What moral philosophers can learn from the history of moral concepts.

The essays in this volume belong to many disciplines: moral philosophy, intellectual history, theology and historical linguistics. They are gathered together in the conviction that these disciplines have much to learn from one another, and in particular that moral philosophy, my own discipline, has much to learn from all the others. It is this latter conviction that I want to defend in this essay, which can be regarded as a manifesto for the volume as a whole.

Moral philosophers, at least in the English speaking world, have generally taken little interest in the history of moral concepts. There are a few honourable exceptions – Bernard Williams, Alasdair MacIntyre and Charles Taylor spring to mind – but their example has by and large not been followed by the younger generation. What explains this neglect? Institutional pressures of a kind familiar across the academic world – growing specialisation, the lure of formal methods – have played a part, no doubt; but beyond this there are causes specific to moral philosophy as this has been practiced in Britain and America over the last century.

The "realists" who dominated British moral philosophy between the wars – Ross, Prichard, Ewing and Broad – regarded morality as reducible to a small set of changeless principles, knowable by direct intuition. Morality has no history. To be sure, *theorising* about morality has a history, but this must be regarded as a history of more or less erroneous descriptions of the same timeless subject matter. (By analogy, the world of numbers has no history, though its investigation by humans obviously does.) It is characteristic of this approach that Plato and Aristotle were seen as attempting to get clear about the very same concepts that preoccupy modern moral philosophers, only failing rather badly. Such smugness drove R. G. Collingwood (for one) to distraction. "It was like having a nightmare," he wrote in his autobiography,

> about a man who had got it into his head that τριήρης was the Greek for
> 'steamer', and when it was pointed out to him that descriptions of triremes in
> Greek writers were at any rate not very good descriptions of steamers, replied
> triumphantly, 'That is just what I say. These Greek philosophers … were
> terribly muddle-headed, and their theory of steamers is all wrong'.[1]

---

[1] R. G. Collingwood, *An Autobiography* (Clarendon Press, Oxford: 1939), 64.

Realism was swept aside by logical positivism in the 1930s and 40s, but its ahistoricism was perpetuated, indeed accentuated, by the new movement. Ross and Prichard had at least given some thought to the history of their subject, but the generation that followed felt at liberty to ignore this entirely. Russell and Wittgenstein had taught it that philosophy's task is to uncover the universal (even "transcendental") form of language, including moral language. The differences between our moral terms and those of Homer are no more philosophically significant than the differences between (say) inflected and uninflected languages; they merely serve to conceal an underlying sameness of logical structure. This doctrine found its monument in R. M. Hare's *The Language of Morals*, published in 1952. Note "language" in the singular, and the definite article. There is, by implication, only *one* language of morals, though its dialects may vary.

Illusions of translatability

The idea of a universal "language of morals" is on the face of it rather implausible. After all, it is notorious that several core terms of that language, including "right", "duty", "ought" and "moral" itself, have no exact counterparts in many historical languages. "If someone professes to be expounding Aristotle and talks in a modern fashion about 'moral' such and such, he must be very imperceptive if he does not constantly feel like someone whose jaws have somehow got out of alignment", wrote Elizabeth Anscombe famously.[2] "Moral" cannot be translated directly into ancient Greek. Neither can any of the other words in the list above.

The universalist will be unimpressed by this response. Concepts are not tied in a one-to-one fashion to particular words, he will insist; they can be expressed by a variety of words in combination. For instance, Russian has no one word for blue. Instead it has two words, *sinie*, "dark blue", and *goluboi*, "light blue". But this does not mean that monoglot Russian speakers lack the *concept* of blue, since anything we can say using "blue" they can say using a disjunction of *sinie* and *goluboi*. More precisely, any state of affairs in which "*x* is blue" is true will be one in which the corresponding Russian sentence "*x* is *sinie* or *goluboi*" is true, and any state of affairs in which "*x* is blue" is false will be one in which the corresponding Russian sentence "*x* is *sinie* or *goluboi*" is false. And that, on one common view of the matter, is all it means to say that Russians possess the concept of blue.

---

[2] G. E. M. Anscombe, "Modern Moral Philosophy", *Philosophy* 33/124 (Jan., 1958), 1-19, 2.

The universalist will make a similar point about moral concepts. English, like other modern European languages, draws a distinction between *shame*, which can be provoked by any kind of failure or shortcoming, of one's own or of one's associates, and *guilt*, which is specifically a response to one's own acts and omissions. Ancient Greek made no such distinction. Here, one word, *aidōs*, usually translated "shame", covered the whole territory that we divide between "shame" and "guilt". But we needn't deny ancient Greeks the *concept* of guilt, the universalist will insist, so long as they were able (as they clearly were) to pick out in words the kind of *aidōs* which looks specifically to one's own misdoings and is bound up with a fear of retribution and a desire to make amends. In short, the ancient Greeks possessed the concept of guilt *implicitly*, insofar as they had the verbal resources to spell it out circuitously. That they lacked a distinct word for it is beside the point.

The universalist thesis can now be stated more precisely: all intellectually mature humans understand the core moral concepts, either explicitly (they have special words to express them) or implicitly (they can convey them using a variety of words in combination). Either way, it does not matter. The important point is that the philosophical analysis of the core moral concepts can proceed as it has always done in the analytic tradition, without reference to the variety of human languages and cultures.

I mistrust this appeal to implicit understanding. It seems to me a mistake to say that anyone who can string together some conjunction or disjunction of terms extensionally equivalent to "guilt" possesses the concept of guilt. Words matter. To see why, consider "blue" again. Even if it is translatable *salva veritate* as "*sinie* or *goluboi*", it does not follow that anyone who understands "*sinie* or *goluboi*" thereby has the concept of blue, for to have the concept of blue is to see all shades of blue as shades of one and the same colour, which is precisely what those who think in terms of *sinie* and *goluboi* cannot do. "Blue" directs our attention to similarities that would otherwise go unnoticed, or (as Russians would doubtless have it) *sinie* and *goluboi* focus our gaze on differences that would otherwise be obscured.[3] The contrast is the familiar one between "lumpers" and "splitters". English speakers are lumpers when it comes to blue but splitters when it comes to pink and red. From the lumper's standpoint, pink just is light red, yet it is hard for us to see it that way. For one thing, pink has acquired a raft

---

[3] There is some scientific evidence that Russian and English speakers do indeed differ in the way they discriminate shades of blue. See *New Scientist Daily News*, 1 May 2007: https://www.newscientist.com/article/dn11759-russian-speakers-get-the-blues/

of symbolic meanings – think of Barbie dolls etc. – unimaginable in a linguistic culture that spoke only of "light red".

Similar considerations apply to shame and guilt. Even if "guilt" is translatable *salva veritate* as "*aidōs* of such and such a type, occasioned by actions of such and such a type", we cannot assume that anyone who grasps this latter phrase possesses the concept of guilt, for to see an emotion as a type of *aidōs* is to make of it something other than it would be if taken as a thing in its own right. "Even though some reactions in [Greek] societies were structured in the same way as our reactions of guilt", wrote Bernard Williams,

> they were not simply guilt if they were not separately recognised as such; just as shame is not the same when it does not have guilt as a contrast. What people's ethical emotions are depends significantly on what they take them to be. The truth about Greek societies, and in particular the Homeric, is not that they failed to recognise any of the reactions that we associate with guilt, but that they did not make of those reactions the special thing that they become when they are separately recognised as guilt.[4]

Williams himself was a lumper when it came to shame and guilt. He thought that making a "special thing" of the reactions associated with guilt obscured rather than clarified those reactions, by detaching them from a sense of the self and its relation to others. "Shame can understand guilt, but guilt cannot understand itself", he wrote memorably.[5] But that is a further point. My claim here is simply that we cannot sensibly analyse the concepts of guilt and shame in abstraction from the history of the words by means of which they have been expressed.

The words *aidōs* and "guilt" are at any rate not co-extensional, but even where two words are co-extensional we cannot assume that they express the same concept. Take "homosexual" and "gay". I suppose there is a sense in which all homosexuals are gay and all gays homosexual, yet the two words – the one coldly medicalising, the other cheerfully affirmative – belong to different worlds of thought and feeling. One can understand why many gay people refuse the label "homosexual". Conversely, the late Allan Bloom insisted that he was

---

[4] Bernard Williams, *Shame and Necessity* (Berkeley and Los Angeles: University of California Press, 1993), 91.
[5] Williams, *Shame and Necessity*, 93.

not "gay", though he slept with men. "Invert" was his preferred self-description.[6] Homosexuality – or whatever one wants to call it – is not something independent of the various names it travels under. It depends significantly on what it is taken to be.

The culprit in the background here is a certain conception of meaning popular among analytic philosophers, according to which two names or descriptive phrases mean the same just in case they pick out the same class of objects. That conception of meaning may be useful for the analysis of scientific and mathematical language – the purpose for which it was originally devised – but it is woefully inadequate to the language of human action and feeling. This is because human actions and feelings, unlike numbers and molecules, are not wholly independent of the words we use to articulate them; when those words change, they change too. The planet Venus is what it is regardless whether we refer to it as the Morning or the Evening Star, but a person's state of feeling may differ significantly depending on whether he thinks of it as *aidōs* or guilt. This is just one aspect of the reflexive or "hermeneutic" character of human self-understanding, on which much ink has been spilt.

A couple of clarifying points. First, I am not rejecting *all* appeals to implicit understanding. Many languages lack words corresponding to "voluntary", but insofar as speakers of those languages are inclined to excuse people for doing things by mistake or under compulsion they may be said to possess the concept of the voluntary implicitly. A similar point applies to *action*, *intention* and *identity*. All these are philosophers' concepts, devised for special analytic purposes. We can impute them to speakers who lack the corresponding words without fear of foisting on them some culturally specific weltanschauung – or at least so it may be argued, though in fact the boundaries of the *weltanschaulich* are unclear and many purportedly neutral analyses conceal substantive ethical assumptions.

Second, I am not denying that concepts such as shame, gratitude and courage correspond to *something* universal in human experience. We immediately understand the shame of Adam and Eve when they were seen naked by God after eating the fruit; if we did not, the Genesis story would mean little to us. This primordial *something* to which all conceptions of shame etc. relate is what permits us to say that some of those conceptions are uniquely *true*, insofar as they disclose that something more directly and luminously than others. Philosophy need not cede its place to cultural history entirely. But more on this anon.

---

[6] At least, if Saul Bellow's thinly fictionalised account is to be trusted. See Saul Bellow, *Ravelstein* (London: Viking, 2000), 160.

Concepts as abstractions from linguistic practice

Few contemporary analytic philosophers believe that there is a universal language of morals whose logic we can hope to uncover, yet they continue for the most part to analyse concepts such as *duty*, *right* and *welfare* after the fashion of Hare et. al., without regard to their origin and development. This may look like sheer intellectual inertia, but an argument can be made – has been made – in defence of the approach. Most contemporary analytic philosophers think of concepts as abstractions from linguistic practice, whose identity is determined by their contribution to the truth-value of propositions of which they form a part. Concepts thought of in this way cannot change, any more than numbers or shapes can change. Words, however, can change in the concepts that they express; hence to say that a concept has changed is just a misleading way of saying that a word which once expressed one concept now expresses another. A division of intellectual labour thus suggests itself: philosophers have responsibility for the analysis of concepts, historians for the history of words. A similar division has been proposed by some historians themselves. "The various transformations we can hope to chart will not strictly speaking be changes in concepts at all", Quentin Skinner has written. "They will be transformations in the applications of the terms by which our concepts are expressed."[7]

This picture of concepts has allowed analytic philosophers to persevere in their traditionally ahistorical style of analysis but without the universalism that once accompanied it. Their job, as they see it, is to get clear about what "we" – more on that treacherous pronoun in a moment – think about *justice*, *welfare* etc., in full awareness that others in the past have thought differently. Concepts are, so to speak, stills from the moving picture of human speculation, set apart for purposes of formal analysis. They are changeless not because they are eternal but because change has been excluded from them by definition.

The first thing to say about this style of analysis is that it represents a massive retreat from the classical ambition of philosophy, which was to say something universally true about the basic structures of reality and of human life. This ambition still inspired earlier versions of conceptual analysis, insofar as they took as their object the universal logic or "grammar" of human thought. But when analysis confines itself to what "we" would say about x, y or z, its

---

[7] Quentin Skinner, *Visions of Politics*, Vol. 1 (Cambridge: Cambridge University Press, 2002), 179

only function is to give precision to the vagaries of contemporary usage. Philosophy appears to have become empirical sociolinguistics.

Yet if this really is empirical sociolinguistics, why is it done so badly? Claims about popular linguistic intuitions abound, but evidence to support them is nowhere to be found. All we get are blank assertions along the lines of "we would say ...", which sounds suspiciously like philosophers generalising from what *they* would say – a flagrant fallacy, as exponents of "ex-phi" are quick to point out. But perhaps the "we would say ..." formula should be understood not as a generalisation about what we actually do say but as an invitation to reflect on what we ideally *should* say, the "we" here referring not to some empirical collectivity but to (as it were) "you and I, dear reader." But in that case, concepts are not abstractions from current usage after all but *norms* to which current usage can conform or fail to conform – the very Platonism that modern analytic philosophy has officially left behind.

This sort of underhand normativity about concepts runs through much recent work in analytic philosophy, of which Fred Feldman's *What is this Thing Called Happiness?* is representative. The ostensible target of Feldman's analysis is "the ordinary, somewhat obscure thing that the unperverted English-speaking person on the street means when he or she says that someone is 'happy'."[8] That suggests an empirical investigation into average English usage, though as usual no evidence for such usage is ever produced. But look again at that oddly moralistic adjective "unperverted". What is it doing here? We find the answer some fifty pages earlier: "I think 'happy' is used in the evaluative sense only by philosophers and others whose good linguistic instincts have been perverted by them. Perhaps this usage harks back to a fumbled attempt to find a good English word to translate the Greek *eudaimonia*."[9] But by what right does Feldman dismiss as "perverted" the usage of many *bona fide* English speakers, some of them uncorrupted by philosophy? Couldn't his own hedonistic understanding of happiness be put down with equal justice to the "perverting" influence of Benthamite utilitarianism? Or is Feldman perhaps claiming some special "intuition" (using that word now in its classical sense) into what happiness *really is*? But if so, how could he hope to vindicate such a claim?

This style of conceptual analysis is fatally incoherent. It claims to take its bearing from the linguistic intuitions of a given population, but where those intuitions conflict, as in all

---

[8] Fred Feldman, *What is This Thing Called Happiness?* (Oxford: Oxford University Press, 2010), 186.
[9] Feldman, *What is This Thing Called Happiness?*, 136.

interesting cases they do, it has no way of resolving the conflict consistent with its own premises. For it, intuitions are simply data. If we reject some, it can only be because they are incompatible with others we wish to affirm, not because they are inherently "perverted". The decision which intuitions to affirm and which to reject is ultimately arbitrary; one person's *modus ponens* is another's *modus tollens*. I have argued the case with reference to happiness, but I could equally have focused on *virtue, vice, prudence, freedom* and *justice*. In all these cases, intuitions differ systematically in a way that seems beyond the power of philosophical argument to settle. But there is no need to labour the point. It has been made with great force by Alasdair MacIntyre.

It is just here that a historical approach to concepts can advance matters, in two respects. First, it can *explain* conflicts in our intuitions, by uncovering their origins in our divided cultural inheritance. And second, it can (at least sometimes) help us to *resolve* such conflicts, by revealing one side or the other as a product of social and intellectual developments we have reason to reject. A purely synchronic approach is helpless on both fronts. From its point of view, intuitions must remain a brute given, incapable of either explanation or evaluation. Our only task is to "save" the strongest or most numerous of our intuitions, discarding the rest in the interests of coherence ("biting the bullet" as it is inelegantly called). Philosophy is thereby reduced to the systematisation of prejudice. There is an analogy to be drawn with the way in which neo-classical economics, by abstracting from the process of want formation, deprives itself of all resources for explaining and judging wants, leaving only the question of their satisfaction.

But enough generalities. Let me provide a concrete example of the power of the historical method, focussing again on the concept of happiness. A historical analysis can, I will argue, give substance to Feldman's claim that certain uses of "happy" are "perverted", though it will turn out that these are precisely the uses which he regards as exemplary.

Happiness in historical perspective

What little Feldman says about the history of the concept of happiness is exactly wrong. The evaluative sense of "happiness" was *not* an invention of philosophers struggling to translate *eudaimonia*. On the contrary, it was the original, common meaning of the term. To be happy was to enjoy good "hap" or luck, an association preserved in stock phrases such as "happy chance" and "happy day". By the Elizabethan period, the implication of good luck had receded somewhat, but the word still referred to an objectively desirable condition of being,

not just a pleasant state of mind. "We few, we happy few", says Shakespeare's Henry V to his troops before Agincourt, knowing that many of them will be maimed or killed. "If ye suffer for righteousness' sake, happy are ye", runs the King James version of 1 Peter 3:14. "Happy" in this sense is a pure "verdictive"; it signifies success in life, but carries no implications as to what that success might consist in. The same is true of *eudaimon* as used by Plato, Aristotle and other ancient Greek philosophers. Thus when English translators of ancient philosophy sought an equivalent of *eudaimonia*, "happiness" came readily to hand.

Beginning in the late sixteenth century, however, "happiness" took on its standard modern meaning of "a state of pleasurable contentment of mind" (O.E.D.). Shakespeare provides us with an early example. "To sour your happiness," says Cornelius to Cymbeline, "I must report the queen is dead."[10] "Happiness" here must refer to a state of mind, not "good hap", or Cymbeline would not need to be *informed* of the death of his queen in order to lose it; he would already have lost it by the mere fact of her death. Some time later, "happy" acquired the colloquial sense of "glad" or "pleased", which (assuming the normal relationship between adjectives and their cognate nouns) must have further strengthened the psychological associations of "happiness". By the mid-eighteenth century, this development was complete. Samuel Johnson defined "happiness" firstly as "felicity; state in which the desires are satisfied" and only secondly as "good luck; good fortune".[11] Jeremy Bentham, writing in 1781, was content to define utility as "that property in any object, whereby it tends to produce benefit, advantage, pleasure, good, or happiness, (all this in the present case comes to the same thing)."[12]

Parallel developments occurred in other European languages. The German word *Glück* originally meant simply luck or good luck, acquiring the additional sense of contentment or joy in the late Middle Ages. (Both senses remain active to this day, as Ute Frevert reminds us in this volume.) The indistinguishability of *Glück* (= happiness) and *Glück* (= luck) troubled philosophers like Leibniz, who lighted upon an alternative word, *Glückseligkeit,* to designate

---

[10] *Cymbeline* act V, sc. V, line 27. See O.E.D.. The O.E.D. lists a couple of examples of "happiness" in this sense dating from 1500 and 1534, but they both seem to me to admit of interpretation as "good hap".

[11] Samuel Johnson, *Dictionary of the English Language* (London: Strahan, 1755), vol. 1, 965.

[12] Jeremy Bentham, *An Introduction to the Principles of Morals and Legislation*, chap. 1.3. In Jeremy Bentham, *A Fragment on Government and An Introduction to the Principles of Morals and Legislation* (Oxford: Blackwell, 1967), 126.

happiness as an enduring inner state, invulnerable to fortune.[13] *Glückseligkeit* soon established itself as the standard counterpart to *eudaimonia* in philosophical German, yet its ordinary sense was always more spiritualised than that of the Greek term, and became still more so over time (Lutheranism was an influence here, as was the fallacy, widespread in the eighteenth century, that the *selig* in *glückselig* derives from *Seele*, "soul").[14] All this contributed to the revolt against eudaimonism in the late eighteenth century. Kant, the leader of that revolt, saw no essential difference between *Glückseligkeit* and *Glück*; both for him were mere psychological states, incapable of grounding duty. Morals, as he famously put it, "is not properly the doctrine of how we are to *make* ourselves happy but of how we are to become *worthy* of happiness".[15]

The main French word for happiness, *bonheur*, also originally signified good luck, as its etymology (*bon heur* = good hour) might suggest. However, from the fifteenth century onwards it too came to denote a state of joy or contentment, drawing close in meaning to "pleasure".[16] "I am enjoying heavenly happiness (*bonheur*)", cries Éradice, a character in the popular eighteenth-century pornographic novel, *Thérèse Philosophe*, as she is ravished by the lecherous Father Dirrag.[17] Pleasure is "the happiness of an instant, an element of happiness", runs an entry in Pierre-Benjamin Lafaye's 1858 *Dictionnaire des Synonymes*.[18]

Today, the psychological sense of "happiness" is unquestionably the dominant one, though the original evaluative meaning remains active under the surface. Some well-known thought experiments help bring this out. We hesitate to describe as "happy" a woman who is deeply deluded about core aspects of her life, however much fun she is having.[19] And we would probably not say that a doctor who performed a lobotomy on a new-born baby after hearing its parents express a wish for its happiness had correctly interpreted their meaning.[20] In an episode of the TV series, *Mad Men*, ad executive Roger Sterling grumbles at the chilly

---

[13] See *Deutsches Wörterbuch von Jacob Grimm und Wilhelm Grimm*, ed. Theodor Kochs, Joachim Bahr and others (Leipzig, S. Hirzel: 1958), vol. 4, 347.

[14] See the article on "*Glück*" by Christian Helmreich, in Barbara Cassin (ed.), *Dictionary of Untranslatables: A Philosophical Lexicon* (Princeton: Princeton University Press, 2014), 397-402.

[15] Immanuel Kant, *Critique of Practical Reason*, trans. Mary Gregor (Cambridge: Cambridge University Press, 1997), 108.

[16] See "Bonheur" in *Le Grand Robert de la Langue Française* (Paris, Le Robert: 1985), Vol. 2, 67-8.

[17] *Thérèse philosophe*, ed. Florence Lotterie (Paris: Flammarion, 2007), 98.

[18] Pierre-Benjamin Lafaye, *Dictionnaire des Synonymes de la Langue Française* (Paris: Librairie Hachette, 1884), 408.

[19] See Julia Annas, *Intelligent Virtue* (Oxford: Oxford University Press, 2011), 141

[20] See Philippa Foot, *Natural Goodness* (Oxford: Oxford University Press, 2001), 85-6.

reception handed out to him and his new bride, a vacuous young gold-digger. "It's a mistake to be conspicuously happy: some people don't like it", he says to his friend Don Draper. "No one thinks you're happy", replies Don. "They think you're foolish."[21] Since Roger clearly *is* happy in the ordinary, smiley-face sense, Don's use of the word here must be understood as verdictive, yet it comes over as perfectly natural and idiomatic. Certainly no one could accuse this suave and worldly adman of having had his linguistic instincts perverted by philosophy!

How should we describe this semantic situation? One superficially plausible view is that "happy" is a homonym. It has two quite different senses: a newer one, now standard, and an older one, now invoked only rarely. On this view, Don Draper is simply punning, as he might be had he said: "No one thinks you're gay. They think you're miserable". But we need only to put the point this way to see that it is false. "Happy" is not homonymous, and Don Draper is not punning. When Roger says that he is happy, and Don says that he is not, they are *disagreeing*, not playing on words. Something Bernard Williams says about liberty is relevant here. "It is important", Williams writes

> that the disputes that have circled around the various definitions and concepts of liberty do not just represent a set of verbal misunderstandings. They have been disagreements about *something*. There is even a sense in which they have been disagreements about some one thing. There must be a core, or a primitive conception, perhaps some universal or widely spread human experience, to which these conceptions relate. ... This core or primitive item ... can, and must, explain how these various accounts of the value of freedom are elaborations of the same thing, that these various interpretations are not just talking past one another.[22]

As with liberty, so with happiness. This too must be "one thing" if we are to make sense of our impression that Roger and Don and all the others who have debated the concept are in genuine disagreement with one another, and not just at cross purposes. But what is this "one

---

[21] *Mad Men*, Session 3, Episode 3.

[22] Bernard Williams, "From Freedom to Liberty: The Construction of a Political Value", in Bernard Williams, *In the Beginning was the Deed: Realism and Moralism in Political Argument* (Princeton: Princeton University Press, 2005), 76.

thing"? A clue is offered by the verb "felicitate", from the Latin word for happy.[23] To call someone happy is, at a very minimum, to felicitate, to congratulate him, to say "well done!" It is not necessarily to say anything more than this. However, "happy" and its cognates have always been prone, by a natural process of association, to incorporate particular ideas about what *merits* felicitation – ideas which have varied according to the values of the society in question. In the popular language of fourth-century BC Athens, to call someone *eudaimon* was generally to imply that he was rich.[24] When the heroines of eighteenth-century English novels look forward to, or wish each other, "happiness", we know without further comment what they are talking about. For women in their situation, success in life means marriage.

My claim, then, is that happiness's new meaning of joy or contentment should be understood as a *specification* of its old meaning of success in life, not as an alternative to it. It represents a particular interpretation of what success in life *is*. That is why it remains appropriate to refuse to apply the word to characters like Roger, who, though joyful, contented etc., are in some obvious sense failures. Of course, not all uses of "happy" are evaluative. It would be absurdly solemn to respond to "I was happy to meet Sarah yesterday" with, "Ah, but were you *really* happy?" "Happy" here means nothing more than "glad". But where the word is used deliberately, with emphasis – as in "I found happiness at last" or "I had a happy life" – it retains something of its original verdictive force, making the question "but did you really ...?" occasionally appropriate.

These brief reflections on "happiness" illustrate a more general point. A word does not pass through the phases of its history like a train through railway stations. Its former uses remain latent within it, preserved in the classics of the literature and in those little fossilised idioms that are a feature of every natural language. This is why conceptual analysis cannot proceed in a purely synchronic fashion, as though the current meaning of a word could be completely sequestered from its previous meanings. It must proceed diachronically, or genealogically.[25] Analytic philosophers have something to learn from linguists here, as well as from thinkers in the hermeneutic tradition of Dilthey and Gadamer.

---

[23] See Timothy Chappell, "Eudaimonia, Happiness, and the Redemption of Unhappiness", *Philosophical Topics* 41/1 (Spring 2013), 36. Chappell points out that "felicitate" has an almost exact counterpart in the ancient Greek verb *eudaimonizein.*

[24] Kenneth Dover, *Greek Popular Morality in the Time of Plato and Aristotle* (Indianapolis: Hackett, 1994), 174.

[25] Hannah Dawson makes a similar point in her essay in this volume, "Shame: from Sin to Sociability", ...

But back to "happiness". I have suggested that its dominant modern meaning of "joy" or "contentment" represents a specification or interpretation of its original meaning of "success in life". But the question remains: why *this* interpretation? Why did the meaning of the word become narrowed in this way and not some other? Some light on this question is shed by the founders of modern philosophy, Descartes, Hobbes and Locke. Their reflections on happiness are of particular relevance here, not so much for their subsequent influence – though this was undoubtedly great – as for the light they cast on intellectual motives *already* at work in the spontaneous evolution of the term's meaning. They show us, more clearly than other sources, just why Europeans of the early modern period began to think of success in life as a matter of achieving certain states of mind.

The Descartes/Hobbes/Locke view of happiness can be represented as the conclusion of an argument, the first premise of which is conventionally Aristotelian: happiness is the final and comprehensive object of human desire. Everything we want we want either as a means to, or as a part of, happiness. But (second premise) the real object of our desires is only ever a mental state. We may say we want to eat an apple or see a movie, but what we *really* want is the pleasure that accompanies eating the apple or seeing the movie. From these two premises, the conclusion follows directly: if the real object of desire is always a mental state, then happiness, as the final and comprehensive object of desire, must itself be a mental state. The argument as a whole is, of course, a characteristic product of the Cartesian revolution. It reflects a picture of the mind as occupied exclusively with its own contents, receiving only causal impacts from the world outside – a picture that dominated modern philosophy from Descartes through to Heidegger and Wittgenstein.

Variations on this argument are implicit in all three of our seventeenth-century witnesses. Descartes' late work, *The Passions of the Soul*, characterises desire, conventionally enough, as involving a "volition to acquire some good or avoid some evil". It adds, however, that this good lies ultimately in the mental state of joy, "for in fact the soul receives no other benefit from all the goods it possesses; and as long as it derives no joy from them, we may say that it does not enjoy them any more than it would if did not possess them at all."[26] This "motivational internalism" (as we can call it) also underlies a letter of Descartes to Princess

---

[26] *The Philosophical Writings of Descartes, Vol. I*, trans. John Cottingham, Robert Stoothooff, Dugald Murdoch (Cambridge: Cambridge University Press, 1985), 365 and 360.

Elizabeth of Bohemia, in which he insists, against Seneca and the whole ancient tradition, that happiness is something distinct from the supreme good. The supreme good is "undoubtedly the thing we ought to set ourselves as the goal of all our actions", but it is only insofar as it is correlated with happiness, defined as "contentment or satisfaction of mind", that we have a motive to seek it out in practice:

> Suppose there is a prize for hitting a bull's-eye: you can make people want to hit the bull's-eye by showing them the prize, but they cannot win the prize if they do not see the bull's-eye; conversely, those who see the bull's-eye are not thereby induced to fire at it if they do not know there is a prize to be won. So too virtue, which is the bull's-eye, does not come to be strongly desired when it is seen all on its own, and contentment, like the prize, cannot be gained unless it is pursued.[27]

Hobbes' *Leviathan*, written at about the same time as Descartes' *Passions of the Soul* and bearing clear marks of Cartesian influence,[28] contains a statement of the same motivational internalism: all desire or appetite is for what is good, or, more precisely, for what promises pleasure – pleasure being simply "the appearance, or sense of Good".[29] Happiness ("felicity" Hobbes calls it) is nothing but the pleasure of satisfying desire – not just once and for all, but over and over, because "Life it selfe is but Motion, and can never be without Desire":[30]

> For there is no such *Finis ultimus*, (utmost ayme,) nor *Summum Bonum*, (greatest Good,) as is spoken in the books of the old Morall Philosophers. Nor can a man any more live, whose Desires are at an end, than he, whose Senses and Imaginations are at a stand. Felicity is a continuall progress of the desire, from one object to another; the attaining of the former, being still but the way to the latter.[31]

Locke, building on Descartes and Hobbes, states the crucial premise even more clearly: "We *love*, *desire*, *rejoice*, and *hope*, only in respect of pleasure; we *hate*, *fear*, and *grieve*, only in respect of pain ultimately: in fine, all these passions are moved by things only as they appear

---

[27] Letter from Descartes to Princess Elizabeth, 18 August 1645. In *The Philosophical Writings of Descartes, Vol. III: The Correspondence*, trans. John Cottingham, Robert Stoothoff, Dugald Murdoch, Anthony Kenny (Cambridge: Cambridge University Press, 1991), 262.

[28] See Ross Carroll's essay in this volume, "How Contempt became a Passion", …

[29] Thomas Hobbes, *Leviathan* (Cambridge: Cambridge University Press, 1996), 40

[30] *Leviathan*, 46

[31] *Leviathan*, 70

to be the causes of pleasure and pain, or to have pleasure or pain some way or another annexed to them."[32] Happiness, as the ultimate object of our desires, must be "the utmost pleasure we are capable of, and misery the utmost pain."[33] Locke sees that this view of the matter makes idle all traditional debates about the "real" nature of happiness:

> The philosophers of old did in vain inquire, whether *Summum bonum* consisted in riches, or bodily delights, or virtue, or contemplation. And they might have reasonably disputed, whether the best relish were to be found in apples, plums or nuts; and have divided themselves into sects upon it. For as pleasant tastes depend not on the things themselves, but their agreeableness to this or that particular palate, wherein there is great variety; so the greatest happiness consists in the having those things which produce the greatest pleasure, and the absence of those which cause any disturbance, any pain. Now, these to different men are very different things.[34]

In sum, Descartes, Hobbes and Locke all continue to use the word "happiness" in the traditional sense, as a name for our heart's desire. But because they conceive of desire as something that can only ever be directed towards states of mind, they must perforce conceive of happiness too as a state of mind; and thus the old evaluative sense gets overlaid with a new descriptive one. I'm not claiming that this exact argument passed explicitly through the mind of every seventeenth-century Jo Blogs who used the word "happiness" in the new sense. But we must assume that it, or something like it, was there implicitly if we are to make sense of the transformation in the term's meaning. Moreover – and this is the key point – the argument is a flagrantly wrong-headed one. States of mind are *not* always what we really want. Sometimes they are, as when we set out to get drunk or to feel weepy, but such cases are recognisable by contrast with more typical instances of human striving. "We don't start love-affairs to secure the stimulation of our pubic nerve", writes Michael Frayn. "(That would be an insanely long way round to go, like buying a car to get the use of the cigar lighter.)"[35] But I'm not going to argue the point further here. I shall trust to my readers' good sense and move on.

---

[32]John Locke, *An Essay Concerning Human Understanding*, Vol. I (London: Dent, 1961), 192.

[33] Locke, *Essay*, Vol. I, 213.

[34] Locke, *Essay*, Vol. I, 222.

[35] Michael Frayn, *Constructions* (London: Wildwood House, 1974), 144.

But isn't this all just a matter of old history, my adversary might at this point protest. The confusions surrounding the birth of a concept need not accompany it forever, or else (to revert to my old analogy) the modern meaning of "gay" might be rejected on the grounds that it grew out of a mistaken belief that all homosexuals are cheerful. But "happiness" is not like "gay", as I've said already. Its older and newer senses are not two separate things. The one is hidden in the other like a Jack-in-the-box, meaning that every time we call someone happy in the smiley-face sense we risk implying, fallaciously, that he has achieved his life's aim. The confusions of the seventeenth century are also our confusions.

So how should we respond to this predicament? We could just *stipulate* that "happiness" be confined to purely descriptive uses, introducing a quasi-technical substitute like "welfare" or "flourishing" to designate success in life. But that would be a mistake, for two reasons. First, "happiness" as a term of ordinary language is not ours to define at will. Trying to limit its meaning by fiat will only create confusion, in our own minds as well as others' (think of the mischief done with "nonsense" by A. J. Ayer or with "selfish" by Richard Dawkins). Second, I am not sure that "happiness" in the verdictive sense *has* any exact equivalents. "Flourishing" and "welfare" are both too determinate. They pick out the specifically "prudential" good, the good which is *for* him or her. "Happiness" is not restricted in this way. "If ye suffer for righteousness' sake, flourishing/faring well are ye" sounds all wrong, and not just for stylistic reasons.

My own preference would be to resolve the ambiguity in the other direction, by reserving "happy" for situations where felicitations really are in order. I am not of course suggesting that we tick people off for saying "I'm happy to drive you to the station" etc.; that would be intolerably pedantic. But where the word can be taken to imply a verdict, a false verdict, we do well to look for alternatives. Hugh Hefner no doubt had a pleasant and fun life. I am not sure he had a happy life.

Conclusion

I hope to have shown two things in this essay. First, our moral concepts are not universal or timeless. "Guilt", "shame", "happiness", "duty" – such words can change, and when they do, the things they stand for change too. If there is a permanent core behind this flux, it is visible only darkly and obliquely, perhaps in certain characteristic human experiences and gestures, perhaps also in the great narratives of world literature. Second, moral philosophers cannot hope to insulate themselves from this flux by confining their attention to an ideal present.

Moral concepts are a product of history, intelligible only in relation to history. Conceptual analysis implies historical analysis.

Does this represent a surrender to relativism? Not at all. On the contrary, it is precisely the "time-slice" method currently in favour among analytic philosophers that represents the real victory of relativism, for where philosophy confines itself to harmonising our "intuitions" – meaning, to put it plainly, the conventional prejudices of modern academic philosophers – it tacitly abandons all aspiration to a deeper and more permanent point of view. We are condemned to a kind of provincialism, a shuffling back and forth among the small circle of currently licenced commonplaces. Only history can liberate us from this provincialism, by uncovering the sources of what passes among us for common sense. (I have sketched what such an "uncovering" might look like in relation to *happiness*; no doubt a similar approach could be extended to other moral concepts.) Philosophy *can* be something more than "its own time apprehended in thought", but only if it first becomes conscious of the extent to which it *is* its own time apprehended in thought. It cannot transcend time by simply ignoring it, in the manner of most analytic philosophy.