακτά· τῷ δὲ διεφθαρμένῳ δι' ἡδονὴν ἢ λύπην εὐθὺς οὐ φαίνεται ἀρχή, οὐδὲ δεῖν τούτου ἕνεκεν οὐδὲ διὰ τοῦθ' αἱρεῖσθαι πάντα καὶ πράττειν· ἔστι γὰρ ἡ κακία φθαρτικὴ ἀρχῆς.* (*AE* B. 5, 1140^{b}11–20)

And this is why we call moderation by this name, as preserving practical wisdom. Now what it preserves is a judgement of this kind. For it is not

every judgement that the pleasant and painful corrupt [διαφθείρει] and twist around [διαστρέφει], for instance that the triangle has or has not angles equal to two right angles, but those concerning what is to be done. For the first principles [ἀρχαί] of things to be done consist in that for the sake of which they are done; but for someone who has been corrupted [διεφθαρμένῳ] by pleasure or pain, straightaway the first principle [ἀρχή] fails to appear, nor [does he see] that he ought to choose and do everything for the sake of this and because of this. For vice is corrupting [φθαρτική] of the first principle [ἀρχῆς].

[B] ἡ δ' ἕξις τῷ ὄμματι τούτῳ γίνεται τῆς ψυχῆς οὐκ ἄνευ ἀρετῆς, ὡς εἴρηταί τε καὶ ἔστι δῆλον· οἱ γὰρ συλλογισμοὶ τῶν πρακτῶν ἀρχὴν ἔχοντές εἰσιν, ἐπειδὴ τοιόνδε τὸ τέλος καὶ τὸ ἄριστον, ὁτιδήποτε ὄν (ἔστω γὰρ λόγου χάριν τὸ τυχόν)· τοῦτο δ' εἰ μὴ τῷ ἀγαθῷ, οὐ φαίνεται· διαστρέφει γὰρ ἡ μοχθηρία καὶ διαψεύδεσθαι ποιεῖ περὶ τὰς πρακτικὰς ἀρχάς. (*AE* B. 12, 1144ª29–36)

And this eye of the soul acquires its formed state not without the aid of virtue, as has been said and as is clear; for the syllogisms which deal with things to be done involve a first principle [ἀρχήν], viz, 'since the end and what is best is of *this* sort', whatever it may be (let it be as it may, for the sake of argument); and this does not appear except to the good person; for wickedness distorts [διαστρέφει] and produces errors about the first principles [ἀρχάς] of action.

Aristotle's account of the subhuman brute is another context in which it is important for him to distinguish the cognition characteristic of the *kakos*:

[C] ἔλαττον δὲ θηριότης κακίας, φοβερώτερον δέ· οὐ γὰρ διέφθαρται τὸ βέλτιον, ὥσπερ ἐν τῷ ἀνθρώπῳ, ἀλλ' οὐκ ἔχει . . . ἀσινεστέρα γὰρ ἡ φαυλότης ἀεὶ ἡ τοῦ μὴ ἔχοντος ἀρχήν, ὁ δὲ νοῦς ἀρχή. (*AE* C. 6, 1150ª1–5)

Brutishness is not as big a deal as badness, though it is more frightening, because it is not that the better [part of the soul] has been corrupted [διέφθαρται], as in a human being, but that it is not present . . . for the badness of what does not have a first principle [ἀρχήν] is always less destructive, and reason [νοῦς] is a first principle [ἀρχή].

And so is the crucial passage already discussed in Section 3 above, in which he outlines the key contrast with the akratic:

[D] ἐπεὶ δ' ὃ μὲν τοιοῦτος οἷος μὴ διὰ τὸ πεπεῖσθαι διώκειν τὰς καθ' ὑπερβολὴν καὶ παρὰ τὸν ὀρθὸν λόγον σωματικὰς ἡδονάς, ὃ δὲ πέπεισται διὰ τὸ τοιοῦτος εἶναι οἷος διώκειν αὐτάς, ἐκεῖνος μὲν οὖν εὐμετάπειστος, οὗτος δὲ οὔ· ἡ γὰρ ἀρετὴ καὶ μοχθηρία τὴν ἀρχὴν ἣ μὲν φθείρει ἣ δὲ σῴζει, ἐν δὲ ταῖς πράξεσι τὸ οὗ ἕνεκα ἀρχή, ὥσπερ ἐν τοῖς μαθηματικοῖς αἱ ὑποθέσεις· οὔτε δὴ ἐκεῖ ὁ λόγος διδασκαλικὸς τῶν

ἀρχῶν οὔτε ἐνταῦθα, ἀλλ' ἀρετὴ ἢ φυσικὴ ἢ ἐθιστὴ τοῦ ὀρθοδοξεῖν περὶ τὴν ἀρχήν. σώφρων μὲν οὖν ὁ τοιοῦτος, ἀκόλαστος δ' ὁ ἐναντίος. ἔστι δέ τις διὰ πάθος ἐκστατικὸς παρὰ τὸν ὀρθὸν λόγον, ὃν ὥστε μὲν μὴ πράττειν κατὰ τὸν ὀρθὸν λόγον κρατεῖ τὸ πάθος, ὥστε δ' εἶναι τοιοῦτον οἷον πεπεῖσθαι διώκειν ἀνέδην δεῖν τὰς τοιαύτας ἡδονὰς οὐ κρατεῖ· οὗτός ἐστιν ὁ ἀκρατής, βελτίων <ὢν> τοῦ ἀκολάστου, οὐδὲ φαῦλος ἁπλῶς· σῴζεται γὰρ τὸ βέλτιστον, ἡ ἀρχή. (*AE* C. 8, 1151^{a}11–26)

Since one is the kind of person who pursues bodily pleasures which are excessive and contrary to reason but not out of conviction, while the other *is* convinced, because he is the kind of person who pursues them, the former easily changes his mind while the latter does not. For virtue and vice respectively preserve and corrupt [*φθείρει*] the first principle [*τὴν ἀρχήν*], and in actions the final cause is the first principle [*ἀρχή*], as the hypotheses are in mathematics; neither in that case is it argument that teaches the first principles [*τῶν ἀρχῶν*], nor is it so here—virtue either natural or produced by habituation is what teaches right opinion about the first principle [*τὴν ἀρχήν*]. Such a one as this, then, is temperate; the contrary type is the self-indulgent. But there is a kind of person who is carried away as a result of passion and contrary to right reason—one whom passion rules so that he does not act according to right reason, but does *not* rule to the extent that he is the sort to be convinced that he *ought* to pursue such pleasures without reserve. This is the akratic, who is better than the self-indulgent person, and not bad without qualification; for the best thing in him, the first principle [*ἀρχή*], is preserved.

Finally, a passage from the *Eudemian Ethics* also touches on the same points, albeit in a somewhat confusing dialectical context:

[E] *ἄτοπον γὰρ εἰ τὴν μὲν ἐν τῷ λογιστικῷ ἀρετὴν μοχθηρία ποτὲ ἐγγενομένη ἐν τῷ ἀλόγῳ στρέψει καὶ ποιήσει ἀγνοεῖν, ἡ δ' ἀρετὴ ἐν τῷ ἀλόγῳ ἀγνοίας ἐνούσης, οὐ στρέψει ταύτην καὶ ποιήσει φρονίμως κρίνειν . . .* (*EE* 8. 1, 1246^{b}19–23, MSS text without supplements)

For it would be bizarre if, on the one hand, when badness has arisen in the non-rational [soul] it will overturn [*στρέψει*] the virtue in the rational [soul] and make it ignorant, but on the other hand virtue in the non-rational [part], when ignorance is present [in the rational part], will not overturn this and cause it to judge wisely . . .[46]

[46] The context is a discussion of the paradoxical possibility that folly plus *akrasia* might amount to virtue (cf. *AE* C. 2, 1146^{a}27–31), and the text and sense are disputed. But for our purposes, all that matters is the opening supposition here, which is clear enough and not called into question by the discussion which follows: when badness arises in the irrational soul it *does* overturn any excellence in the rational soul, rendering it ignorant.

Aristotle's varied but overlapping formulations in these passages revolve around a pair of related theses; confusingly, he uses the crucial term *archē* to express both. I have translated *archē* as 'first principle' throughout, but these passages exploit the deep ambiguity of this ancient term: it must be understood as 'starting-point' in some occurrences and 'governing-principle' in others, often in very close conjunction. Sometimes, as in [C] and the last lines of both [A] and of [D], *archē* evidently refers to an enduring part or feature of the soul itself—namely, the natural ruling principle of our actions, reason or *nous* and its virtue [C], which is 'the best thing' (*τὸ βέλτιστον*) in us [D]. In the bad person, this has been twisted around, overturned, corrupted ([A], end; [C]; [D], end). But elsewhere, Aristotle clearly uses *archē* to mean a kind of judgement [A]: specifically, the starting-point of deliberation, comparable to a hypothesis in mathematics [D] or the major premiss of a syllogism [B]. And this is identified with the end or that 'for the sake of which' (*οὗ ἕνεκα*) action is to be done, which only the good person gets right ([A], [B], [D]). What is particularly confusing is that Aristotle also speaks of *this* kind of *archē* as 'corrupted' and 'twisted around' (*διαφθείρει*, *διαστρέφει*, [A]; *φθείρει*, [D]; *διαστρέφει*, [B]; and note that [B] fudges the question of what exactly gets corrupted), though it is not immediately clear what this could mean. Presumably the idea is that the true and correct first principle required for deliberation is *supplanted* by a faulty, defective one; so the point is the same one Aristotle makes by saying that to the bad person, the (correct) starting-point 'fails to appear' (*οὐ φαίνεται*, [A] at $1140^{b}18$; [B] at $1144^{a}34$).

In sum, the corruption passages make two claims, one about dispositional corruption and one about episodic failure:

(1) *Dispositional corruption:* In the bad person, (practical) reason (*nous*, *logos*) as the ruling principle (*archē*) of the soul has been damaged and perverted.

(2) *Episodic failure:* In particular deliberative contexts, the bad person does not grasp the correct first-principle (*archē*) of action; it 'does not appear' to him.

Not only does Aristotle use *archē* to express both these claims: he scarcely seems to see the two as distinct. (This is particularly evident in the way that both [A] and [C] pivot abruptly from *archē* as occurrent deliberative starting-point to *nous* as the *archē* of the

soul.) The explanation must be that he sees a tight explanatory connection here, running from (1) to (2): it is *because* the ruling-principle is corrupted that the starting-point does not appear. Aristotle as good as says so in [A], which concludes: 'For [γάρ] vice is corrupting of the *archē*.' This is only explanatory, as the γάρ purports, if it does not simply repeat the preceding claims about episodic failure; it must be pointing to the corruption of the soul's governing principle as the *explanation* of this. In short, the dispositional corruption of the bad person is the *cause* of her episodic deliberative ignorance. And this corruption is in turn, as Aristotle tells us in all of [A], [B], and [D], the result of bad habituation—of repeatedly experiencing pleasure and pain in the wrong ways.

Aristotle does not go into detail about what this corrupted state involves. But one possibility we can exclude, I think, is that it consists in the complete *destruction* of practical reason. *Diaphtheiresthai* certainly *can* mean 'destroyed altogether', as of a person killed.[47] Yet it cannot mean that in this context. This is where passage [C] is important: Aristotle is quite explicit that the person from whom practical reason is absent altogether is the brute (ὁ θηριώδης), and he emphasises that this person is different in kind from the vicious agent (*AE* C. 6, 1150^{a}1–5; cf. *AE* C. 1, 1145^{a}25–7). It would also, I think, be a mistake to suppose that 'corruption' amounts merely to a kind of perceptual inadequacy or *weakness* on the part of practical reason. Being damaged and twisted around is something different—and worse—than being weak; it is also necessarily an acquired state, while a weak perceptual capacity could be innate. (We might even think it is the natural state of the immature, unhabituated agent.) And again, there is the contrast with the brute, already noted: if a merely defective perceptual capacity were the

[47] *Diaphtheirein* and cognates cover the range from complete destruction (e.g. of a person killed, Antiphon 2. 2. 5; Aristotle *AE* A. 11, 1138^{a}13; *Pol*. 3. 16, 1287^{a}39–40; 7. 16, 1335^{a}18) to any kind of spoiling (*Pol*. 3. 15, 1286^{a}38; of natural objects which are damaged, *Rhet*. 3. 18, 1419^{b}4–5; *Meteor*. 1. 14, 352^{b}29), including the kinds of ethical and political defectiveness for which our 'corruption' is an excellent match (e.g. of officials impaired by too quick a rise to power at *Pol*. 5. 8, 1308^{b}14; of decision-making distorted by anger at *Pol*. 3. 15, 1286^{a}34–5—n.b. the apparently interchangeable use of *diastrephein* at 3. 16, 1287^{a}31—and of those who take pleasure in what is not really pleasant at *NE* 10. 5, 1176^{a}23–4). This use of *diaphtheirein* for intellectual, moral, and political ruination is standard and longstanding: cf. its use in regard to judgement at Aeschylus *Ag*. 932, and for 'the corruption of the young' (25 A, 30 B) in Plato's *Apology* (see also LSJ s.v.).

problem, the brutish person would be the limit case of the bad person. Aristotle's equally frequent language of being twisted, overturned, distorted, perverted (*diastrephein*, cf. LSJ s.v.) implies that a corrupted reason acts against its own nature—that is, in some sense *irrationally*—in the bad person. And this can only be because, like the enslaved reason of Plato's bad person in *Republic* 8–9, it is suborned to ends other than its own.

So it is not Aristotle's view that a destroyed or enfeebled reason stands by while the non-rational soul does its thing: reason is an active collaborator in vicious action. Though the division of labour in Aristotelian deliberation is enormously controversial, two things must both be true here: that the non-rational soul *primarily* determines the ends of the bad person, and that the corruption of reason *enables* it to do so by some characteristic perverse activity. Otherwise it could not be true to say both, as Aristotle does, that vice is a state of the non-rational soul and that rational corruption is a necessary condition of it. What makes this corruption mysterious, or at least hard to specify, is that it is curiously *negative* in its effect. What it produces is a failure: the correct end simply 'does not appear' (*οὐ φαίνεται*). To understand what is really distinctive about the truly bad person, then, we will need to identify the kind of exercise of reason which has that effect. Unfortunately, Aristotle never gives us an example of corrupted reason in action, or even any useful metaphors or hints. And since his general remarks on deliberation are hopelessly problematic and controversial, there is no readymade account for us to extend to the bad case. In any case the corruption passages make clear that the distinctive activity of a corrupted reason must be exercised *prior* to and outside deliberation proper, since it bears on the starting-point from which deliberation begins.

So at the heart of Aristotle's thinking about badness is a kind of silence, or a black box. We know that what distinguishes the bad person is a certain way of thinking which is both rational and irrational; that this is a necessary condition for the reliable adoption of the wrong ends; and that it is produced by bad habituation, and so ultimately by pleasure and pain. But all these amount to specifications that the right picture of corrupt practical reasoning must meet, rather than giving us the picture itself. However, if we ask ourselves what Aristotle *should* say here, what conception of corrupted reasoning his claims imply, a quite determinate answer

stares us in the face. And it is the same as the answer we get if we turn to the fact of the matter, and ask what kind of thinking *does* do all this.

To see how, let's press a bit on the closest Aristotle comes to a direct depiction of the work of corruption: his rather haunting little locution the end 'does not appear'. By putting the result of corrupted reasoning in these terms, he implies, again, that its work is negative. The non-rational soul is evidently sufficient by itself for the positive task of coming up with the wrong ends; what a corrupted reason contributes can only be the suppression of the alternatives. It blocks recognition of the rightness of the right end; it makes moral truths invisible. (This implies that the crucial activity of a *correctly* functioning practical reason is critical: it must offer a kind of scrutiny of the ends proposed by the non-rational soul, leading to the replacement of any ends which fail to withstand moral reflection with others which do. But I cannot here explore these interesting implications for the virtuous case.)

At this point, if we can get past the unhelpfully vague and moralistic language of 'corruption', what Aristotle means to pick out should be obvious enough. For nothing could be more familiar. *Of course* there is a kind of reasoning which bad people use to ward off recognition of salient moral truths, an activity in which reason is made to serve ends other than its own, a kind of thinking prior and peripheral to deliberation proper which enables it to reliably start in the wrong place. Corrupted reasoning is just the vicious form of what we now call *motivated* reasoning: the varieties of rationalization, denial, confabulation, and self-deceived excuse-making that enable vicious agents to reach the conclusions they want. Motivated reasoning is far from exclusive to bad people, nor is it restricted to contexts of moral deliberation; but Aristotle seems to me right to hold that the truly bad person has to do a lot of it, and that this is a necessary and characteristic feature of his badness—the explanation for the gap in culpability, noted above, between him and the akratic. Sustained success at rationalization and the like is indeed what makes it possible for someone to be consistently and complacently unjust, or cowardly, or self-indulgent—what sets the truly deplorable apart from their less bad akratic counterparts. For motivated reasoning is what produces and sustains their distinctive self-complacency, by eliminating any relevant cognitive dissonance—the unease, or worse, induced by any nascent awareness that something

is wrong with the way they are accustomed to think and behave. We can see this more clearly if we focus on the trainee, the person in whom vice is the product of successful inculcated habituation. After all, what *is* the rational ability instilled by a society in which young people are trained to do the wrong thing? (Think, for instance, of someone raised to privilege in a slaveholding or apartheid society.) The answer seems obvious: they learn how to block any recognition of that wrongness, so that, just as Aristotle says, they can simply 'fail to notice' their own moral condition. And they learn to do *that* by learning to produce rationalizations which eliminate cognitive dissonance, by rendering the wrongness of their habitual ends invisible to them. It seems to me that Aristotle is right to hold that any kind of fully vicious state is dependent on reasoning of this kind, and that this dependency is both psychological and evaluative: that is, it is true *both* that, as a matter of psychological fact, we cannot be *reliably* bad without the suborning of reason in this way, *and* that the person who falls short in this regard—the person whose reason continues to press him with urgent and inconvenient truths—counts as less fully vicious.[48]

We can now see why it would be a mistake to identify the Aristotelian vicious person with the misguided enkratic. The point is closest to explicit in text [D] of the corruption passages, where the point is that the bad person is said to do what he does confidently and without internal conflicts simply because he has become that 'kind of person' (τοιοῦτος)—the kind of person, that is, who does that kind of thing. Sarah Broadie, noting the significance of this passage (and of the contrast with the akratic at *AE* C. 8, 1151^{a}13–14, 22–4), adds that 'the voluptuary's "being persuaded" has been effected by passions and the habit of going along with them. His intellectual conviction that he acts as he should is hardly

[48] Rationalizations and other culpable forms of motivated reasoning (denial, confabulation, etc.) belong to a grey zone of complex, non-transparent psychological states; they are hard to describe without paradox and apparent contradiction, for the person engaged in rationalizing her behaviour must at some level 'know better', but that recognition must be kept from full consciousness if the rationalization is to do the job. Aristotle is on very general metaphysical grounds hostile to the very idea of contradictory beliefs being held by the same thinker (*Metaph. Γ.* 3–4); and in contrast to Plato, he is reluctant to commit to a real division of the soul into ethically salient parts (*NE* 1. 13). But without some kind of partitioning of the self and its beliefs, motivated reasoning is extremely difficult to explain or even describe. And this is perhaps why Aristotle's explanations of vice stop short exactly where they do.

more than the ideological face of a non-intellectually established practical bent'.[49] Aristotle's analysis in fact suggests a kind of debunking of the person who *presents* as a misguided enkratic. Bad people don't act wrongly because of adopting the wrong theory: they adopt the wrong theory because it defends the comfort they have come to feel in doing the wrong thing. And we can also now see why Aristotle is so confident that brute habituation is sufficient for the acquisition of vice. If we are inclined to think of the bad agent as a misguided enkratic, we will be tempted to assume a kind of 'two-track' model, on which the trainee must be simultaneously guided into bad habits and indoctrinated into a false theory. But Aristotle has in mind something subtly yet importantly different. As the corruption passages insist, what twists reason around is just the very same history of pleasure and pain which instils viciousness in the non-rational soul.

This understanding of rational corruption as a commitment to motivated reasoning can also help to explain Aristotle's hard-line stance on culpability for vice in *Nicomachean Ethics* 3. 1–5. Aristotle grants that the bad person always acts in ignorance (ἄγνοια) of what he ought to do; but, he insists, that ignorance is not exculpating. As he says in affirming the voluntary character of vice, 'every wrongdoer [μοχθηρός] is ignorant [ἀγνοεῖ] of what he ought to do, and what he ought to avoid, and because of this sort of error people become unjust and in general bad [κακοί]' (*NE* 3. 1, 1110^{b}28–30).[50] But the key to making sense of this position is to see that ignorance in the relevant sense is something more than a mere absence of

[49] S. Broadie, '*Nicomachean Ethics* VII.8–9 (1151^{b}22): *Akrasia*, *enkrateia*, and Look-alikes', in C. Natali (ed.), *Aristotle's* Nicomachean Ethics, *Book VII* (Oxford, 2009), 157–72 at 161. Likewise Gauthier and Jolif: '[I]f the intemperate person does the wrong thing with conviction, it is not because of some purely intellectual persuasion, so that if informed of his mistake he would immediately become a moral exemplar. The corruption of his judgment is nothing but the upshot of his corrupted character.' ([S]i l'intempérant est persuadé qu'il doit faire le mal, ce n'est pas d'une persuasion purement intellectuelle, de telle sorte qu'il suffirait de lui dire qu'il se trompe pour qu'il devienne immédiatement un modèle de bonnes oeuvres! La dépravation de son jugement n'est que la conséquence de la dépravation de son caractère, R. A. Gauthier and J.-Y. Jolif, *L'Éthique à Nicomaque*, 2nd ed. (Louvain-La-Neuve, 2002), vol. ii. 2, 598–9 ad *AE* C. 3, 1146^{a}33.)

[50] ἀγνοεῖ μὲν οὖν πᾶς ὁ μοχθηρὸς ἃ δεῖ πράττειν καὶ ὧν ἀφεκτέον, καὶ διὰ τὴν τοιαύτην ἁμαρτίαν ἄδικοι καὶ ὅλως κακοὶ γίνονται. Further to n. 11 above, note what looks like a difference here between *mochthēros*, used for any agent going wrong situationally, and *kakos*, used for the character type eventually produced by such behaviour.

information.[51] It refers not to a condition of mental blankness but to a constructive state—a kind of error or false thinking, as he says in [B]—which is expressive of one's character. In fully bad people, this ignorance seems to be irreversible; at any rate they could not, as presently constituted, act otherwise than they do. But given that this state is one achieved by their past actions, and sustained by their ongoing efforts, it is not hard to see why he deems it culpable.[52]

5. Conclusions

In the end, then, Aristotle's bad person turns out to be a distinctive ethical type, a new portrait to hang in the gallery: the *rationalizer*. He neither chooses the bad lucidly, like the Satanic agent; *nor* acts on sincere intellectual error, like the misguided enkratic; *nor* does wrong in opposition to his reason, like the akratic; *nor* lacks its input altogether, like a madman or a brute. And in eschewing these cruder pictures, Aristotle escapes an ugly philosophical dilemma. For suppose we ask: is the evil person governed by reason or not? If he is, as by being a misguided enkratic, then by Aristotle's lights he is getting something—almost the most important thing—*right*. And Aristotle would owe us much more of an account than he gives of how reason itself can lead us astray (a possibility which he, like Plato, generally prefers

[51] Cf. Kontos, 'Non-Virtuous', 215 and n.b. *Top*. 6. 9, 148^{a}8–9: ignorance is not mere privation of knowledge, but involves positive error.

[52] One might object that there is a problem with claiming, as I do, both that Aristotle has the trainee or conformist bad person in view *and* that he takes motivated reasoning to be a necessary condition of their badness. For the more all-pervasive a vicious ideology is in some society, the less cognitive 'work' is required of the individual: in the limit case, we might think, the perfectly habituated trainee in a wholly vicious society will experience no cognitive dissonance at all. Now if this strikes us as exculpating, it confirms that we *do*, like Aristotle, think of badness as involving corrupt cognitive effort; but it seems to exclude the trainee as a paradigm case. However, I doubt that Aristotle would grant this last point: he seems to be an epistemic optimist (like, in their different ways, Socrates before him and the Stoics after), confident that we *all* have inalienable epistemic access to an in-principle sufficient fund of moral truths. His extreme severity about the culpability of bad agents for their own condition, apparently regardless of context, is presumably a side effect of this optimism. (Aristotle's optimism also raises the question of whether the soul of the vicious person really *can* ever be *fully* harmonious; for a reading of Aristotle on which he, like Kant and the Socrates of the *Gorgias* (482 B), is committed to denying this possibility, see S. Engstrom, 'Virtue and Vice in Aristotle and Kant', in Kontos, *Evil*, 222–39.)

to ignore). If on the other hand the bad agent acts *against* his reason, then his badness seems to be incomplete, and of a venial and forgivable akratic type. If these were the only two possibilities, it is hard to see what true badness could be. The awkward, allusive locutions of the corruption passages represent Aristotle's struggle to articulate a third option, one which threads this needle. The bad person is neither simply rational nor irrational. In him, reason formally governs; but it does so as a puppet regime, always busying itself to find some way to endorse what non-rational vice demands.

So Aristotle's principal substantive claim about vice is, simply, that this sort of collaborationist regime is ethically worse than either of the cruder options. For him, no one can count as truly bad unless they exhibit the intellectual dishonesty, the determined lack of self-awareness, involved in the sustained rationalization of wrong actions. This seems to me an insightful and elegantly simple thesis, accurately capturing what we find appalling about the very worst people. Aristotle's other principal claim on the topic is psychological—perhaps even, in principle, empirical. It is that people acquire this kind of rational corruption, and with it vice in the strict and proper sense, through brute habituation, rather than innocent intellectual error or even indoctrination. The vicious are compromised cognitively, just as they are ethically, by the knock-on effects of repeated bad action itself.

Ironically, given how little he has to say about it, Aristotle's account of habituation in the end turns out to work better for vice than for virtue. As I noted earlier, habituation into virtue is somewhat mysterious, and interpreters have had to build in enrichments about which Aristotle himself says strikingly little. But in the case of vice, it is hardly controversial that brute habituation can do the job: that we become what we do, and that someone in whom bad habits are inculcated by bad people in a bad society will, in fact, tend to become bad. And Aristotle is surely right that this process is not just a matter of doing the wrong thing with increasing autonomy and reliability, but of experiencing it as ever more natural, and becoming ever more resourceful at keeping the recognition of its wrongness at bay.

Seeing this can in turn help to rebalance our sense of where Aristotelian virtue comes from. For the enrichments which have been attributed to virtuous habituation by generations of sophisticated and charitable interpreters have arguably obscured the very point Aristotle is trying to make. Namely: mere rote repeated action is

extremely powerful! Brute habituation does more than you would think! Particularly in *Nicomachean Ethics* 2, Aristotle is answering 'Meno's question' about how virtue is acquired; and his answer is not only diametrically opposed to that of Plato's Socrates, but, I suspect, deliberately one-sided and polemical—even meant to shock.[53] He is arguing for an alarming, pessimistic, anti-intellectualist paradox: that *action itself*, however caused, is the cause of character, and not (or not only) the other way around. In the beginning was the deed.

This picture of badness is very unlike what we find in other ancient philosophical accounts. The worst person in Plato's *Republic*, the tyrannical man, is a paranoid mess; the Stoics argue that *any* bad person must be prey to fear, distress, and inconsistent behaviour. By contrast, Aristotle's conception is positively antidramatic, and remarkably modern-looking: as a kind of learned, self-serving moral obliviousness, Aristotelian vice is more reminiscent of modern concepts like the 'banality of evil', 'affected ignorance', and 'white ignorance' than of anything in his own philosophical tradition.[54] A natural worry would be that this picture is *too* familiar, and that there's something anachronistic about the reading I've offered. But the rationalizer figure often enough holds the stage in ancient tragedy; this convergence is one of the reasons that Sophocles and Euripides can so easily be reset in a contemporary tyranny or dystopia.[55] And if Aristotle's account of vice looks modern, so does its foundation: his stark commitment to the decisive powers of brute habituation for good and for ill. Social science is only now beginning to investigate those powers;[56] Aristotle can help us to a better understanding of how they might generate something as powerful and manifold as human evil.

[53] For 'Meno's question', cf. Plato, *Meno* 70 A; for Plato's Socrates, at least as Aristotle himself interprets him, *EE* 1. 2, 1216^{b}2–10. By 'deliberately one-sided', I have in mind that the postponement of any discussion of *phronēsis* to *AE* B might be a matter of deliberate authorial strategy, precisely in order to initially push the attribution of virtue and vice to habituation as far as it will go. But of course puzzles about the composition of Aristotle's ethical works make any such claims highly speculative.

[54] H. Arendt, *Eichmann in Jerusalem: A Report on the Banality of Evil* (New York, 1963); M. M. Moody-Adams, 'Culture, Responsibility, and Affected Ignorance', *Ethics*, 104 (1994), 291–309; C. W. Mills, 'White Ignorance', in S. Sullivan and N. Tuana (eds.), *Race and Epistemologies of Ignorance* (Albany, 2007), 11–38.

[55] Examples include Odysseus in Sophocles' *Philoctetes*, Creon in his *Antigone*, Jason in Euripides' *Medea*, Pentheus in the *Bacchae*...

[56] For a popular overview with references, cf. C. Duhigg, *The Power of Habit* (New York, 2012).

BIBLIOGRAPHY

Angier, T. P. S., 'Aristotle', in id. (ed.), *The History of Evil in Antiquity: 2000 BCE–450 CE* (London, 2019), 145–62.

Annas, J., 'Virtue, Skill and Vice' ['Virtue'], *Etica & Politica*, 17 (2015), 94–106.

Arendt, H., *Eichmann in Jerusalem: A Report on the Banality of Evil* (New York, 1963).

Brickhouse, T., 'Does Aristotle Have a Consistent Account of Vice?' ['Account'], *Review of Metaphysics*, 57 (2003), 3–23.

Broadie, S., *Ethics with Aristotle* [*Ethics with Aristotle*] (Oxford, 1991).

Broadie, S. '*Nicomachean Ethics* VII.8–9 ($1151^{b}22$): *Akrasia*, *enkrateia*, and Look-alikes', in C. Natali (ed.), *Aristotle's* Nicomachean Ethics*, Book VII* (Oxford, 2009), 157–72.

Burnyeat, M. F., 'Aristotle on Learning to be Good', in A. Rorty (ed.), *Essays on Aristotle's Ethics* (Berkeley, 1980), 69–92.

Corcilius, K., 'Aristotle's Definition of Non-Rational Pleasure and Pain and Desire', in J. Miller (ed.), *Aristotle's* Nicomachean Ethics*: A Critical Guide* (Cambridge, 2011), 117–43.

Curzer, H., *Aristotle and the Virtues* [*Virtues*] (Oxford, 2012).

Di Muzio, G., 'Aristotle on Improving One's Character', *Phronesis*, 45 (2000), 205–19.

Duhigg, C., *The Power of Habit* (New York, 2012).

Elliott, J. R., 'Reply to Müller: Aristotle on Vicious Choice' ['Reply'], *British Journal for the History of Philosophy*, 24 (2016), 1193–203.

Engstrom, S., 'Virtue and Vice in Aristotle and Kant', in Kontos, *Evil*, 222–39.

Fermani, A., '*To kakon pollachōs legetai:* The Plurivocity of the Notion of Evil in Aristotelian Ethics', in C. Baracchi (ed.), *The Bloomsbury Companion to Aristotle* (London and New York, 2014), 241–59.

Gauthier, R. A. and Jolif, J.-Y., *L'Éthique à Nicomaque*, 2nd ed. (Louvain-La-Neuve, 2002).

Hursthouse, R., 'Moral Habituation: A Review of Troels Engberg-Pedersen, *Aristotle's Theory of Moral Insight*', *Oxford Studies in Ancient Philosophy*, 6 (1988), 201–19.

Irwin, T., 'Vice and Reason' ['Reason'], *Journal of Ethics*, 5 (2001), 73–97.

Jimenez, M., 'Aristotle on "Steering the Young by Pleasure and Pain"', *Journal of Speculative Philosophy*, 29 (2015), 137–64.

Kenny, A., *The Aristotelian Ethics*, 2nd ed. (Oxford, 2016).

Kontos, P., 'Non-virtuous Intellectual States in Aristotle's Ethics' ['Non-Virtuous'], *Oxford Studies in Ancient Philosophy*, 47 (2014), 205–43.

Kontos, P., 'Radical Evil in Aristotle's Ethics and Politics' ['Radical'], in id., *Evil*, 75–97.

Kontos, P. (ed.), *Evil in Aristotle* [*Evil*] (Cambridge, 2018).
Kraut, R., 'The Political *kakon*: The Lowest Forms of Constitutions', in Kontos, *Evil*, 170–88.
Lawrence, G., 'Acquiring Character: Becoming Grown-up', in M. Pakaluk and G. Pearson (eds.), *Moral Psychology and Human Action in Aristotle* (Oxford, 2011), 233–84.
Lewis, C. S., *The Screwtape Letters* (London, 1942; repr. Glasgow, 1977).
McDowell, J., 'Deliberation and Moral Development in Aristotle's Ethics', in S. Engstrom and J. Whiting (eds.), *Aristotle, Kant, and the Stoics: Rethinking Happiness and Duty* (Cambridge, 1996), 19–35.
Mills, C. W., 'White Ignorance', in S. Sullivan and N. Tuana (eds.), *Race and Epistemologies of Ignorance* (Albany, 2007), 11–38.
Moody-Adams, M. M., 'Culture, Responsibility, and Affected Ignorance', *Ethics*, 104 (1994), 291–309.
Moss, J., *Aristotle and the Apparent Good: Perception,* Phantasia*, Thought, and Desire* (Oxford, 2012).
Mulhern, J. J., '*Kakia* in Aristotle', *Mnemosyne*, suppl. 307 (2008), 233–54.
Müller, J., 'Aristotle on Vice' ['Vice'], *British Journal for the History of Philosophy*, 23 (2015), 459–77.
Nielsen, K., 'Vice in the *Nicomachean Ethics*' ['Vice'], *Phronesis*, 62 (2017), 1–25.
Reeve, C. D. C., 'Good and Bad in Aristotle' ['Good'], in Kontos, *Evil*, 17–31.
Roochnik, D., 'Aristotle's Account of the Vicious: A Forgivable Inconsistency' ['Inconsistency'], *History of Philosophy Quarterly*, 24 (2007), 207–20.
Sherman, N., *The Fabric of Character: Aristotle's Theory of Virtue* [*Fabric*] (Oxford, 1989).
Stewart, J. A., *Notes on the Nicomachean Ethics of Aristotle*, 2 vols. (Oxford, 1892).
Stocker, M. 'Desiring the Bad: An Essay in Moral Psychology', *Journal of Philosophy*, 76 (1979), 738–53.
Velleman, J. D., 'The Guise of the Good', *Noûs*, 26 (1992), 3–26.
Yack, B., *The Problems of a Political Animal: Community, Justice, and Conflict in Aristotelian Political Thought* (Berkeley, 1993).