Topoi (2013) 32:125–128
DOI 10.1007/s11245-013-9151-5

Plato's *Philebus*

Ronald de Sousa

Published online: 2 March 2013
© Springer Science+Business Media Dordrecht 2013

Philebus is a work that ostensibly addresses the question of the rightful role of pleasure in the good life, particularly in its relative importance when compared to knowledge and reason. Its striking originality, as I shall try to explain, lies in a series of arguments made within it for the conclusion that some pleasures can be false.

The way in which it does this may at first inspire resistance, for it may appear excessively subservient to two fashionable trends—albeit the two trends might be thought to exist in mutual tension. One is the tendency—sometimes decried as "scientism"—to envisage aspects of consciousness from a reductionist, biological point of view, thus minimizing the specifically subjective essence of consciousness. The other somewhat modish view that appears to be espoused in this work is that every state of consciousness can somehow be construed as cognitive. The first tendency is manifest in the extension of the reference of terms designating states of consciousness to bodily processes of which we are wholly unaware. The second takes the form of arguing that pleasure, a state of consciousness ostensibly characterized entirely in terms of its intrinsic quality, is actually a representational state that can be correct or incorrect, "true" or "false". As we shall see, however, once we allow for certain excesses, the author makes a surprisingly convincing argument for both of these views, as well as for their compatibility.

The work is offered in the form of a dialogue—in that respect, at least, it is not "trendy"—between Socrates and an interlocutor called Protarchus. The latter defends a position attributed to one Philebus, a curiously ghostly character who unaccountably "leaves the field" on the first page, intervening again only three or four times to reiterate a simple-minded praise of pleasure as the supreme good. That is not the only one of this work's quirks of style. Its organization is somewhat confusing. Its central pages contain an analysis and classification of pleasures, in the course of which Socrates attempts to persuade his interlocutor that some pleasures are false, but little is done in the dialogue to relate this claim to the work's announced topic. One is left to assume, I suppose, that falsity might detract from the claim of any conditions to be life's chief good. In this review, I shall not attempt to deal with the somewhat messy structure of the whole; neither shall I attempt to canvas all the topics that come up only to be desultorily dropped. I shall concentrate instead on the arguments adduced for the claim that pleasures can be false.

The groundwork is laid with schematic sketches of the nature of sensation, memory, pleasure, and desire (31–36). The order in which these are taken up is somewhat involved: sensation gets defined because it is used in the definiens of memory, which is needed for the definition of desire, which is involved in a certain class of pleasures in which in turn memory also plays a part. Central cases of *sensation* are said to be those in which some change or disturbance affects both body and mind; but the majority of physiological disturbances never reach consciousness. Since this means that most of the processes in which sensation consists are actually hidden from awareness, one might dub this the "iceberg" theory of sensation. *Memory* is the preservation of sensation: but what it preserves is, to continue the analogy, only the tip. (On this point, despite Plato's scientific bent, he has neglected much evidence, from psychological experiments with "priming", showing that memories capable of influencing thought and behavior

R. de Sousa (✉)
Department of Philosophy, University of Toronto, 170 St George St., #424, Toronto, ON M5R 2M8, Canada
e-mail: sousa@chass.utoronto.ca

do get preserved even when a "sensation" never actually reaches awareness. See e.g., Bargh et al. 1996.)

In conformity with this essentially biological point of view, pleasure is then defined as a bodily process, without essential reference to consciousness. That process cannot, however, be defined in itself, but only in tandem with pain; for the two respectively correspond to deviations from and restorations of the body's natural condition of homeostasis.

But what sort of "correspondence" is this? A first formulation implies mere correlation—saying merely that bodily pleasure and pain occur under certain conditions. This falls short of a definition. But a few lines later it is rephrased as an identity: "destruction [of homeostasis] is pain, but the way back again is pleasure" (32). The passage from the former to the latter formulation, the reader may well feel, is where the rabbit gets into the hat. For in the stronger formulation, but not the weaker, pleasure and pain are species of sensation: a bodily process strong enough to penetrate to the soul just *is* either a "disruption" or a "restoration".

This seems to present a dilemma: either such processes are merely *occasions* of pleasure and pain, and we still don't know what pleasure and pain *are*; or else pleasures and pains consist in the bodily processes themselves. In the latter case, how can they be any more true or false than any other bare physiological process? But on the former interpretation, one still hasn't been told what pleasure and pain actually *are*; and in the absence of that information we have no way of assessing whether it is appropriate to speak of them as being true or false. On the "iceberg" model, the question arises for sensation generally: does sensation consist in just the tip, as suggested by the talk of memory as preservation of sensation (presumably only the tip can be preserved), or is it constituted by the whole process of disturbance or restoration?

The solution, I believe, is one that Plato here borrows from a number of other advocates of "naturalizing" knowledge, notably Fred Dretske (1995) and Jesse Prinz (2004). If the tip of a physiological change that reaches consciousness can be construed as having an indicative function as a result of natural selection (whether or not it is itself part of the same causal process), it makes sense to think of it as cognitive. The underlying notion of function at work here is the aetiological one, made familiar by a number of writers in the last half-century and refined by Ruth Millikan (1984). In this perspective, it can be correct to say *both* that the consciousness of pleasure is identical with (part of) the physiological process involved, *and* that the former "means" or refers to the latter. As such, that conscious experience can mislead. That is the basic strategy that allows Plato to argue that pleasures can be true or false.

But what is it that pleasure might be indicating or referring to? This will become clearer once we introduce one more mental states that gets defined in terms of those already posited. *Desire* is introduced in connection with a class of pleasures that "belong to the soul alone" and "always involve memory" (33). Thirst, for example, is said to be "desire for drink", or more precisely: "for replenishment by drink" (35). Desire in general, on Plato's view, always involves two contrasting states: an apprehension of the actual state of bodily depletion, and a simultaneous memory of replenishment. (This analysis is plausible, so long as we do not ask about the very first desire in a person's life: must there not have been a first instance of thirst or hunger before the formation of any memories? But perhaps what happens in the mind of an infant when it first suckles at its mother's breast does not yet quite merit the name of desire.)

Let us grant, then, that Plato has established a minimal sense in which states of consciousness experienced as pleasant might be construed as representational. That still doesn't explain exactly how such states can be false. One problem is that there are alternative paradigms for the construal of 'true' and 'false'. Some uses of 'true' and 'false' are unrelated to cognition: 'a true friend', for example; a 'false bottom' is not a bottom that refers to a non-fact, but simply something that is *not a bottom at all*; a 'false front' is a real front that misleads, but doesn't refer to anything. In the work under review Protarchus asserts that only propositions can be literally true and false. But in the cases under consideration the only available proposition is *I am experiencing pleasure,* and surely about that I could never be wrong: "neither in dreams nor waking, not even a crazy person out of his mind would ever think he is enjoying, but not be enjoying at all" (36). Socrates concedes this point, but only momentarily, as we shall see.

First, however, Plato brings up two other types of false pleasure: false pleasures of anticipation and pleasures distorted by comparisons. Pleasures of anticipation involve judgments about the future. If Plato is to establish that the pleasure has as much claim as the judgment to be called true or false intrinsically, rather than derivatively, he can allow (1) that the pleasure depends on the judgment, but must show (2) that the pleasure's relation to its object is the same as, or at least parallel to, the relation of the judgment to its object.

To establish (1), we need to distinguish taking pleasure in something *about which* I happen to have a certain belief, from taking pleasure *in* a belief. Bernard Williams (1959, 66) gives this example: I may take pleasure in the contemplation of a picture; though I believe it to be a Giorgione, a reattribution would not spoil my pleasure. This is taking pleasure in something about which I have a certain belief. But if I am taking pleasure in the (supposed) fact that it is a Giorgione, then my pleasure would not survive disillusionment. Only in this second case does the pleasure

"depend on" the belief. Now in order to show (2), Plato makes use of an intriguing simile to explain the analogy between belief and pleasure. It is, he says, as if the scribe in my head that records the judgment were to hand it over to a painter whose job it was to depict its content; that depiction, and not the judgment itself, is the direct object of the pleasure.

In their role as objects, however, the painter's images cannot be *mere* images: for images cannot differentiate anticipation from fantasy or memory. It is only in so far as the representation is linked to a *belief*, that it can be regarded as false. But then Protarchus can still claim that truth value attaches uniquely to the belief, and that the pleasure is at one remove, since "it is true or false only according to the truth or falsity of the belief with which it corresponds" (39). The work of the painter is directly an object of pleasure, just as the work of the scribe is directly an object of belief; but the former is only indirectly what the latter is directly, a truth-value bearer.

There is, however, a twist that might save Socrates' doctrine. It is to insist that the only proper truth-bearer is an abstract object: a proposition, distinct from both the scribe's inscription and the painter's image. The proposition is related to both in the same way, mediated, albeit by a picture in one case and a linguistic entity in the other. If so, the truth-value claim is no weaker in one case than it is in the other. The picture-pleasure and the inscription-belief are now exactly *on a par* with respect to their relation to the truth bearer. It does not matter that the work of the painter still *causally* depends on that of the scribe: for what matters to the parallel is the relation of each to its truth-bearing object, not to its causes. (Similarly one belief can be causally dependent on another, logically related belief. But its truth-value is not thereby "merely derivative".) The parallel has been secured by slightly *loosening* the connection between belief and truth value. This seems an independently correct move anyway, since propositions can be true or false whether or not they are believed, and whether or not they are being enjoyed.

But why, one might ask, do we need something like the "painter" in the first place?

Here again, Plato seems to be drawing on recent neuropsychology, though he persists in being economical with citation. Paul Slovic (2007) has shown that words in general are virtually powerless, as compared to images, in moving us to compassion and more generally to emotional response. That, I believe, explains the emphasis on the simile of the painter. In order to generate pleasure or pain, when those do not arise directly from a bodily state, belief is not enough: it has to be supplemented by imagination.

Now the fact that imagination is able to generate pleasure and pain serves to remind us that the relation between pleasure and belief is a reciprocal one. The causal arrow can face either way. In the cases so far considered, belief generated pleasure; false pleasures of this type, Socrates remarks, were "infected" by the falsity of a belief. But we should attend also to "the opposite" cases, where pleasure generates beliefs (42). Those cases resemble comparative illusions, such as mis-estimation of size in the context of a misleading contrast. In the well-known Muller-Lyer illusion, for example, a line looks longer or shorter depending on the arrowheads that frame it.

But the description of a pleasure as "misleading" in such cases again seems to beg the question. If the quality of a sensation is considered independently of any reference to anything else, then it won't get assessed in terms of its indicative value.

In reply, Plato reminds us that his conception of pleasure has physiological roots: it tells us about beneficial or harmful changes taking place inside our bodies. It parallels *perception*, not "mere sensation" (43). The indicated state of the body (*qua* harmful or beneficial) can be considered the objective referent of the pain. On this model, illusion is possible, no less than for any other sensations—e.g. of colour or of size—that mislead because of their contrastive setting.

One might object that cases where context affects pleasure or pain introduce not mistakes, but merely *change*. A drink on a hot day when I am really thirsty does not mistakenly seem more pleasant: it *is* more pleasant. The pleasure is what it is, not falsified but genuinely increased by the context. That is surely how we should describe things in the light of the view that the cognitive function of pleasure is to tell me what is happening in my body. That this is what Plato intends is confirmed by such passages as the one where he writes that what reaches consciousness are "big changes" (43), which it seems reasonable to interpret as meaning ones that matter for the survival and welfare of the organism, or as Plato puts it "processes concerned with the preservation or decay of living things" (25). This biological perspective is re-affirmed with the assertion that distress or pleasure that misleads "cannot be called right or useful" (37).

The same rationale is implicitly invoked to justify Plato's most paradoxical thesis. At the beginning of the dialogue Socrates had originally *granted* to Protarchus that a pleasure is no less a pleasure for being false. Now that is precisely what is being denied. The "most false" pleasures, he writes, are those where what is mistaken for a pleasure is not a mixture of pleasure and pain, but a neutral condition. Some pleasures, it is now claimed, are false not in the sense of corresponding to no fact, but in the sense of being unreal, of not existing at all. They are, we might say, not merely illusory, but hallucinatory: they are really non-pleasures.

The argument for this conclusion is conducted in terms of the paradigm of pain and pleasure familiar from the

outset: (bodily) pains originate in large disruptions of homeostasis, pleasures in its restoration. Since not all bodily processes are conscious, this allows for times of relative equilibrium when nothing is experienced. Yet these times are sometimes experienced as pleasant: people "think, at the very time that they are not feeling pain, that they are feeling pleasure" (44). This seems preposterous. For to claim that the subject is experiencing a "false pleasure" seems to entail that he is experiencing *some* sort of pleasure—which contradicts the assertion that there is no pleasure. Why not say instead that the subject is mistaken at the level of meta-cognition, in holding a false *opinion* about the nature of his current experience?

The objection is plausible. Nevertheless, I think it can be answered. We have seen that for Plato, pleasure is a form of sensing. But sensing is ambiguous: in one sense, it refers merely to a certain quality of experience; in another, it refers to *perception*. Plato could be clearer about this; but if we are willing to grant him that distinction, then we can construe the "hallucinatory" type of false pleasure as a state that fools you into thinking that it is a *perception of replenishment*. Since a perception of replenishment is alone properly called a 'pleasure', this state of consciousness seems to be but is not really a pleasure at all. Thus can a false pleasure be a non-pleasure.

A more sweeping objection might be made. In none of the three kinds of alleged false pleasures, is there any conceivable standard of objectivity. While it might be admitted that pleasures and pains are indeed informative about present or future states of the body, one might still deny that they *refer* to those states. The indicative relation is a causal relation, and therefore one of "natural meaning"; but for truth and falsehood to be in question we need to be able to ascribe "non-natural meaning" in the sense articulated by Paul Grice (1989).

From the biological point of view that underlies Plato's view, however, I have already suggested that the causal relation in question can be described as a functional one providing it exists as a result of its adaptive advantages. The applicable standard of objectivity is not hard and fast, but it relates to the overall goal that clearly underlies Plato's investigation: to find a characterization of the role of pleasure in human life most likely to enhance an agent's thriving in accordance with general facts about human nature.

In this perspective, Plato might have been better advised to speak not of truth and falsity, but of rationality and irrationality. That would enable us to make sense of a passage which is otherwise bound to strike many a reader as puzzling. I refer to the claim that a morally bad person is more likely than a morally good one to be subject to false pleasures of anticipation: "a man may often have a vision of a heap of gold, and… of himself mightily rejoicing over his good fortune"; but while "the bad, too, have pleasures painted in their fancy… theirs are false pleasures" (40). On the face of it, this is absurd. Bad men are not necessarily worse predictors than good ones, either about whether they are likely to amass a pile of gold or about whether they will enjoy it if they do. What is possible, however, is that "goodness", at least if we understand it loosely as something like "wisdom", is more likely to yield *proportionate* assessment of desirability. If we regard pleasures of anticipations as perceptions of the desirability of future states, they can be mistaken in a more interesting way than a mere error of factual forecasting*: a pleasure of anticipation* can fail of *rationality* because it is not proportionate to *anticipated pleasure*. This says nothing about how much one should enjoy anticipation: that might vary from one person to another. It requires only consistent proportionality. Pleasures are perceptions that move us to action: an intense pleasure in the anticipation of something trivial might cause us to act imprudently. Biologically, disproportionate pleasures of anticipation will be poor guides in the search for future satisfactions. A large pleasure in the contemplation of what will actually give me little pleasure, or pain in foreseeing pleasure, are likely to thwart me in my planning.

In short, the argument Plato makes for the possibility of false pleasure may not be wholly convincing if one takes the term literally. It is best taken as a suggestion about the biological role of pleasure as a guide to present health or future satisfaction, and as a comment on the rational conduct of life. Thus understood, this is a work that makes good use of some of the important ideas that philosophers, neuroscientists and psychologists have developed over the past 50 years. Although it is rather short on detail about mechanisms, *Philebus* nicely illustrates the importance for philosophers of keeping up with contemporary science.

References

Bargh JA, Chen M, Burrows L (1996) Automaticity of social behavior: direct effects of trait construct and stereotype activation on action. J Pers Social Psychol 71(2):230–244

Dretske F (1995) Naturalizing the mind. MIT Press, Cambridge

Grice P (1989) Meaning. In Studies in the way of words. Harvard University Press, Cambridge

Millikan R (1984) Language, thought, and other biological categories. MIT Press, Cambridge

Prinz J (2004) Gut feelings: a perceptual theory of emotion. Oxford University Press, New York

Slovic P (2007) When compassion fails. New Sci 194 (2598):18

Williams B (1959) Pleasure and belief. Proc Aristotelian Soc, Suppl Volume 33:57–72