

University of Groningen

Plato on pleasure and illusion

van Zoonen, Derek

DOI:
[10.33612/diss.250286363](https://doi.org/10.33612/diss.250286363)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2022

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

van Zoonen, D. (2022). *Plato on pleasure and illusion*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen. <https://doi.org/10.33612/diss.250286363>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

2. Tricked by Pleasure: Hedonic Fallibilism in Plato's Republic

In *Republic* 9, Socrates presents three 'proofs' or 'demonstrations' (ἀποδείξεις) of the superiority of the just life. Together with book 4's discussion of justice as a kind of psychological health, these arguments are designed to undermine Thrasymachus's radical attack on conventional morality along with his underlying ideas about what makes for a life worth living. Because the second and third of these arguments both invoke pleasure and argue for one and the same conclusion,¹⁴⁴ they seem best understood as a larger, two-pronged strand of argumentation which runs from 580c10 to 588a6 and which I will call the Pleasure Argument.¹⁴⁵

Whereas book 9's first argument had attacked ordinary lives on psychological grounds—it basically proposes that a person's degree of psychological harmony and integration decides whether their life counts as worth living or not—the Pleasure Argument changes tack and assails ordinary, non-ideal lives on hedonic rather than psychological grounds. In the earlier bits of this argument, Socrates appeals to the idea that the rationally integrated agent—who has tempered the lower parts of his soul and made himself 'one instead of many' (παντάπασιν ἓνα γεγόμενον ἐκ πολλῶν, 443e1–

¹⁴⁴ Cf. 583a1–3 ('the one of us in whom [the rational part] rules has the most pleasant life (ὁ τούτου βίος ἡδιστός)') and 587b9–10 ('the tyrant will inevitably live most unpleasantly and the [ideal agent] most pleasantly (ἀηδέστατα ... ὁ τύραννος βιώσεται, ὁ δὲ βασιλεὺς ἡδιστα)').

¹⁴⁵ Cf. Nussbaum (1986: 141).

2)¹⁴⁶—counts as a ‘fully human’ person ‘on whom nothing is lost,’ as Abraham Maslow and Henry James would later capture this psychological ideal.¹⁴⁷

Not only does this ideal Platonic agent possess comprehensive first-person experience of the full range of humanly possible pleasures, he combines such experiential input with an exceptional ability to confront it with facts, logic, and rationality. This makes him a ‘competent judge,’ as J. S. Mill would later call a similar figure in his *Utilitarianism*, who ‘speaks the truth (ἀληθέστατα λέγει, 582a1)’ and praises ‘competently’ or ‘with authority (κύριος, 583a4).’ More specifically, he is right when he claims that, of all possible pleasures, the pleasures ‘of the part of the soul with which we understand (τὸ ... ᾧ μανθάνει ἄνθρωπος, 580d9)’ are superior the pleasures that derive from our embodied nature as biological and social, other-directed creatures who care about material and reputational goods.

In the later stage of the argument, this line of thought gets amplified: Socrates suggests that ‘besides the pleasure of the φρόνιμος, other pleasure is neither entirely true nor pure, but like some sort of *trompe l’oeil* painting (οὐδὲ παναληθῆς ἐστὶν ἢ τῶν ἄλλων ἡδονὴ πλὴν τῆς τοῦ φρονίμου οὐδὲ καθαρὰ, ἀλλ’ ἐσκιαγραφημένη τις, 583b3–5).’ Plato’s thesis here, I take it, is that when it comes to non-ideal pleasures, what presents itself to us as a pleasure may not always be the real thing.¹⁴⁸ Put differently, ordinary, non-ideal pleasures—or at least some of them—are not just inferior, they are deceptive and not what they pretend to be: they are mere

¹⁴⁶ Throughout this chapter, references are to the *Republic* (unless otherwise noted) and I more or less follow Reeve’s (1992) translation of the Greek (provided by Slings (2003)) with small modifications here and there.

¹⁴⁷ Maslow (1971: 27–29) and H. James (1977: 133).

¹⁴⁸ This gloss on the infamous notion of false pleasure comes from Lovibond (1989: 215).

semblances, counterfeits, or simulacra (εἰδωλα) of true pleasure, to borrow Plato's own terminology. More than that, and rather puzzlingly, the discovery that ordinary pleasure turns out to be deceptive and defective is hailed as the 'strongest and most decisive (μέγιστόν τε καὶ κυριώτατον, 583b6–7)' strike available against Thrasymachus's immoralist challenge.

I believe that neither of these features of the argument have been properly appreciated by the readings currently dominating the literature. The argument itself—and especially its putative defence of what I call hedonic fallibilism—tends to be brusquely dismissed as a philosophical failure which clashes with the commonsensical thought that we cannot err in our self-ascriptions of pleasure. This dismissive reaction, in turn, explains why most existing treatments simply skate over the question of how book 9's critical account of ordinary pleasure is supposed to integrate with the *Republic* as a whole.¹⁴⁹

It is against the backdrop of these typical reactions that the project of this chapter is to be understood: I want to make sense of Plato's hedonic fallibilism—the view that what presents itself as a pleasure may not always be the real thing—in the expectation that a proper understanding of this view might also elucidate our understanding of book 9's place in the economy of the *Republic*.

Before proceeding, three points of interpretive housekeeping are in order. First, I should admit that it is difficult to get the precise nature of Plato's charge against normal, non-ideal pleasure in view. Indeed, there might be some truth in the oft-heard complaint that Plato's thinking pulls in different directions in book 9 of the *Republic* and that this renders the Pleasure Argument itself fragmentary,

¹⁴⁹ Even those who have tried to come to grips with the exact role of the Pleasure Argument in the economy of the dialogue tend to be critical. For such a sustained yet highly dismissive interpretation, see e.g. Kraut (1992).

ambiguous, and perhaps even inconsistent.¹⁵⁰ Throughout the argument, pleasure is most frequently called not ‘pure’ and not ‘true,’ but Socrates also uses a wide variety of other derogatory labels to denounce it and it is not obvious how these labels should be unpacked or how they hang together.¹⁵¹ And even if we bracket Socrates’s other labels for the moment, the central charge that ordinary pleasure is not true, not pure, and like a *trompe l’oeil* painting seems rather obscure and has thus given rise to a variety of interpretations.

Still, although there is much that defies easy interpretation, it strikes me that there is good reason to believe that the Pleasure Argument is—at the very least partially—in the business of arguing that non-actualized, non-integrated agents can be radically mistaken about the affective state they take themselves to be in. On what seems to be one of the most straightforward, most natural reading of this bit of text, Socrates’s claim that ‘the pleasure of other people’—such as Thrasymachus’s beloved tyrant—‘is neither pure, nor true, but like a *trompe l’oeil* shadow-painting (οὐδὲ παναληθῆς ... οὐδὲ καθαρὰ, ἀλλ’ ἐσκιαγραφημένη τις, 583b3–5)’ should be unpacked as the

¹⁵⁰ For such criticisms, see. e.g. Cross and Woosley (1964), Frede (1985), Frede (1992), and Gosling and Taylor (1982). Erginel (2011) is an excellent attempt to defuse these charges.

¹⁵¹ Such pleasure is explicitly or implicitly called ‘far removed [from true pleasure] (πάνυ πόρρω, 581e1),’ ‘merely necessary (τῷ ὄντι ἀναγκαίως, 581e3)’ (in the sense that it is unavoidable), ‘worthless (οὐδενὸς ἀξίαν, 581d2),’ ‘vulgar (φορτικὴν, 581d6),’ ‘worthless trash (καπνὸν καὶ φλυαρίαν, 581d7),’ ‘untrustworthy’ or ‘unreliable (ἀπιστοτέρως, 585e4),’ not ‘stable’ or ‘secure (βεβαίου, 586a6),’ connected with ‘insatiability (ἀπλησίαν, 586b3),’ ‘[a] mere simulacrum of true pleasure and like a *trompe l’oeil* painting (εἰδῶλοις τῆς ἀληθοῦς ἡδονῆς καὶ ἐσκιαγραφημέναις, 586b8),’ ‘a simulacrum of pleasure in respect of reality three stages removed from that other [pleasure] (ἡδονῆς τρίτῳ εἰδῶλω πρὸς ἀλήθειαν ἀπ’ ἐκείνου, 587c8–9),’ ‘alien (ἀλλοτρίαν, 587a6)’ rather than ‘authentic’ or ‘one’s own (οἰκείως, 586e1),’ ‘spurious (νόθαιν)’ rather than ‘genuine (γνησίως, 597b15–16),’ and like ‘a [two-dimensional] plane figure (ἐπίπεδον, 587d6)’ (rather than a three-dimensional solid figure).

radical proposal that what presents itself as a pleasure may not always be the real thing. Unlike the ideal Platonic agent, the majority of people are what I call hedonically fallible: they often mistake a merely apparent, fake pleasure for a genuine, authentic one and this means, more plainly speaking, that there are situations where they think they are experiencing pleasure, although they are not.¹⁵²

Even though this mistaken self-ascription interpretation of the argument is one of the most straightforward readings available, most interpretive alternatives have taken issue with it—mainly on grounds of interpretive charity. This brings me to the second caveat. Because commentators are typically driven by the attempt to save Plato from committing the philosophical blunder of defending hedonic fallibilism, my approach in this chapter is more or less indirect. If we can make sense of the text on the most natural, most obvious way of reading it—on which the argument is indeed suggesting that we can go wrong in our sincere belief that we are experiencing pleasure—the interpretive alternatives should lose their appeal. Put differently, I will more or less *assume* that Plato is trying to defend hedonic fallibilism in *Republic* 9 and then try to make sense of this idea, operating under the further assumption that we have not yet run out of exegetical and philosophical resources for developing an alternative interpretation of this admittedly puzzling line of thought.

There is a third and last caveat. Like the larger argument to which it belongs, Plato's defence of hedonic fallibilism itself is fragmentary and varied. Following other commentators,¹⁵³ I think the later part of the Pleasure Argument can be divided into two strands of argumentation—the first running from 583b1 to 585a7, the second

¹⁵² For similar construals of this argument, see J. Butler (1999), Erginel (2006), Wolfsdorf (2013-a), Wolfsdorf (2013-b).

¹⁵³ Cf. e.g. Erginel (2019: 98), Mooradian (1992: 87), White (1979: 229), Wolfsdorf (2013-a: 110), Wolfsdorf (2013-b: 63).

running from 585a8 to 588a10—both of which articulate and defend the thesis that we can go wrong in our self-ascriptions of pleasure.

The focus of this chapter will be on the first of these different argumentative strands, though, the one running from 583b1 to 585a7, which I call the Deceptive Pleasure Argument. This bit of argumentation has not received the attention it deserves, even though it contains a cogent and interesting line of thought worth taking seriously.¹⁵⁴ What is more, unlike the later stretch of argumentation, it does not lean on the contested heavy-duty metaphysics of the middle books, it does not force us to take a stand on what it entails to cognize the Forms, and it is not embodied in a convoluted and problematic bit of text.

What should emerge is a textually and exegetically sound, philosophically attractive, yet unexplored interpretive option concerning Plato's Deceptive Pleasure Argument and its defence of hedonic fallibilism. Here is the plan for the chapter. In section 1, I discuss a standard objection against hedonic fallibilism and show how this objection has given rise to two ways of reading the argument. In sections 2 and 3, I offer a close reading of the Deceptive Pleasure Argument. Finally, by way of conclusion, section 4 places the argument in a wider context and discuss some of its ramifications.

1. Hedonic Fallibilism, a Standard Objection, and Interpretive Alternatives

¹⁵⁴ Plato himself seemed to have thought so at any rate: the *Philebus*—whose treatment of pleasure is sometimes taken to be a polished version of the rougher discussion found in *Republic* 9—takes up the first strand of argumentation but leaves the second line of argumentation behind.

The Platonic proposal that agents are hedonically fallible is typically met with resistance. According to a standard objection, Plato's defence of hedonic fallibilism rests on a philosophical mistake. Because our pleasures and other *πάθη* are transparent to us, the thought goes, there is no room whatsoever for doubt, ignorance, or error in the affective domain: human beings are hedonically omniscient and thus hedonically infallible. Put differently, I am always in a position to know whether I am experiencing pleasure or something else; whenever I am experiencing pleasure, I will also *believe* (or *know*) I am experiencing pleasure; and—conversely and most importantly—whenever I *believe* I am experiencing pleasure, I *am* in fact experiencing pleasure. On this view, pleasure is essentially subjective and there is no appearance reality gap when it comes to hedonic experiences: something *is* pleasant just in case the subject of the pleasure candidly reports that it *seems* pleasant to them.¹⁵⁵

This standard objection has generated two typical reactions to Plato's alleged defence of hedonic fallibilism in the Pleasure Argument. A first family of interpretations—Dismissive Readings, as I call them—holds that because hedonic fallibilism is an untenable position Plato *should* not hold, we can dismiss this argument as a failure. As Rosen puts it, to give just one example more or less taken at random from a larger sample:

[P]lato does not, and in my view cannot, explain the difference between a fantasm or illusion of pleasure and pain and the real thing. If I believe that I am in pain, then I am, even if the *source* of the pain is imaginary.¹⁵⁶

Summarizing the scholarly *communis opinio*, Crombie similarly claims

¹⁵⁵ For good discussion of this objection, see Erginel (2006) and Annas (1980: 296) and Annas (1981: 307).

¹⁵⁶ Rosen (1966: 336–37).

that, in the affective domain, the distinction between appearance and reality is ‘quite out of place,’¹⁵⁷ Grote resists the proposal because a pleasure ‘is what it seems, neither more nor less; its essence consists in being felt,’¹⁵⁸ both Gosling and Taylor and Urmson blame Plato’s mistake on a naïve and unsophisticated theory of pleasure,¹⁵⁹ Cross and Woosley think that so much as taking the pleasure arguments of book 9 seriously would be tantamount to paying Plato the ‘doubtful compliment’ of undeserved philosophical attention,¹⁶⁰ and Gibbs is even willing to entertain the possibility that ‘the rhetorical, fragmented, and ultimately inconsistent character of [*Republic* 9’s pleasure arguments] reflects Plato’s own intellectual limitations.’¹⁶¹

Exercising interpretive charity, a second family of interpretations—Evasive Readings, as I call them—tries to develop alternative ways of understanding book 9’s project which manage to find a way around Plato’s alleged defence of hedonic fallibilism. This type of interpretation suggests that, upon closer scrutiny, it will turn out that hedonic fallibilism is a position Plato in fact *does* not hold. Whatever grounds it might have for stigmatizing ordinary pleasure, the Deceptive Pleasure Argument is not defending the palpably false thesis that ordinary people can go wrong in their honest reports that they are experiencing pleasure.

Thus, when Socrates at some point in the argument concludes that ordinary, non-ideal pleasure involves ‘some kind of trickery (γοητεία τις, 584a10)’ as well as ‘untrustworthy or deceptive appearances (οὐδὲν ὑγιές τούτων τῶν φαντασμάτων, 584a9)’ detached from the ‘reality’ or ‘truth of pleasure (ἡδονῆς ἀλήθειαν,

¹⁵⁷ Crombie (1962: 139).

¹⁵⁸ Grote (1875: 602).

¹⁵⁹ Gosling and Taylor (1982: 451) and Urmson (1984: 213).

¹⁶⁰ Cross and Woosley (1964: 265).

¹⁶¹ Gibbs (2001: 33).

584a9),’ the evasive reading maintains that the argument is actually merely suggesting that ordinary pleasure fails to get at genuine value, say, or that it goes wrong in some other way.¹⁶² In what follows, I aim to find a middle way between these interpretive alternatives. The reading I defend here takes Plato at his word yet also attempts to bring out what is most cogent and most interesting about the Deceptive Pleasure Argument (understood as a defence of hedonic fallibilism).

2. The Deceptive Pleasure Argument

2.1. Misleading Pleasure

The Deceptive Pleasure Argument opens with the sweeping claim that all pleasure besides the pleasures of the ideal agent is ‘neither true nor pure but like a *trompe l’oeil* shadow-painting (οὐδὲ παναληθῆς ἐστὶν ἢ ... ἡδονὴ ... οὐδὲ καθαρὰ, ἀλλ’ ἐσκιαγραφημένη τις, 583b3–5).’ Because these charges of falsity and impurity constitute the core of Socrates’s attack on non-ideal pleasure, let me briefly gloss these notions. Plato typically calls pleasures pure when they ‘are not derived from pains (οὐκ ἐκ λυπῶν, 584a12),’ that is to say, when they do not just consist in ‘an escape from pain (τὴν λύπης ἀπαλλαγὴν, 5854b9–c1)’ or ‘[a] cessation of pain (παῦλαν λύπης, 584b2).’¹⁶³

¹⁶² Below, in section 2.3, I will discuss (different sub-types of) this interpretive option more carefully.

¹⁶³ The notion of impurity also has important religious and spiritual connotations: in eschatological contexts, it is often associated with the polluting influences of the body. Cf. e.g. *Phaedo* 67a5–6, discussed extensively in chapter 1, where Socrates argues that we should ‘keep ourselves pure from [the body] (καθαρεύωμεν ἀπ’ αὐτοῦ)’—especially its lowly pleasures and desires. For good discussion of this puritanism, see Dodds (1951) and Morgan (1992).

The notion of truth operative in *Republic* 9 is harder to capture, if only because the word has a variety of meanings. Crudely speaking, something can be called true or false in terms of what it *represents* or in terms of what it *is*.¹⁶⁴ A belief or utterance is called false, for instance, if it fails to represent the world correctly, but we can also speak of a true van Gogh painting rather than a forgery, if it really has all the properties of a van Gogh painting.¹⁶⁵ To capture some sort of intuitive unity between these different senses, I want to adopt a fruitful suggestion developed by other scholars who claim that, for Plato, a pleasure is false just in case it deceives us about some way the world really is—where this includes deception about the hedonic experience *itself*—and that a pleasure is true just in case it does not.¹⁶⁶

This understanding of the falsity of pleasure harmonizes nicely

¹⁶⁴ For this characterization, see Wolfsdorf (2013-b: 63).

¹⁶⁵ When understood representationally, the property of being true or false is attributed to beliefs, belief-like states, or utterances whose propositional content represents—or fails to represent—the world correctly; when understood ontologically, the property of being true or false is attributed to some *F* that has—or fails to have—all the properties of really being an *F*. In these cases, ‘real’ or ‘genuine’ and ‘fake’ or ‘unreal’ can be substituted for ‘true’ and ‘false’ without loss of meaning (e.g. one might speak of the true Beatles as opposed to a cover band). It is worth noting, especially in the current context, that characterizing some *F* as ‘true’ or ‘untrue’ or ‘false’ often amounts to claiming that the untrue or false *F* *merely appears to be F*, without really being *F*, whereas the true *F* *really is F*. For more guidance, see especially Wolfsdorf (2013-a: 122–129) and Szaif (1998) and (2018).

¹⁶⁶ For this take on what Plato means by calling pleasures false, see especially Whiting (2014: 24) but also Fletcher (2018-b: 395–396). As LSJ (s.v. 2) attests, the adjective ψευδής is often used for lying or deceitful persons; in the *Sophist*, Plato connects ‘falsity (τὸ ψευδός)’ with ‘deception (ἀπάτη, 260c6)’ as well as ‘simulacra, imitations, and appearance (εἰδῶλων τε καὶ εἰκόνων ἤδη καὶ φαντασίας, 260c8–9)’; likewise, in *Republic* 10 (598e5–599a3), imitations and appearances are juxtaposed with manifestations of reality or truth and linked with deception and trickery; and when speaking of perceptual illusions, Aristotle uses the verb διαψεύδω (e.g. at *On Dreams* 460b22).

with the fact that Plato's suspicion of pleasure stems—to a significant degree at least—from his conviction that pleasure is inexorably connected with many types of illusion and deception.¹⁶⁷ In an important paper, Jessica Moss has convincingly argued that the intimate association between pleasure and illusion plays a pivotal role in Plato's moral psychology—especially in the *Republic*, where 'the association between pleasure and illusion that functioned almost as a background assumption of [other dialogues] is developed into a full-fledged theory.'¹⁶⁸

Earlier, at 413b4–c3, Socrates had claimed that people often abandon 'true belief' because they are 'victims of trickery (τοὺς γοητευθέντας)' who are 'under the spell of pleasure (ὕφ' ἡδονῆς κηληθέντες).'¹⁶⁹ And later, in book 10, Socrates will locate our susceptibility to perceptual illusions and our desire for ordinary pleasure in the same part of the soul. As Moss summarizes the moral psychology underlying the *Republic* and its emphasis on pleasure's tendency to mislead: '[t]he *Republic* characterizes appetite, the pleasure-desiring part of the soul, as *illusion-bound*.'¹⁷⁰ (my emphasis)

Although this is very often overlooked, the close link between pleasure and illusion also plays a pivotal role in the Deceptive Pleasure Argument. Indeed, as I hope to show, appreciating this connection will provide the key to unlocking the most promising interpretation of the argument. Socrates maintains that those who err hedonically are 'getting deceived (ἀπατῶνται, 585a5),' he calls their defective pleasure 'fakes,' 'semblances,' 'simulacra,' or 'counterfeits

¹⁶⁷ See Moss (2006) and, before her, Mooradian (1992).

¹⁶⁸ Moss (2006: 523). For a more detailed discussion of the *Republic*, with which I will engage below, see Moss (2008).

¹⁶⁹ Pleasure deceives, Socrates points out, and 'everything that deceives seems to bewitch (ἔουκε ... γοητεύειν πάντα ὅσα ἀπατᾶ, 413c4).'

¹⁷⁰ Moss (2006: 517).

(of true pleasure) (εἰδῶλοις τῆς ἀληθοῦς ἡδονῆς, 586b8 and 586c4),’ compares them to optical illusions or ‘*trompe l’oeil* shadow-paintings (ἐσκιαγραφημέναις, 583b5 and 586b8),’ claims that they are caused by ‘some kind of trickery’ or ‘magic (γοητεία τις, 584a10),’ and argues that defective pleasure is the product of ‘untrustworthy, illusory appearances (οὐδὲν ὑγιὲς τούτων τῶν φαντασμάτων, 584a9)’ rather than a manifestation of the ‘truth’ or the ‘reality of pleasure (ἡδονῆς ἀλήθειαν, 584a9).’

2.2. Platonic Affective Psychology – Five Key Ideas

Appreciating the connection between pleasure and illusion does not just get us to the heart of Plato’s thinking about pleasure in *Republic* 9, I also want to argue that the Deceptive Pleasure Argument identifies and examines a new type of hedonic deception or illusion. To see how Plato develops this new type of hedonic trickery, let us have a closer look at the text. In collaboration with his interlocutor Glaucon—who is supposed to ‘answer questions (σοῦ ἀποκρινομένου)’ while Socrates ‘seeks (ζητῶν, 583c1)’—Socrates develops some basic ideas about our affective psychology. Here is the text:

[T2.1] [Socrates] Tell me, then, don’t we say that pain is the opposite of pleasure (ἐναντίον λύπην ἡδονῆ)?

[Glaucon] Yes.

[S.] Isn’t there also a state of feeling neither pleasure nor pain (τὸ μήτε χαίρειν μήτε λυπεῖσθαι)?

[G.] There is.

[S.] Isn’t it intermediate between these two, in the middle (μεταξὺ τούτων ἀμφοῖν ἐν μέσῳ ὄν), a sort of calm of the soul (ἡσυχίαν τινὰ ... τῆς ψυχῆς) where these are concerned (περὶ ταῦτα)? Or wouldn’t you describe it that way?

[G.] I would. (583c3–9)

We can extract the following claims about human affectivity from this bit of text:

- (1) Pleasure and pain are opposites or contraries.
- (2) There is a third, intermediate affective experience, which is located 'in between' or 'in the middle of' pleasure and pain.
- (3) Unlike pleasure and pain, which are affectively positive and affectively negative respectively, this intermediate affective experience is affectively neutral—it consists in 'neither experiencing pleasure nor pain (τὸ μήτε χαίρειν μήτε λυπεῖσθαι).'
- (4) This affectively neutral intermediate experience is 'a sort of calm of the soul (ἡσυχίαν τινὰ ... τῆς ψυχῆς),' that is, it is a state.

Slightly later, at 583d9–10, Socrates adds a crucially important fifth idea:

- (5) Pleasure and pain themselves are 'a sort of motion or change of the soul (τό ἡδὺ ἐν ψυχῇ γιγνόμενον καὶ τὸ λυπηρὸν κίνησις τις ἀμφοτέρω),' that is, they are a process.

On the picture we are given here, pleasure and pain are two opposite ends of a scale of affective experience with positive valence (pleasure) on the one extreme of the spectrum, negative valence (pain) on the other, and neutral or indifferent valence (calm) in the middle, situated in between pleasure and pain. Apart from their differences in hedonic tone—whether they feel good, bad, or neutral—these affective experiences also differ, more essentially, in respect of the fact that pleasure and pain are *processes* while calm is a *state*, characterized more specifically—and privatively, to boot—as the *absence* of pleasure

and pain and other affective upheaval.¹⁷¹

2.3. Socrates's Fallibility Cases

Between 583c10 and 583d9, the argument turns to a set of problematic cases. Although these deviant cases or fallibility cases, as I will call them, appear to undermine the argument's underlying affective psychology and the clean distinctions between three different affective experiences that have just been drawn, Socrates instead treats them as evidence for our hedonic fallibility.

There are people, he points out, who mistake the affectively indifferent state of calm for an affectively positive state of pleasure when it gets juxtaposed with an affectively negative state of pain. Sick people, for instance, hold that 'nothing gives more pleasure than being healthy (οὐδὲν ἐστὶν ἥδιον τοῦ ὑγιαίνειν, 583c13)' – they even go so far as to claim that health is 'the most pleasurable thing there is (ἥδιστον ὄν, 583d1)' – and those who are 'in great pain (τῶν περιωδυνία ἐχομένων)' report that 'nothing is more pleasant than putting an end to their suffering (οὐδὲν ἥδιον τοῦ παύσασθαι ὀδυνώμενον, 583cd3–4).'

In the next step, Socrates claims, more sweepingly, that this

¹⁷¹ Surprisingly enough, the hedonist Aristippus of Cyrene is reported to have developed a strikingly similar theory of human affect which nicely illustrates the model Plato is working with here: 'There are three states in our constitution. In one, which is like a storm at sea (ἐοικυῖαν τῷ κατὰ θάλασσαν χειμῶνι), we feel pain (ἀλγοῦμεν); in another, which is similar to a smooth undulation stirred by a favorable breeze (τῷ λείῳ κύματι ἀφομοιουμένην ... οὐρίῳ παραβαλλομένην ἀνέμῳ), we feel pleasure (ἠδόμεθα) (for pleasure is a smooth motion (λείαν κίνησιν τὴν ἠδονήν). The third state, in which we feel neither pain nor pleasure (οὔτε ἀλγοῦμεν οὔτε ἠδόμεθα), is in the middle (μέσην εἶναι κατάστασιν) and is like a calm sea (γαλήνην παραπλησίαν οὔσαν).' SSR 4b.5 = Eusebius, *Preparation for the Gospel* 14.18.32.

mistake of mixing up calm with pleasure is actually widespread and generalizes to many other cases—including the defective, non-ideal pleasures of most people besides the φρόνιμος:

[T2.2] [T]here are many other similar circumstances in which you find people who, when they are in pain (ὅταν λυπῶνται), praise not experiencing pain or freedom from that sort of thing as most pleasant rather than experiencing pleasure (τὸ μὴ λυπεῖσθαι καὶ τὴν ἡσυχίαν τοῦ τοιοῦτου ἐγκωμιάζουσιν ὡς ἥδιστον, οὐ τὸ χαίρειν). (583d6–9)

After a back and forth with his interlocutor, in which Glaucon's deflationary and infallibilist take on these deviant cases is scrutinized, Socrates reaffirms his conclusion that these people are getting tricked: 'there is nothing trustworthy in these [hedonic] appearances as far as the truth about pleasure is concerned, only some kind of trickery (οὐδὲν ὑγιὲς τούτων τῶν φαντασμάτων πρὸς ἡδονῆς ἀλήθειαν, ἀλλὰ γοητεία τις, 584a9–10).'

At the same time, though, he also warns Glaucon not to stray in the other direction and overgeneralize the force of these examples by offering an all too sweeping reductive theory of all and every pleasure as nothing but 'a cessation of' or 'relief from pain (παῦλαν λύπης, 584b2)'. Socrates's point is just that there are some cases where people fall prey to deceptive pleasure, *not* that all pleasures are just a trick of the mind.¹⁷²

Influential evasive readings, most of which have entered the literature as a response to Plato's problematic defence of hedonic fallibilism, interpret these remarks as follows. Going from Socrates's

¹⁷² Such a far-reaching reductive take on the nature of any and all pleasure can be found in Speusippus and 'the grumpies' in the *Philebus* who boldly deny the existence of pleasure—all so-called pleasures are in fact nothing but an escape from pain—and dismiss pleasure's power as 'a trick (γοήτευμα, 44c8).'

warning to Glaucon and his observations about the deeper ‘nature (πεφυκέναι, 584b1)’ of pleasure, a first type of evasive reading claims that Socrates is merely taking issue with rival explanatory accounts of the *nature* of pleasure. Thus, Dorothea Frede has suggested that Socrates is basically ‘ruling out the paradox (he calls it ‘trickery’) that pleasure is merely the end of pain, and pain the end of pleasure.’¹⁷³

In a similar vein, Reeve claims that ‘[Plato] is not arguing that [the subjects in Socrates’s deviant cases] wrongly find this state to be pleasant. His argument is about the nature of pleasure, about what pleasure is, about what account of it is true’¹⁷⁴ and Warren reads the passage as ‘a neat dialectical argument against a potentially troubling set of opponents’ who deny the existence of an intermediate affective state and offer an alternative to Socrates’s own account of pleasure and pain.¹⁷⁵

A second type of evasive reading suggests that Socrates is merely repudiating ordinary pleasure on evaluative or ethical grounds, the basic charge being that such pleasure fails to get at genuine value or misinforms us about what matters. According to Nussbaum, to give a prominent example of such a reading, Platonic talk of ‘true pleasures’ should be construed as talk of ‘truly valuable activities’ and these, in turn, refer to:

those activities that are chosen in harmony with *true* beliefs about value or worth, as opposed to those in which agents take pleasure because they falsely believe them to have worth.¹⁷⁶

¹⁷³ Frede (1985: 158).

¹⁷⁴ Reeve (1988: 305).

¹⁷⁵ Warren (2011: 113).

¹⁷⁶ Nussbaum (1986: 141).

In stark contrast, ordinary activities or ‘false pleasures’ – including all appetitive activities – should be rejected ‘as lacking in true or intrinsic value,’¹⁷⁷ which is to say that the judgment that they are worth something is mistaken because it is ‘thoroughly based on illusion.’¹⁷⁸

Likewise, on Clerk-Shaw’s account, the basic problem *Republic* 9 identifies with most pleasure is that it is the root cause of pernicious ethical errors:

Felt pleasures do not merely offer an incomplete picture of what is noble and good; they also offer a distorted picture of what is noble and good, [so] the fact that we are subject to pervasive hedonic error rules out using felt pleasures as a reliable guide to what is good for us.¹⁷⁹

Daniel Russell, finally, suggests that Socrates calls certain pleasures less ‘true’ than other pleasures just in case they mistakenly suggest to those who experience these pleasures that their life is going better than it really is and this suggests that *Republic* 9’s critique of pleasure is in fact just a covert critique of inferior ways of life.¹⁸⁰

Going from Socrates’s initial set-up of the argument, a third and last type of evasive reading treats Socrates as some kind of behavioural economist *avant la lettre* who merely points out that people are bad at affective forecasting due to flaws in their rationality. From the skewered perspective of being ill, for instance, sick people might come to the false conclusion that being healthy is pleasant. Likewise, I might misestimate the unpleasantness of filing my tax

¹⁷⁷ Tellingly, she understands Socrates’s concluding remark – that there is οὐδὲν ὑγιὲς τούτων τῶν φαντασμάτων πρὸς ἡδονῆς ἀλήθειαν – as if he is talking about the distorted *views* ordinary people take of *the good*. (1986: 146).

¹⁷⁸ Nussbaum (1986: 146).

¹⁷⁹ Clerk-Shaw (2015: 151).

¹⁸⁰ Russell (2005: 127–35).

return when I fail to compare it, more rationally, with a long-term vision of its hedonic consequences in my life as a whole.¹⁸¹ In this vein, Murphy claims that Plato's argument is merely 'attacking statements about [the] prevalence [of pleasure and pain] made at a [temporal] distance and the generalizations of those who compare one condition of life with another'¹⁸² and Annas grants Plato the 'acceptable point' that 'we do sometimes use the language of appearance and reality about pleasures when we are contrasting long and short term views of some pleasures and pains.'¹⁸³

What binds these various evasive interpretive strategies is the conviction that the hedonic mistake Socrates is examining here does not concern one's current affective experience and whether one is in a given affective state or not. On evasive readings, the agents targeted by Socrates's argument do not go wrong in their honest, immediate, present-tense self-ascriptions of pleasure, they merely have false beliefs about other things which, somehow, bear on pleasure—albeit in a more derivative way.

Importantly, the falsity of these more derivative hedonic beliefs—about what explanatory account of pleasure is correct, about the pleasantness of future states of affairs, or about the evaluative status of the things one enjoys—does not impinge on the hedonic experiences *themselves*. This is crucial: if Socrates is not espousing hedonic fallibilism in this stretch of text but merely trying to expose relatively benign pleasure-related mistakes, located more downstream from one's hedonic experiences, it follows that the pleasures themselves remain unscathed and this means, even more importantly, that 'the usual charge against [the defective pleasure argument]—of violating the obvious rule that pleasures and pains ...

¹⁸¹ Annas (1981: 311).

¹⁸² Murphy (1953: 212–213).

¹⁸³ Annas (1981: 311).

cannot be unreal since their *esse* is *percipi*—need [not] be allowed weight,’ as Murphy concludes.¹⁸⁴

Evasive readings—especially the one put forward by Murphy—have been fairly influential, not in the last place because they are seen as a welcome solution to the putative problem posed by the Platonic proposal that we are hedonically fallible. Although there is much to say against these evasive readings, I will confine myself to a few observations and my focus will be on Murphy whose interpretation is arguably the most widely followed line in the scholarly literature.¹⁸⁵

To be sure, Murphy’s evasive reading finds some support in the text—especially Socrates’s initial remarks lend credence to the idea that he is merely arguing that people tend to go wrong in their predictions or generalizations about pleasure instead of repudiating their immediate self-ascriptions of pleasure.¹⁸⁶ I nevertheless believe that this type of reading faces serious difficulties, so let me give three strong reasons to think that we are instead dealing with mistaken self-ascriptions and deceptive pleasures which are not what they pretend to be.

First of all, note that, throughout the argument, Socrates takes issue with the pleasures themselves rather than with the higher-order, derivative beliefs one might *entertain* about such affective states. He calls the defective pleasures *themselves* untrue, compares them to

¹⁸⁴ Murphy (1953: 212–213).

¹⁸⁵ Cf. J. Butler (1999: 286n5): ‘Contemporary interpreters, such as Irwin, Frede, Gosling and Taylor, and Cross and Woosley, all follow Murphy’s lead to some degree.’

¹⁸⁶ It is perhaps worth stressing that although I defend a mistaken self-ascription reading of the argument here, I am open to the possibility that there are other hedonic mistakes at play which are captured and emphasized by alternative interpretations. My basic point is the following: *even if* Socrates is interested in other hedonic errors as well, he is *also* defending hedonic fallibilism—and this will be the focus of this chapter.

deceptive *trompe l'oeil* paintings, and denounces them as mere εἰδῶλα—flimsy, insubstantial, pale semblances of the real deal, ‘in respect of reality three stages removed from [true pleasure] (ἡδονῆς τρίτῳ εἰδῶλῳ πρὸς ἀλήθειαν ἀπ’ ἐκείνου, 587c8–9).’ The most natural, most straightforward way of understanding these charges of falsity, deception, and ontological defectiveness, I take it, is to unpack them as the proposal that it is possible to conflate a merely apparent pleasure for a genuine pleasure, thereby mistakenly self-ascribing a pleasure that is not really there.

Secondly, the exchange between Glaucon and Socrates that takes place in the course of the argument strongly suggests that the error of the relevant agents concerns immediate self-ascriptions of affective experiences rather than other, more derivative types of hedonic error. Having been confronted with Socrates’s fallibility cases, Glaucon pushes back in the following way:

[T2.3] [a] [Glaucon] That may be the case because at that moment the state of calm turns out to be¹⁸⁷ pleasant (τοῦτο γὰρ τότε ἡδὺ ἴσως καὶ ἀγαπητὸν γίγνεται, ἡσυχία).

[b] [Socrates] So [, if you are right,] after someone has stopped experiencing pleasure, the absence of pleasure will be painful (καὶ ὅταν παύσῃται ἄρα, εἶπον, χαίρων τις, ἢ τῆς ἡδονῆς ἡσυχία λυπηρὸν ἔσται).¹⁸⁸

[G.] Probably (ἴσως). (583d10–e3)

Part [a] might be taken to suggest that Glaucon is describing a future state of affairs in which—then and there, at that very moment—the calm will prove itself to be pleasant as soon as the preceding pain has

¹⁸⁷ I will explain and defend my choice for this translation of γίγνεται below.

¹⁸⁸ Note that ἡσυχία (with a genitive) does double duty here: it refers to the absence of or freedom from pain as well as the middle state of calm.

subsided.¹⁸⁹

This is confirmed by part [b], which is less ambiguous and more perspicuous. In it, Socrates complements Glaucon's remark and tries to sum up what his interlocutors think is going on in the deviant cases: '[on your account], then, after someone has stopped experiencing pleasure, the absence of pleasure will be painful (ὅταν παύσῃται ἄρα χαίρων τις, ἢ τῆς ἡδονῆς ἡσυχία λυπηρὸν ἔσται).' Although this is often missed,¹⁹⁰ the Greek suggests that we are dealing with two future states of affairs located at different times in the future: an anterior event expressed in the dependent clause and a posterior event expressed in the main clause.

That is to say, Socrates's gloss on Glaucon's claim expresses a situation in the future where the pleasure has come to an end, after which the calm will feel painful.¹⁹¹ Conversely, Glaucon's proposal—which Socrates is after all merely complementing—is best understood in a parallel way: he must be talking about a similar future situation

¹⁸⁹ A defender of a Murphy-style affective forecasting interpretation might object that Glaucon's 'at that moment' (τότε) instead refers to the moment at which the irrational people formulate their false predictions about the future, the idea being that from a skewered, painful perspective, the future state of calm might look pleasant, although it is not. The way in which Socrates takes up Glaucon's suggestion rules this reading out, though, and decides in favour of my interpretation.

¹⁹⁰ E.g. in the translations by Reeve (1992) and Shorey (1935). In their defence, it is worth noting, though, that the English language often uses what is known as a concealed future—a present-stem form—in temporal clauses referring to future states of affairs.

¹⁹¹ Cf. Rijksbaron (2002: 80): 'ἐπίν, ὅταν etc. + aorist subjunctive [rather than present subjunctive]: the state of affairs of the dependent clause is anterior to that of the main clause.' For a parallel construal, cf. Euripides, *Iphigeneia in Tauris* 625: 'τάφος δὲ ποῖος δέξεταί μ', ὅταν θάνω;' ('what tomb shall receive me, after I have died?'). I want to thank Stephen White and especially Gerry Wakker for useful discussion of these grammatical issues and the importance of this bit of text for my mistaken self-ascription reading.

where the pain has come to an end, after which the calm will manifest itself as something pleasant.

What this suggests, then, is that when the irrational people engage in sloppy affective forecasting and claim that putting an end to pain is pleasant, they are doing so for the simple reason that this is how they are in fact going to experience their convalescence against the backdrop of their current state of uncomfortable sickness. They, so to speak, practice what they preach: when the moment is there, their recovery will really manifest itself as a pleasure—albeit a false and deceptive one—and it will really *feel* good.¹⁹²

That we are dealing with mistaken immediate self-ascriptions rather than mistaken predictions or generalizations seems to be confirmed by what follows.¹⁹³ When Socrates summarises his results, he echoes his earlier exchange with Glaucon¹⁹⁴ and offers a more general observation about these cases, using present-tense forms: when juxtaposed with pain, the affectively neutral state of calm ‘merely seems—but is not really—pleasant (οὐκ ἔστιν ..., ἀλλὰ φαίνεται, 584a7).’

Of course, this could again involve the juxtaposition of a present state of pain with a future state of calm—as if we are dealing with generalizations and predictions about the future made from a blurry current perspective—but the way in which Socrates explicates these mistakes by means of an analogy rules out such an evasive

¹⁹² Thus, the affective forecasting reading goes hand in hand with a mistaken self-ascription reading: in Socrates’s example, the hedonically mistaken individual predicts that their convalescence is going to be pleasant (although it is neutral) and when the time is there, they will experience their convalescence as a pleasure (although it is neither a pleasure nor a pain, but an affectively indifferent state).

¹⁹³ Cf. J. Butler (1999: 292).

¹⁹⁴ Adam (1902: 350) rightly notes that ‘οὐκ ἔστιν κτλ. contradicts Glaucon’s suggestion τοῦτο γὰρ—ἡσυχία 583d,’ the point being that ‘there is no *reality* about this (as *you* [Glaucon] suggest): it is only a φάντασμα.’

construal (584d1–585a7). When it comes to pleasure, Socrates suggests, ordinary people are like people who live in a house with three floors, although they are only aware of the basement and the ground floor. This tricks them into thinking they have reached the highest floor when they have in fact moved from the basement to the ground floor.¹⁹⁵

Likewise, non-ideal agents are ‘deceived (*ἀπατῶνται*, 585a5)’ into believing that they are experiencing pleasure, even though an affectively negative state of pain merely gets juxtaposed with an intermediate state of affective neutrality or painlessness: ‘whenever [non-ideal agents] are moved from pain to the intermediate state, they strongly’—but wrongly—‘believe they are in a state of filling and pleasure (*σφόδρα μὲν οἴονται πρὸς πληρώσει τε καὶ ἡδονῇ γίγνεσθαι*, 585a2-3).’¹⁹⁶ Even though the intermediate state of calm is—by definition (as per 583c5)—not a pleasure, one might mistake this state of painlessness for a state of pleasure thanks to a contrast effect and the juxtaposition with pain.¹⁹⁷

It seems fairly obvious, then, that irrational people go wrong

¹⁹⁵ This example comes from Warren (2011: 133).

¹⁹⁶ Following the translation of Wolfsdorf (2013-a: 117). For this translation of *γίγνεσθαι* + *πρὸς*, see LSJ s.v. II.3.c: ‘to be near something,’ ‘to be at something.’ For the analogy between the topological and the hedonic domain to make sense, this translation seems required: in the topological example, someone ‘standing in the middle’ (the intermediate affective state) believes himself to be in the upper region (pleasure); likewise, someone in an intermediate state thinks they are in a pleasant state.

¹⁹⁷ Socrates develops a second analogy (585a3–5): the victims of this hedonic mistake are like people who take themselves to be seeing something white, even though they are merely looking at a patch of grey which gets confused for something white because it appears against the backdrop of patch of black.

in their first-person, present-tense self-ascriptions of pleasure.¹⁹⁸ Even though they sincerely believe and report otherwise, they are not experiencing pleasure: in reality (τῶ ὄντι, 585a1), they find themselves in an affectively neutral psychological condition of pain- and pleasureless calm which they *mistake* for a pleasure.

Finally, there is a third difficulty for the evasive readings put forward by Murphy and others. Such readings are arguably in a weaker position to deliver a cogent story about how Plato's demonstration of the superiority of the rationally integrated life is meant constitute the 'strongest and most decisive' strike against Thrasymachus's sceptical project and make trouble for the underlying rival theory of the good life. It is hard to see, in other words, how sloppy generalisations, poor affective forecasting skills, adherence to faulty accounts of the nature of pleasure, or defective evaluative beliefs would constitute the promised knock-down argument against the life of the tyrant preferred by Thrasymachus and other implicit and explicit hedonists,¹⁹⁹ unless such derivative hedonic errors would ultimately bleed into the hedonic experiences themselves and make them less pleasant than they seem to be and hence not worth having.

It is telling that when Socrates targets Thrasymachus's beloved tyrant towards the end of the Pleasure Argument (587a4–588b5), he is not concerned with the derivative mistakes identified by the evasive readings. Instead, he goes to great lengths to point out how impoverished, flimsy, and one-dimensional the tyrant's pleasures really are, basically arguing that the tyrant is seriously mistaken about his pleasures. The tyrant *thinks* he is well off, hedonically speaking,

¹⁹⁸ As Wolfsdorf (2013-a: 117) puts it: 'The mover is not predicting where he will stand as he moves from the bottom. Rather, he believes he is at the top when is standing at the midpoint.' Cf. J. Butler (1999) and Erginel (2006) for the same point.

¹⁹⁹ For the claim that Plato's tyrant should be understood as a pleasure-seeker, see Johnstone (2015) and Moss (2005).

although he is not. The crucial point here—and the point that should bother Thrasymachus who seems to flirt with such a way of life on hedonistic grounds—is that the tyrant’s pleasures are not what they pretend to be: they are insubstantial, defective simulacra of the real deal, although this bit of information about the contents of his very own mind is not introspectively available to the psychologically fragmented tyrant whose blurred vision makes him a stranger to himself.²⁰⁰

2.4. Glaucon’s Infallibilism

In the next stretch of the argument, Glaucon pushes back against Socrates’s proposal and offers his own infallibilist explanation of the deviant cases. Glaucon objects that the people in these cases merely say these things for the simple reason that ‘at that moment, this is perhaps what turns out to be pleasant and desirable—[the] calm (τοῦτο τότε ἤδὺ ἴσως καὶ ἀγαπητὸν γίγνεται, ἡσυχία, 583d10–11).’ Put differently, Glaucon holds that Socrates’s examples do not count as instance of serious hedonic error: the people in his deviant cases do in fact experience pleasure, just as they claim. If that is the case, though, Socrates adds, it would also follow—conversely and from the same premises—that ‘after someone has stopped feeling pleasure, this calm or freedom from pleasure will be painful (ὅταν παύσηται ἄρα χαίρων τις, ἢ τῆς ἡδονῆς ἡσυχία λυπηρὸν ἔσται, 583e1–2).’

Like many contemporary students of Plato, Glaucon assumes

²⁰⁰ Both in the *Gorgias* and in *Republic*, the tyrant is used as a prime illustration of the Platonic view that we are not transparent to ourselves and that we can only discover what we really believe, desire, and feel by leading an examined life. See Brickhouse and Smith (1994: 73–102), Lear (1998: 59–60), and Parry (1992) for discussion of the tyrant’s deeply impoverished self-knowledge.

that there is a more parsimonious and more down-to-earth explanation of Socrates's fallibility cases available which does not commit us to the counter-intuitive notion that people are hedonically fallible. This alternative take on Socrates's fallibility cases seems best understood as driven by some kind of phenomenal subjectivism about our affective psychological states. Such affective phenomenal subjectivism, in its turn, boils down to the thesis that there is no appearance reality gap in the affective domain. When it comes to pleasure, pain, and certain other *πάθη*, things are to any person just as they appear to that person.²⁰¹ A pleasure's immediate phenomenological quality—how it *seems* to the subject of the experience or how it *feels*—exhausts its reality or how it *is*. Pleasure is just a bit of felt experience—neither more nor less. Crucially, this picture obviously guarantees our hedonic infallibility: if the ontology of pleasure is a first-person, *esse est percipi* ontology, any situation in which it strikes me *as if* I am experiencing pleasure counts as a situation in which I *am* experiencing pleasure.²⁰²

²⁰¹ Following Chappell (2005: 62 and 241).

²⁰² Plato was not unaware of this difficulty. In the *Theaetetus* (179c2–4), Socrates claims that '[s]o long as we keep within the limits of that immediate present experience of the individual (τὸ παρὸν ἐκάστῳ πάθος) which gives rise to sensory experiences and to sensory judgments (αἰ αἰσθήσεις καὶ αἰ κατὰ ταύτας δόξαι), it is more difficult (χαλεπώτερον) [but not impossible] to convict these latter of being untrue (οὐκ ἀληθεῖς).' In the *Philebus*, Socrates wrestles with Protarchus' infallibilism (there construed as the claim that pleasures cannot be false): not only does he mention people who are dreaming, mad, or suffering from other delusion as possible candidates for hedonic fallibility (36e5–8), he also develops an argument—briefly touched upon below—which suggests that people can go wrong in their immediate self-ascriptions of pleasure (42c5–44a11).

2.5 Two Arguments Against Glaucon's Infallibilism

In what follows, Socrates aims to show that Glaucon's alternative explanation of these putative cases of hedonic fallibility falls apart under sharper scrutiny. Here is the text in full, broken down in four parts:

[T2.4] [a] [Socrates] [If what you, Glaucon, are saying is right,] then the thing we described as being intermediate between the two, the calm, this thing will be both—both pain and pleasure (ὁ μεταξύ ... ἀμφοτέρων ... , τὴν ἡσυχίαν, τοῦτό ... ἀμφοτέρω ἔσται, λύπη τε καὶ ἡδονή).

[Glaucon] That's what it looks like.

[S.] Now, is it possible for that which is neither to be both (δυνατὸν τὸ μηδέτερον ὄν ἀμφοτέρω γίνεσθαι)?

[G.] Not in my view (οὐ μοι δοκεῖ)

[b] [S.] More importantly, when pleasure and pain take place in the soul, they are both some kind of process (τὸ ἡδὺ ἐν ψυχῇ γιγνόμενον καὶ τὸ λυπηρὸν κινήσις τις ἀμφοτέρω ἔσπτον).

[G.] Yes.

[S.] But didn't what is neither painful nor pleasant come to light just now as a state of calm and something in the middle between them (τὸ δὲ μήτε λυπηρὸν μήτε ἡδὺ οὐχὶ ἡσυχία μέντοι καὶ ἐν μέσῳ τούτοις ἐφάνη ἄρτι)?

[G.] Yes, it did.

[c] [S.] How can it be correct, then, to think that not feeling pain is pleasant or that not feeling pleasure is painful (τὸ μὴ ἀλγεῖν ἡδὺ ἡγεῖσθαι ἢ τὸ μὴ χαίρειν ἀνιαρόν)?

[G.] There's no way it can be.

[S.] Then it isn't correct. [d] But when the calm is juxtaposed with the painful it merely appears pleasant without really being pleasant (οὐκ ἔστιν ... τοῦτο, ἀλλὰ φαίνεται) and when the calm is juxtaposed with the pleasant it merely appears painful without really being painful (παρὰ τὸ ἀλγεῖν ἡδὺ καὶ παρὰ τὸ ἡδὺ ἀλγεῖν τότε ἢ ἡσυχία), but there is nothing trustworthy in these deceptive appearances as far as the truth about pleasure is concerned, only some kind of trickery (οὐδὲν ὑγιὲς τούτων τῶν φαντασμάτων πρὸς

ἡδονῆς ἀλήθειαν, ἀλλὰ γοητεία τις).

[G.] That's what our account (λόγος) suggests at any rate. (583e4–584a11)

Parts [a] and [b] contain what I take to be two separate yet closely interconnected arguments against Glaucon's infallibilist response, designed to safeguard and reaffirm the earlier conclusion—forced upon us by the fallibility cases—that we can go wrong in our self-ascriptions of pleasure.²⁰³ As Socrates puts it in part [c], his arguments are meant to (re-) establish the conclusion that it is not correct 'to think that the absence of pain is pleasant (τὸ μὴ ἀλγεῖν ἢδὲ ἡγεῖσθαι, 584a4).' Part [d], finally, develops an explanation as to why people fail to recognize what affective state they are in, which pivots on a distinction between appearance and reality or seeming and being and introduces the notion of hedonic illusions or deceptive pleasures. These parts are best discussed in turn.

2.5.1. Sub-Argument 1: the Dialectical Argument

In part [a], Socrates offers his first-sub-argument against Glaucon's infallibilism. At first blush, this argument seems undeveloped and

²⁰³ Although Socrates also describes cases where the calm merely seems to be painful without really being painful, the overall focus of the argument lies on cases of misleading and merely apparent pleasure. The argument is not just explicitly concerned with the falsity, impurity, and deceptiveness of pleasure rather than pain, Socrates also claims that self-ascriptions of pain almost never go wrong: when non-ideal agents 'think they are in pain when they ascend to the painful, ... they really are in pain (ὅταν ... ἐπὶ τὸ λυπηρὸν φέρωνται, ἀληθῆ τε οἴονται καὶ τῷ ὄντι λυποῦνται, 585a1).' This is an important qualification, because Plato's critics—and other optimists about introspection—often use the example of pain to reduce their opponents' pessimism about self-knowledge to absurdity. This point is well made by Schwitzgebel (2008: 259).

commentators have struggled to make sense of it.²⁰⁴ This interpretive struggle seems to stem from our inability to get a decent handle on the puzzling principle that ‘what is neither [*F* nor *G*] cannot γίγνεσθαι both [*F* and *G*]’ (583e7) which, somehow, allows Socrates to reduce Glaucon’s initial position to absurdity and forces him to give up his hedonic infallibilism (at 584a6).

This principle is often taken to invoke book 4’s Principle of Opposites, the thought being that the principle claims that it is impossible for what is neither (an) *F* nor (a) *G* (where *F* and *G* are opposites or contraries) to *become* both *F* and *G* (at least not in the same respect, at the same time, and relative to the same thing). Understood that way, Socrates’s thought here would be that nothing whatsoever—including the calm—can become both pleasant and painful, at least not in the same respect, at the same time, or relative to the same thing.²⁰⁵

Socrates must be driving at something else, though. As we learn in the *Parmenides*, there is usually nothing ‘puzzling (ἄξιον θαυμάζειν, 129c3)’ about the compresence of contrary properties. To unravel the superficial clash between opposites, we often just have to dig a bit deeper and add the relevant qualifiers. In this sub-argument, Socrates even supplies these qualifiers himself: he explicitly states that the calm is painful when it gets juxtaposed with pleasure and pleasant when it gets juxtaposed with pain which goes to show that the calm is not pleasant and painful in the same respect, at the same time, and towards the same thing.

Fortunately, there is a more parsimonious and more elegant reading of this passage available which does not run into these

²⁰⁴ Wolfsdorf (2013-a: 113) calls it ‘the puzzling argument.’

²⁰⁵ For readings along these lines, see e.g. J. Butler (1999: 291–92), Warren (2011: 121), White (1979: 229).

problems.²⁰⁶ Note that at three points in the argument Plato shifts nonchalantly and without explanation between instances of εἶναι and γίγνεσθαι. This happens at 583e3, where Socrates's ἔσται takes up Glaucon's γίγνεται almost in the same breath. It happens slightly later as well: when Socrates posits his principle at 583e7, γίγνεσθαι refers back to and is meant to block the inference that the state of affective calm ἔσται both pleasure and pain (583e5). Finally, this equation occurs again in the last part of our passage, where Socrates echoes Glaucon's initial suggestion and again picks up the latter's γίγνεται (as well as his own ἔσται) by an ἔστιν: 'the calm *is* not pleasant, it merely *seems* pleasant (οὐκ ἔστιν ... ἀλλὰ φαίνεται, 584a7).'²⁰⁷

It is obvious, then, that Socrates equates εἶναι and γίγνεσθαι in his first sub-argument against Glaucon's infallibilism. What is not so obvious, is what we should make of this equation. Let me offer a promising way to understand this otherwise puzzling and problematic move: because the *Aktionsart* of the verb εἶναι is stative or durative rather than kinetic, mutative, or dynamic, forms of the verb γίγνεσθαι are sometimes used in a dynamic sense of εἶναι.²⁰⁸ When this happens, instances of γίγνεσθαι are used not in the meaning of 'coming into being,' 'becoming *F*,' or 'coming into a certain state *F*,' but in the meaning of 'being *F* as result,' 'showing or proving itself to be *F*,' 'turning out to be *F*,' or simply 'being *F*' (with the dynamic connotation suppressed).²⁰⁹

²⁰⁶ I want to thank Tamer Nawar, Katja Vogt, and especially Job van Eck who have significantly improved my understanding of this tricky passage.

²⁰⁷ Adam (1902: 350).

²⁰⁸ See e.g. Kahn (1981: 111) and van Eck (2000).

²⁰⁹ Cf. e.g. *Sophist* 256b3: 'τρία δὴ γίγνεται ταῦτα (these *prove to be* three [classes]),' *Philebus* 64d4–5: 'δι' ἣν ἢ παντὸς ἀξία γίγνεται ἠτισοῦν ἢ τὸ παρὰπαν οὐδενός (thanks to which it *turns out to be* either most valuable or worth nothing at all),' and Euripides, *Medea* 14–15: 'ἥπερ μεγίστη γίγνεται σωτηρία, ὅταν γυνὴ πρὸς ἄνδρα μὴ διχοστατῇ (it *proves to be* the greatest source of safety, when a woman is not at odds

If we construe the principle along these lines, Socrates's line of thought is more or less tautological: because what is not-*F* cannot turn out to be *F* or, more precisely, because what is neither *F* nor *G* cannot turn out to be both *F* and *G*, the non-pleasant and non-painful state of calm cannot turn out to be pleasant and painful—*pace* Glaucon, who had suggested that the neutral, painless, and pleasureless state of psychological calm will sometimes be both pleasant and painful, thanks to the juxtaposition with a contrastive affective state.

More plainly speaking, the guiding idea here must be something like the following. If we agree that an apple, say, is neither a pear nor an orange, it is trivially true that an apple cannot prove itself to be an orange—even in strange circumstances or when juxtaposed with a pear—although it might be mistaken for an orange (because it looks that way, for instance). Likewise, if we agree that the affectively neutral state of calm is neither a pleasure nor a pain, it is trivially true that the calm cannot turn out to be either of them (or both for that matter)—even though it might be mistaken for a pleasure (because it looks that way in funny circumstances, for instance).²¹⁰

A passage from the *Philebus* sheds light on Socrates's line of reasoning. In it, Socrates explains to his interlocutor Protarchus that according to the 'correct explanatory account of the matter (τὸν ὀρθὸν λόγον, 43e10),' those who claim to be experiencing pleasure when they find themselves in a state of painless psychological calm are mistaken. Although they *think* they are experiencing pleasure, they are

with her man).' As Mastrorarde (2002: 165) comments: "'proves to be' or 'shows itself to be' or the like is frequently the best English equivalent for γίγνομαι followed by a predicate noun or adjective.'

²¹⁰ This is confirmed by part [d], where Socrates introduces a distinction between mere appearance and reality. Although an apple cannot (prove or show itself to) *really be* an orange, it might nevertheless *strike me* as an apple—if I look at it in unfavourable perceptual circumstances, for instance, or if my (phenomenal) concept of apples is defective.

not. Put differently, *not* experiencing pain is not, and cannot be, the same as experiencing pleasure—even if there are people who suggest otherwise because they report that they experience pleasure when they are merely in a state of pain- and pleasureless calm.²¹¹ Here is why:

[T2.5] [Socrates] Imagine three sorts of things, whichever you may like, and because these are high-sounding names, let us call them gold, silver, and what is neither of the two (τριῶν ὄντων οὐκ ἡμῖν, ὄντων βούλει, τίθει, καλλίσουσιν ἵνα ὀνόμασι χρῶμεθα, τὸ μὲν χρυσόν, τὸ δ' ἄργυρον, τρίτον δὲ τὸ μηδέτερον τούτων.)

[Protarchus] Consider it done.

[S.] Is there any way conceivable in which this third kind, which is neither of both, could turn out to be the same as either of the other two sorts, gold or silver (τὸ δὴ μηδέτερον τούτων ἔσθ' ἡμῖν ὅπως θάτερον γένοιτο ἄν, χρυσοῦς ἢ ἄργυρος)?

[P.] How could it?

[S.] That the middle kind of life be described as either pleasant or painful would be the wrong thing to think, if anyone happened to think so, and it would be the wrong thing to say, if anyone would say so, at least according to the right explanatory account of the matter (κατὰ γε τὸν ὀρθὸν λόγον). (*Philebus* 43e1–10, my translation, loosely based on Frede)

Unlike most other translators,²¹² D. Frede construes the expression γένοιτο ἄν exactly right—as a dynamic, kinetic, or mutative version of εἶναι. Plato's suggestion is *not* that what is neither gold nor silver

²¹¹ Cf. *Philebus* 43d4–5 for this identity claim: 'S.: Not experiencing pain would not be identical with experiencing pleasure (οὐκ ἄν εἴη τὸ μὴ λυπεῖσθαι ταῦτόν τῳ χαίρειν), right? P.: Certainly not!'

²¹² Cf. Fowler (1962: *ad loc.*): 'Now can that which is neither *become* either gold or silver?' and Gosling (1975 *ad loc.*): 'Could it possibly *become* gold or silver?' (my emphasis) Hackforth (1974: *ad loc.*) is slightly closer to the mark: 'Could we possibly *identify* the third *with* either of the others, gold or silver?' (my emphasis)

(bronze, say) cannot *become* gold; his more reasonable suggestion is that, while remaining itself, bronze, which is neither gold nor silver, cannot *turn out to be* gold or silver, even in special circumstances or when juxtaposed with a different mineral.

Likewise, or so Socrates argues, affectively indifferent psychological calm, which is neither pleasure nor pain but something in between the two, cannot somehow turn out to be a pleasure—even when it gets juxtaposed with a contrastive affective state and even if there are people who report that they are experiencing pleasure when they are not in pain because they find themselves in this state of calm affective indifference. This is so, Socrates adds in a crucially important rider, because ‘not experiencing pain and experiencing pleasure each have a different or independent nature (χωρίς τοῦ μὴ λυπεῖσθαι καὶ τοῦ χαίρειν ἢ φύσις ἐκατέρου, 44a10).’²¹³

The guiding thought here seems to be, then, that there is an intrinsic, essential, underlying feature—or cluster of such features—that is common and peculiar to all pleasures and that sets it apart from other mental states, such as affectively neutral psychological calm. Given this underlying, essential feature of the different affective states—their ‘hidden structure’ or φύσις—it is simply impossible for affectively neutral calm to prove itself to be a pleasure or a pain.

This suggests the following interpretation of Socrates’s first, dialectical strike against Glaucon’s infallibilist gloss on the deviant cases. When laying down his affective psychology, Socrates basically *stipulates* that the intermediate state of affective calm is neither a pleasure nor a pain. (This is thesis 3 of Socrates’s affective psychology, touched upon above). In the remainder of the argument, he then reminds Glaucon that this necessarily entails that the calm cannot prove itself to be a pleasure or a pain—on the hardly disputable

²¹³ Cf. *Republic* 584b1–2, where Socrates argues that, as regards its φύσις (πεφικέναι), pleasure is emphatically *not* merely ‘relief from pain (παῦλαν λύπης εἶναι).’

assumption that it is not possible for what is not-*F* and not-*G* to be both an *F* and a *G*. Taking all this together, then, Socrates's *reductio* of Glaucon's position goes something like this:

1. The affectively neutral state of calm is neither a pleasure nor a pain.
2. Juxtaposed with pleasure, the calm is a pain. [Assumption]
3. Juxtaposed with pain, the calm is a pleasure. [Assumption]
4. The calm is both a pleasure and a pain. [From 2, 3]
5. It is not possible for what is not-*F* and not-*G* to be both an *F* and a *G*. [Assumption]
6. Contradiction. [From 1, 4, 5]

Admittedly, this interpretation of Socrates's first argument against Glaucon might seem a bit deflationary. From a dialectical perspective, however, Socrates's argumentative move makes perfect sense: he is just pointing out an inconsistency in Glaucon's belief set and reminds his interlocutor of his earlier dialectical commitments. When Socrates posited his affective psychology—including the claim that the calm is a different affective state, lying in between pleasure and pain, and characterized as neither of the two—Glaucon did not push back and immediately expressed agreement with the proposed model of human affect.²¹⁴

²¹⁴ Importantly, Plato makes it explicit that Socrates needs Glaucon's help in developing his thesis and that they are in this together: 'With you answering my questions and me seeking, I'll find out [what is wrong with ordinary pleasure].' (583c1) More broadly speaking, when arguing for a contentious position, Socrates tends to begin from premises that are available to his opponent. As he puts it in book 6, if you are arguing with an interlocutor who disputes the truth of what you are saying, your safest best bet is 'to soothe and persuade [your interlocutor] gently,

That is to say, the argument just records Socrates's attempt to trip up Glaucon in light of his earlier commitments: if he believes that calm and pleasure are different, separate affective states and that calm is not a pleasure, he cannot also believe—on pain of contradicting himself—that in some contexts the calm can turn out to be a pleasure. In still other words, and to put it as succinctly and plainly as possible, Socrates points out that the neutral state of calm cannot be a pleasure for the simple reason that it is the neutral state, and not a pleasure.²¹⁵ Even though the calm might deceptively *appear* to be a pleasure, it cannot really *be* a pleasure: the underlying essence of our mental states or psychological conditions decides whether we are experiencing a pleasure or something else—even if introspection suggests otherwise. To put it boldly, Plato's view is that reality—the underlying φύσις or hidden structure—does not care about our feelings.

2.5.2. Sub-Argument [2]: The Substantive Argument

Let us turn to Socrates's second sub-argument located in part [b]. On the face of it, Socrates more or less repeats his previous dialectical strategy: his argument seemingly takes a stipulated theory of our affective psychology for granted—on which pleasure is a κίνησις and the intermediate state of affectively indifferent calm is not—from which it deduces the conclusion that the intermediate state of

hiding from him that he isn't in his right mind (ἐπικρυπτόμενοι ὅτι οὐχ ὑγιαίνει, 476e1–2).' Fine (1990: 87) and Irwin (1977: 136) interpret this methodological statement as 'the dialectical requirement': in his exchange with the lovers of sights and sounds—and perhaps more generally (as suggested by e.g. *Meno* 75d2–7)—Socrates must argue from premises that are acceptable to his interlocutors.

²¹⁵ For a similar point, see Mooradian (1992: 79n1).

experiencing neither pleasure nor pain cannot be a pleasure.²¹⁶ If we understand the argument along these lines, Socrates again merely points out a clash in his interlocutor's belief set: Glaucon's take on the fallibility cases suggests that the calm can be a pleasure, but he *also* believes that something can only be a pleasure if it is a κίνησις, and that the calm is not a κίνησις.

But if this is all there is to Socrates's response to Glaucon's infallibilism, it would render the Deceptive Pleasure Argument rather underwhelming: the existence of false pleasures is derived from merely stipulated theses about human affective psychology which lack further argumentative support. More worryingly, perhaps, such a reading would make the argument vulnerable to argumentative pressure. For what Socrates sees as consequences of his affective psychology, Glaucon might see—or perhaps should have seen—as counter-examples. Rather than accepting that Socrates's affective psychology entails that we are hedonically fallible, Glaucon could push back and argue that Socrates's affective psychology simply fails to carve our affective nature at its joints *precisely because* it entails that we are hedonically fallible.²¹⁷

In sharp contrast to this critical approach, I believe we can read part [b]—and what follows—in a more promising, more substantive, and philosophically more interesting way. The proposal that pleasure is a κίνησις is not just a dogmatic assumption, I will suggest, it is derived from a sophisticated and powerful theory of pleasure—albeit a theory that is still in an embryonic state. The κίνησις passage gives us more to work with than it seems, then, although it requires some effort to extract the relevant information from the text. Additionally, I want to argue that this more detailed account of pleasure sheds

²¹⁶ For such a reading of the argument, see Fletcher (2018-a: 27–28).

²¹⁷ For this concern, see e.g. Gosling and Taylor (1982: 450). Thanks to Roger Crisp for pushing me to address this worry.

important light on various aspects of the Deceptive Pleasure Argument. More specifically, it does not just complement the previous line of thinking, it also lays the groundwork for a more substantive, more cogent argument against Glaucon's infallibilism.²¹⁸

At the heart of part [b] lies Socrates's claim that pleasure is a κίνησις: a process (of movement or change) rather than a state (of calm). This is not just what makes a pleasure a pleasure, it also sets pleasure and affective neutrality sharply apart: the latter is not a process but a state of affectively neutral mental calm and freedom from affective disturbances such as pleasure and pain. The claim that pleasures are—at least partially—picked out by their identity as processes rather than states must find its origin in Plato's restorative theory of pleasure.²¹⁹

Throughout the corpus, Plato puts forward more or less developed formulations of such as restorative model of pleasure, according to which pleasure, somehow, involves a process of restoration or replenishment.²²⁰ In the topological analogy, for instance, Socrates seemingly equates 'filling' or 'fulfilment (πλήρωσις)' and 'pleasure (ἡδονή, 585a3),' at another point in the argument he speaks of pleasure as a 'filling (πλήρωσις, 585b1)' of an

²¹⁸ Note that the passage is introduced by the particle καὶ μὴν: this particle signals that the speaker adds information (καὶ) and that they vouch for the correctness or relevance of this added information, even if their interlocutor may not expect it (μὴν). Generally, its use is progressive—as Denniston (1954: 351–52) writes, 'καὶ μὴν often introduces a new argument, a new item in a series, or a new point of any kind'—but Plato also uses καὶ μὴν to add missing premises to an argument. (353) I thank Gerry Wakker for significant help with these issues.

²¹⁹ Indeed, Socrates's affective psychology as a whole—including the denial that calm can be a pleasure—seems to issue from an underlying restorative theory of pleasure. This point is well made by Mooradian (1992: 79 n1).

²²⁰ 'Somehow involves' is deliberately vague. For more thorough discussion, see Frede (1992), Wolfsdorf (2013-b), and Evans (2007-a).

‘emptiness (κένωσις, 585b9),’ and slightly later, Socrates maintains that ‘being filled with what is appropriate to nature is pleasant (τὸ πληροῦσθαι τῶν φύσει προσηκόντων ἡδύ ἐστι, 585d11).’

Although these formulations fall short of being careful definitions of pleasure—strictly speaking, it looks as if they just pick out one thing among others (getting (ful)filled) that happens to be pleasant—the context, along with the corpus as a whole, indicates that Socrates is trying to give conditions, or at least a necessary condition, on something’s counting as a pleasure. For something to count as a pleasure, it takes—at least—a restorative process. In addition, there are some indications in the text that Plato also thinks that, for something to count as a pleasure, this restorative process must be (quasi-)perceptually registered and enter awareness.²²¹

In Aristotle’s *Rhetoric*, we find a similar definition of pleasure which combines these two elements and usefully illustrates the Platonic idea that is doing most of the heavy lifting here:

[T2.6] We may lay it down that pleasure is a κίνησις, a settling-down by which the soul as a whole is perceptibly brought into its natural state of being; and that pain is the opposite (τὴν ἡδονὴν κίνησιν τινα τῆς ψυχῆς καὶ κατάστασιν ἀθρόαν καὶ αἰσθητὴν εἰς τὴν ὑπάρχουσαν φύσιν, λύπην δὲ τοῦναντίον).²²² (*Rhetoric* 1.11, 1369b33–35)

²²¹ See Erginel (2019: 99) for this ‘perception condition’ in *Republic* 9. Although this requirement receives a more sustained treatment in the *Philebus*, we find textual support in *Republic* 9 as well: for one, pleasure is described as a κίνησις *in* or *of* the soul, for another, while speaking of bodily pleasures, Socrates points out that these pleasures ‘stretch out to the soul through the body (ἄι διὰ τοῦ σώματος ἐπὶ τὴν ψυχὴν τείνουσαι, 584c3–4).’

²²² Cf. Aristotle’s summary of Plato’s take on pleasure in the *Nicomachean Ethics* (1152b13): ‘[on the Platonic account,] every pleasure is a registered/conscious/perceived γένεσις toward [the subject’s] nature (πᾶσα ἡδονὴ γένεσις ἐστὶν εἰς φύσιν αἰσθητή).’

In sum, Plato treats pleasure as a perceptually registered process of restoring a previously disrupted healthy, natural, and harmonious state.

Critics have been quick to write off Plato's restorative theory of pleasure. This is not the place to give a full-fledged defense of this model,²²³ but two closely related observations are in order: one about the nature of this model, another about its scope. Plato's detractors often complain that his restorative model of pleasure is merely in a position to account for crude physical pleasures—the pleasure of eating when hungry, for instance, which is quite literally a filling of an emptiness—although it cannot make sense of more elevated pleasures such as the pleasure of listening to Beethoven's brilliant late string quartets, petting your cat, or learning a new language.

This problem seems to subside, though, if we ascribe a more subtle version of this model of pleasure to Plato. The sharp insight he is trying to capture here, I argue, is that pleasure is awareness of processes of *fulfilling diverse needs*.²²⁴ That is to say, pleasure indicates, signals, tracks, or represents something else—namely some kind of relative *improvement* such as the restoration of a disrupted balance or harmony, a return to our homeostatic state, the satisfaction of our desires, an increase in our net level of welfare, or the fulfilment of our physical or psychological needs.²²⁵

The affective calm, by contrast, can be understood as a stable,

²²³ For some recent attempts to rehabilitate Plato's restorative theory of pleasure, see e.g. Aufderheide (forthcoming), Carpenter (2011), Vogt (forthcoming), van Zoonen (2021)/chapter 4 of this thesis.

²²⁴ Katz (2005).

²²⁵ I am following Katz (2005) here who calls the Platonic view an 'improvement indicator view.' For similar contemporary representational theories of pleasure, see e.g. Millgram (2000), Schroeder (2001), and Schroeder (2004: 71–106).

homeostatic state and a neutral midpoint on the continuum in between pleasure and pain where pain should be understood as a deviation from this neutral middle state and pleasure as a return to it.²²⁶ As such, it makes sense that the state of calm was defined, more or less privatively, as the *absence* of psychological turbulence—such as pleasures, pains, and other πάθη reaching the mind—or ‘freedom from that sort of thing (τὴν ἡσυχίαν τοῦ τοιούτου, 583d8).’

Getting this restorative model of pleasure in view is important for a proper understanding of at least three aspects of the Deceptive Pleasure Argument. First of all, the restorative model of pleasure sheds new light on the idea that pleasure and calm are two different things—the basic insight that was driving the first sub-argument against Glaucon. As Plato’s model of pleasure suggests, the former precisely consists in a reparative process during which the latter is restored.

This is not just what sets the two apart, it also explains why the homeostatic state of calm cannot turn out to be a pleasure. Given the fact that κινήσεις and ἡσυχίαι are essentially different entities with separate and independent natures—an identity claim that is rooted in an elegant bit of theory and not just plucked out of thin air—it follows that an affectively neutral state of calm (a ἡσυχία rather than a κίνησις) can never turn out to be a pleasure, even if looks that way—thanks to juxtaposition with a contrastive condition—and even if gets mistaken for one.

Secondly, it is in this way that we can see why Socrates claims

²²⁶ In the *Philebus*, pleasure and pain are called ‘changes’ (μεταβολαί) and pleasure is described as a settling down, a calming, or quieting (α κατὰστασις) of the natural state (φύσις), that is to say, it is a restorative process in which a previously disrupted (natural and healthy) state (of calm or quietude) is being repaired.

that something only counts as a pleasure if it is a κίνησις—a claim which, in its turn, helps us understand the second sub-argument. Having laid down the claim that pleasure is a κίνησις, Socrates again trips up Glaucon on the basis of his dialectical commitments, but this time he supplies evidence for his claims. If you believe that pleasure is a restorative process and hence a κίνησις and if you also believe that the intermediate state is a ἡσυχία rather than a restorative process and hence *not* a κίνησις, you cannot *also* believe—on pain of contradicting yourself—that, in some contexts, the calm can turn out to be a pleasure. As before, Socrates reduces Glaucon’s thinking to absurdity by pointing out an inconsistency in his belief set:

1. The affectively neutral state of calm is not a κίνησις but a ἡσυχία.
2. If something is not a κίνησις, it is not a pleasure.
3. The affectively neutral state of calm is not a pleasure. [From 1, 2]
4. Against the backdrop of pain, the affectively neutral state of calm is pleasure. [Assumption]
5. The affectively neutral state of calm is a pleasure. [From 4]
6. Contradiction. [From 3, 5]

With his two sub-arguments in place, both of which lean on the claim that pleasure is a restorative process and thus a κίνησις, Socrates concludes that ‘it cannot be right to think that the absence of pain is pleasant (τὸ μὴ ἀλγεῖν ἢ δὴ ἡγεῖσθαι, 584a4–6).’

Thirdly and lastly, appreciating that pleasures are—at least partially—picked out by their identity as a restorative process can also help us come to grips with Plato’s deeply counter-intuitive proposal

that there are such things as deceptive pleasures or hedonic illusions. Spelling this out in more detail is the project of the next section.

3. Deceptive Pleasures or Hedonic Illusions

In part [d] of [T2.5], Socrates offers a substantive account—some kind of error theory—of why agents mistakenly report that they are experiencing pleasure when they actually find themselves in a state of affective neutrality, even though the two are importantly different from each other—as part [a] and [b] had suggested. Whereas Glaucon had cautiously advanced the proposal that these people do in fact experience genuine pleasure—‘at that moment, this, the calm, perhaps turns out to be pleasant (τοῦτο ... τότε ἡδὺ ἴσως καὶ ἀγαπητὸν γίγνεται, ἡσυχία, 583d10–11)’—Socrates rejects his interlocutor’s attempt to defuse these fallibility cases and responds that ‘at that moment, this, the calm, merely *appears* to be pleasant’ without ‘really *being* so (οὐκ ἔστιν τοῦτο ἀλλὰ φαίνεται παρὰ τὸ ἀλγεινὸν ἡδὺ ... τότε ἡ ἡσυχία, 584a7–10).’

Socrates suggests, in other words, that the victims of this hedonic mistake go wrong in their self-ascriptions of pleasure for the simple reason that the calm *seems* to be pleasant without really *being* pleasant.²²⁷ Their pleasure is merely apparent—what presents itself as

²²⁷ Although the verb φαίνεσθαι is notoriously ambiguous between a positive and a negative use, the juxtaposition with εἶναι makes it clear that we are dealing here with unreliable, deceptive appearances rather than a manifestation of reality. For a similar antithesis, cf. *Gorgias* 464a3–4 (‘δοκοῦσαν μὲν εὐεξίαν, οὖσαν δ’ οὐ’), *Republic* 596e4 (‘φαινόμενα, οὐ μέντοι ὄντα γέ που τῆ ἀληθεία’), and 599a3 (‘φαντάσματα ἀλλ’ οὐκ ὄντα’). That the distinction between seeming and being is operative here is corroborated by the fact that Socrates dismisses ordinary pleasure as ‘trickery (γοητεία)’ and sets its ‘unreliable, illusory appearances (οὐδὲν ὑγιὲς τούτων τῶν

a pleasure is not the real thing—and this means that we are dealing here with deceptive pleasures or hedonic illusions, situations where a pleasure’s appearance and its reality come radically apart.

From a Platonic perspective, drawing a connection between pleasure and illusion comes naturally—as I intimated before. In the *Protagoras*, the *Gorgias*, and especially book 10 of the *Republic*, Socrates links our susceptibility to illusions and our appetitive desire for pleasure and locates them in a cognitively impaired part of the soul.²²⁸ As Moss and others have shown, this lowest part of the soul is stuck at the level of εἰκασία which basically means that it is out of touch with reality. Because it only exercises the lowest level of cognition and is unable to avail itself of the rational power of λογισμός, this psychological part is confined to the use of unreliable images which is to say that it is unable to distinguish mere appearances, imitations, or faint adumbrations from what counts as a manifestation of reality.²²⁹

As Socrates explains in *Republic* 10, ‘*trompe l’oeil* shadow-painting, conjuring, and other forms of trickery have powers that are little short of magical (ἡ σκιαγραφία ἐπιθεμένη γοητείας οὐδὲν ἀπολείπει, καὶ ἡ θαυματοποιία καὶ αἱ ἄλλαι πολλὰ τοιαῦται μηχαναί, 602d2–4).’ Exploiting the fact that most human beings have the tendency to naively accept appearances without scrutinizing them by means of their rational powers—‘a weakness in our nature (ἡμῶν τῶ παθήματι τῆς φύσεως, 602d2)’—these forms of deception lure us into mistaking misleading appearances for what is real, thus causing ‘error[s] (πλάνην)’ and ‘all kinds of confusions in our souls (παῖσά τις ταραχὴ δῆλη, 602d1–2).’

φαντασμάτων, 584a9)’ sharply apart from ‘the truth about’ or ‘the reality of pleasure (ἡδονῆς ἀλήθειαν, 584a9).’

²²⁸ Moss (2006: 517).

²²⁹ For discussion of εἰκασία, see e.g. Lorenz (2006), Moss (2008), and especially Storey (2020).

Although a handful of commentators have appreciated the fact that deception plays a pivotal role in the Defective Pleasure Argument, these scholars typically point to pleasure's tendency to distort our view of the good and mislead us about what matters.²³⁰ On this approach, the 'confusion in our soul' caused by pleasure is an evaluative or ethical error: pleasure tricks us by making things *seem* good which are not *really* good. Moss, for instance, offers the following reading of Plato's take on the connection between pleasure and illusion in the *Republic* which is mainly informed by material coming from book 4 and 10:

Plato concludes in the *Republic* [that] desires for pleasure are very deficient as desires for good. Illusion-bound desires, like those of the appetitive soul, are desires for what *appears* good, but these desires ignore the agent's more reliable thought about what really is good.²³¹

On Moss's account, then, the hedonic trickery is located in one's *desires* for pleasure—rather than one's pleasures themselves—and these deceptive desires mislead us about some object beyond ourselves by depicting it as pleasant and thus good (at least for those who are unable to distinguish appearances from reality), although it is in fact not really good. While in the grip of an appetitive desire for pleasure, for instance, a second helping of cake might *seem* good—merely because it *seems* pleasant—although an all-things-considered rational judgment would be able to pierce through this evaluative illusion and tell me that it would be better to abstain (in light of the fact, say, that I am trying to lose weight).

This evaluative reading fits book 4 and book 10 of the *Republic* with no rough edges, but it is clear that book 9's Deceptive Pleasure

²³⁰ See e.g. Moss (2006), Clerk-Shaw (2015: 143–170), and Russell (2005: 120–135).

²³¹ Moss (2006: 532).

Argument is describing another, distinct type of illusion. As I have already argued, Socrates is examining situations where the experiences *themselves* are not what they seem: the point here is not that some pleasures mislead us about what matters, the point is that what passes itself off as a pleasure might, in fact, not be a pleasure at all but merely a deficient imitation or insubstantial simulacrum of what really counts as a pleasure.

Indeed, these cases must constitute the limiting case of hedonic illusions. When people falsely believe they are experiencing pleasure, appearance and reality come apart completely and the victims of this error find themselves behind ‘a falsifying veil’ of illusion which conceals a significant part of their world.²³² It is not just the case that they are deceived about the evaluative status, the size, the intensity, or the intentional object of their pleasure—types of hedonic illusion that receive extensive discussion elsewhere in the Platonic corpus—they are deeply mistaken about their very own affective make-up and the very affective state they take themselves to be in. Although they think they are in a state of pleasure, they actually find themselves in a state of calm which gets mistaken for a state of pleasure thanks to a contrast effect with a state of pain. That is to say, a non-pleasure successfully wears the mask of pleasure and manages to get itself mistaken for one.²³³

To see how the restorative model of pleasure I identified above can elucidate Plato’s proposal that a pleasure’s reality and its appearance can come apart in this radical way, we might start by recalling that Glaucon and most modern infallibilist critics reject

²³² Murdoch (1972: 84).

²³³ This of course suggests that there is a trivial sense in which my reading also counts as evaluative: if we grant that it is good to experience pleasure and that it is indifferent to experience neither pleasure nor pain, cases where affective neutrality deceptively appears to be pleasant also trick us about what matters.

hedonic fallibilism on the grounds that pleasure is nothing more than a chunk of seeming picked out by its immediate phenomenological quality to which we have privileged, infallible epistemic access.²³⁴

On the view lurking in the background of the Deceptive Pleasure Argument, in sharp contrast, there is more to pleasure than meets the introspective eye: Plato suggests that pleasure is more than a feeling and not just a blind surge affect whose reality is exhausted by its immediate felt quality. More precisely, the argument seems to treat pleasure as a complex consisting of two aspects or ‘faces’: there is an internal, subjective, or phenomenal face—the introspectively available features of a pleasure—and there is an external, objective face—the underlying restorative process of having one’s needs fulfilled. Both of these conditions—both faces of pleasure—state different, individually necessary, and perhaps jointly sufficient conditions of pleasure.²³⁵

Crucially, though, the internal aspect—the felt quality or what Socrates calls the ‘appearance’ or φάντασμα—is not sufficient for something to count as a genuine pleasure. In the absence of an underlying restorative process—what I called the external face—we are merely dealing with a deceptive, illusory simulacrum of pleasure which is nothing more than a trick of the mind and a deficient

²³⁴ As D. Lewis (1980: 222) puts it, ‘a theory of what it is for a state to be a pain is inescapably a theory of what it is like to be in that state, of the phenomenal character of that state.’ And according to Kripke (1980: 153), ‘pain ... is not picked out by one of its accidental properties; rather it is picked out by its immediate phenomenological quality.’

²³⁵ In a passage in the *Philebus* (37a1–10), which I will discuss in more depth in chapter 3, Socrates makes a similar yet slightly different point about the nature of pleasure when he suggests that pleasure is a unitary complex consisting of (i) the *act* of taking pleasure (ἡδεσθαί) (the internal face) and (ii) the *content* of that act or that in which one takes pleasure (τό γε ᾧ τὸ ἡδόμενον ἡδεται) (e.g. the fact that I won the lottery) (the external face).

imitation of the real thing. For an affective experience to count as a real pleasure instead, something over and beyond an experience's subjective, raw feel needs to obtain.

More precisely speaking, the appearance itself has to be properly brought about by some kind of underlying restorative process. In order for us to experience genuine pleasure rather than a flimsy, insubstantial semblance of pleasure, it must not just *feel* as if our needs are being met—this is the internal face of pleasure which explains the intuitions of the infallibilist—our needs must also really *be* met—this is the external face of pleasure and Plato's important addition to the infallibilist's phenomenal model of pleasure.

So on the Platonic assumption that pleasure is a veridical perception of a restorative process of return to some healthy, natural, harmonious state, it follows that something only counts as a pleasure if it is a veridical appearance of having one's needs met. In the absence of the underlying process of having one's needs met and being restored to a homeostatic state, the pleasure we think we are experiencing is actually a defective and deceptive simulacrum of the real thing.

This suggests that cases of hedonic illusion are very similar to cases of perceptual illusion where, roughly speaking, some *x* appears other than it really is.²³⁶ Just like a straight stick might look bent while

²³⁶ That this is the right way of understanding this bit of text is corroborated by the fact that Plato compares these defective and deceptive pleasures with σκιαγραφία paintings. Σκιαγραφία—literally 'shadow-painting'—was an ancient painting technique, developed by Apollodorus in the last quarter of the fifth century BC, which was able to create a realistic effect of depth and three-dimensionality by juxtaposing contrasting colors comparable to contemporary *trompe l'oeil* or *chiaroscuro*. Cf. Brill's *Der Neue Pauly* (s.v.), Keuls (1975), and Keuls (1978: chapter 4). This illusionist painting technique—as Plato points out in book 10—exploits the inferior part of our nature and our cognitive and sensory blind spots by juxtaposing different shades of color to create an illusion of depth and three-dimensionality. Similarly, in the hedonic case, the

submerged in water and just like something concave might look convex—to use Plato’s own examples (602c10–d1)—the affectively indifferent calm might appear other than it really is: if it gets juxtaposed with pain, it seems to be pleasant or looks like a pleasure, although it is not.

In the hedonic case, the juxtaposition of an affectively negative state of pain and an affectively indifferent state of calm creates the illusion of there really being an underlying restorative process and thus an affectively positive state of pleasure, although, in reality, there is just the neutral state of psychological calm temporarily masquerading as pleasure and tricking its victims into self-ascribing a pleasure that is not really there. The analogy with perceptual illusion elucidates Plato’s thinking. The mere fact that a stick *looks* bent while submerged in water does not mean it really *is* bent, the mere fact that lukewarm water *feels* cold when contrasted with hot water does not mean it really *is* cold, and the mere fact that—in the Checker shadow illusion—two tiles of the exact same darkness *seem* to be of different color when they are partly shadowed by another object does not mean they really *are* of a different color. In the same way, or so Plato argues, the mere fact that the state of calm *seems* pleasant when juxtaposed with pain does not mean it really *is* pleasant.

Note that Plato can only claim this, as I have argued, because he believes that a pleasure is more than a feeling. In addition to the feeling aspect, an underlying restorative process which guarantees that our needs are really met is required: in the absence of such an objective correlate out there in the world, one can go wrong in self-

juxtaposition of mental states—pain and painlessness, to be precise—exploits the same inferior part of our mind and exploits our affective blind spots to create an illusion of pleasure. In both cases, the deceptive simulacrum seems to be superficial and lacks depth. As Socrates’s own metaphor has it (587d6), deceptive pleasures are like a two-dimensional projection of a solid, three-dimensional object.

ascribing a pleasure—merely in virtue of the fact that there is a felt appearance as of there being such an underlying restorative process.

However, as Socrates concludes in part [d], these hedonic appearances are misleading: they are unreliable and untrustworthy, they do not track the truth about pleasure, and they involve some kind of trickery.²³⁷ In cases of hedonic fallibility, the appearance is there—it strikes the subject as if they are being restored, there is something it is like to experience a misleading pleasure, it really feels good—although the actual restorative process is lacking and that is why we are dealing with a deceptive rather than a truthful pleasure.

4. Conclusion

The Deceptive Pleasure Argument in book 9 of Plato's *Republic* is often dismissed scornfully. One of the argument's most problematic features is its defence of hedonic fallibilism—the counter-intuitive thesis that we can go wrong in our immediate self-ascriptions of pleasure. Because pleasures are neither more nor less than a bit of felt experience, Plato's detractors argue, it is hard to see how the introspectively based belief that I am currently experiencing pleasure could be mistaken.

Driven by considerations of this sort, a large group of interpretations rejects this Platonic argument out of hand—even though it is meant to constitute the 'greatest and most decisive' strike against Thrasymachus's project—whereas another group of interpretations tries to save Plato from himself by developing textually ungrounded readings on which the argument is defending an entirely

²³⁷ Dickie (2001: 44) notes that Plato often uses the analogy of γοητεία (magic, trickery, witchcraft) to describe 'the creation of illusion by making objects appear to be present that are not really there.'

different yet more palatable thesis.

In this chapter, I have tried to find an alternative interpretation. The reading I have put forward here is based on the most natural and most straightforward reading of the text but it also strives to *make sense* of Plato's hedonic fallibilism. The key to unlocking this puzzling position, I have suggested, lies in appreciating the fact that Plato holds that pleasures are not exhausted by their phenomenology. For something to count as a pleasure, it takes something over and above and external to the feeling of the relevant experience to turn that experience into a genuine pleasure. More precisely speaking, in addition to the phenomenal aspect—the pleasurable feel of the experience—it takes an underlying restorative process of having one of our needs fulfilled. By tricking us into adopting the false belief that our needs are met when this not really the case, these deceptive pleasures or hedonic illusions misinform us about some way the world really is and this is what makes them defective, deceptive, merely apparent, or what Plato calls false or untrue.

This theory of pleasure sheds light on the idea that there are such things as deceptive pleasures. Because any representation can misrepresent and any appearance can fail to latch onto reality, Plato seems to believe that there is nothing crazy or counter-intuitive about the idea that there can be errors in our hedonic tone. When not brought about in the usual way—that is, by a genuine underlying restorative process in which one of our needs is satisfied—our so-called pleasures are not what they pretend to be.

If there is the felt appearance of having one's needs fulfilled without the underlying restorative process, we are dealing with a case of deceptive pleasure or hedonic illusion akin to perceptual illusion. It *seems* as if we are experiencing pleasure—as if we undergo some relative improvement—although we are not. In such a situation, as Plato has it, our putative pleasure is a mere trick, εἴδωλον, or *trompe*

l'oeil painting: an insubstantial, deficient, specious, false simulacrum of the real deal which misinform us about some way the world really is.²³⁸ Importantly, the impoverished status of these pseudo-pleasures is not introspectively accessible: like any other illusion, a deceptive pleasure cannot be directly seen for what it is—and this adds to the tragic foolishness of those who are taken in by these false and deceptive pleasures.²³⁹

This interpretive option is attractive and deserves more scholarly attention than it has received. If my reading gets things right, Plato's hedonic fallibilism is not just more intelligible and more cogent than usually thought, it also fits seamlessly into the over-arching strategy of the dialogue as a whole. For unlike the first demonstration of *Republic* 9, and like the arguments of book 1,²⁴⁰ the Deceptive Pleasure Argument constitutes an immanent critique of the Thrasymachus's project.²⁴¹ Like 'the majority' of people—who treat their pleasures as 'important things (μεγάλων τινῶν)' because they believe that without pleasure one is 'hardly living at all (οὐδὲ ζῶντες),' let alone 'living well (εὖ ζῶντες, 329a7–8)'—Thrasymachus seems at least implicitly driven by a tacit commitment to hedonism.²⁴²

With its defence of hedonic fallibility, the argument meets Thrasymachus on his own ground: it shows that the good-making features of his favourite, pleonectic, tyrannical way of life—its

²³⁸ My thinking here is indebted to Schroeder (2001) and Schroeder (2007).

²³⁹ This point is well made by Whiting (2014: 36).

²⁴⁰ Nawar (2018).

²⁴¹ For the plausible suggestion that the third demonstration of book 9 is supposed to bite with an (implicit) hedonist like Thrasymachus, see e.g. Frede (1985: 157–58), Gosling and Taylor (1982: 101), and White (1979: 233–34).

²⁴² For an excellent defence of the idea that Socrates's interlocutors are attracted to the tyrannical way of life because of their (implicit) hedonism, see Moss (2005). Johnstone (2015) defends the related claim that Plato's tyrant is primarily concerned with the pursuit of bodily pleasure.

pleasures—are not what they seem. Crudely speaking, because Thrasymachus mainly seems to care about pleasure, he should be bothered by the possibility that his pleasures are merely apparent rather than genuine.²⁴³

Note that even if Thrasymachus and other implicit or explicit hedonists remain unmoved by this argument—for instance, because they refuse to accept the idea that pleasures are picked out by an underlying restorative process rather than the phenomenal quality of the experience—the later parts of *Republic* 9 will tighten the screw. There, in a dense and vexed passage, Plato seemingly suggests that the φρόνιμος's ecstatic pleasures of 'the soul itself by itself (τὴν τῆς ψυχῆς ἡδονὴν αὐτῆς καθ' αὐτήν, 485d11)' are more pleasant—and thus phenomenally different—from non-ideal pleasures. When the soul transcends the gravitational field of the material plane to cognize the Forms, 'getting near what really is and having intercourse with it (πλησιάσας καὶ μιγείς τῷ ὄντι ὄντως 490b4–5),' this somehow involves a supreme restoration and hence a supreme pleasure—no less than 729 times more pleasant than the sketchy pleasures of ordinary people and tyrants.

That the interpretation I defend here is in a position to explain why Socrates calls the argument the 'strongest and most decisive' of book 9's attempts to defuse Thrasymachus's challenge is an important second benefit of my reading, besides the fact that it can respect the text as it stands. There is a third and final advantage. On the reading defended here, *Republic* 9 develops a historically influential idea that

²⁴³ Interestingly, this suggests that the Deceptive Pleasure Argument is a mirror image of the jellyfish *elenchus* from the *Philebus*: in that argument, Socrates makes trouble for hedonism by forcing his interlocutor to admit that it would be bad to have a pleasure without the higher-order, self-ascribing belief *that* one is experiencing pleasure; here, Socrates pursues a reverse strategy by arguing that it would be bad to have the belief *that* one is experiencing pleasure without the underlying pleasure.

is not just interesting but also worth taking seriously. Plato is basically suggesting that our pleasures are a form of openness to the world: when things go well—when our pleasures are true rather than false, as Plato would have it—they do not cast a veil between us and reality but supply us with a fragment of the world.²⁴⁴ Indeed, I believe we are dealing here with the first attempt in Western philosophy to formulate the powerful idea that, like our beliefs, our pleasures are charged with ‘the critical task of getting things right.’²⁴⁵

²⁴⁴ See McDowell (1994: 111), de Sousa (1987), and Tappolet (2016: 29) for contemporary iterations of a similar idea.

²⁴⁵ Following Evans (2008-a: 123), who detects this proposal in Plato’s *Philebus*. Portions of this chapter, in different stages of development, have been presented in Oxford (*Oxford Graduate Conference*) (special thanks to Roger Crisp for his astute comments and further discussion), Austin, Texas (online) (*Annual Ancient Philosophy Workshop*), Erlangen-Nuremburg (online) (*Lust und Unlust in der antiken Philosophie*) (special thanks to Dorothea Frede who kindly offered to read a whole draft of this chapter), and at two occasions in Groningen. I thank the audiences at these different occasions for their valuable input.