Well-Being in Action

Dr Don Berry

23rd October 2023

Abstract

This thesis consists of a critique of two influential accounts of what is good for a human being: that which satisfies their desires, and that which gives them pleasure. Both accounts are therefore ‘subjectivist’ in that they are centred on the personal values of the agent in question.

In contrast to the conventional philosophical approach, the study will evaluate these two theories in the context of their suitability for guiding our deliberations about what to do and how best to live. It offers an argument against both theories when they are thus put to work as a basis for making our choices.

The argument proceeds by considering what is required in order for us to develop and maintain certain practical and intellectual powers that are essential to our forming an adequate conception of our own good and effectively pursuing this ideal in practice. I will claim that gaining and maintaining these powers requires a certain kind of supportive social context – and, in particular, engagement in certain kinds of formative relationships with others. However, participating in these relationships in the required way necessitates our having an openness to learning from and being influenced by these others that is not possible for an agent who fully commits to either of our two subjectivist views.

Contents

1 Elements of a Philosophy of The Human Good **1**
1.1 Introduction: Action-Guiding and Appraisive Approaches 2
1.2 Preliminary Features of The Enquiry 8
1.3 Guiding Practical Reason and Normative Constructivism 23
1.4 Timescales 32
1.5 Role of Theory: Action Selection 40
1.6 Form of Theory: Life Planning 44
1.7 Intrinsic and Instrumental Value 52
1.8 Conclusion 60

2 Two Subjectivist Accounts **62**
2.1 Decision Theory 63
2.2 Desire-Based Accounts In General 82
2.3 Pleasure and Measurement 93
2.4 Hedonism as a Guide to Action 108
2.5 Conclusion 121

3 The Social Context of Practical Reason **123**
3.1 Practical Reason and its Requirements 124
3.2 The Long Childhood 137
3.3 Formal Education 151
3.4 Practices 164
3.5 The Sapling and the Caterpillar 181
3.6 Physical Infirmity, Mental Illness, and Old Age 192
3.7 Conclusion 208

4 Conclusion **210**

5 Bibliography **214**

Chapter 1

Elements of a Philosophy of The Human Good

> *The only critique of a philosophy that is possible and that proves something, namely trying to see whether one can live in accordance with it, has never been taught at universities: all that has ever been taught is a critique of words by means of other words.*
>
> – Friedrich Nietzsche[1]

This opening chapter offers a discussion of the shape of the coming enquiry: what path it will follow, what assumptions it makes, and what standards of success it aims to realise. The goal of the chapter is to develop a framework for the evaluation of theories of what I will refer to as the 'Human Good'; a phrase intended as a non-committal term of art, in advance of the more detailed specification of what we are after that will come later.[2]

In a departure from the dominant philosophical approach, the framework developed here will aim at evaluating our two key accounts of the Human Good within the context of practical reason, treating them as standards that guide our actions and choices – both in the here-and-now, and in the unfolding of our lives as a whole.

[1] Nietzsche, 1997 [1874], p. 187.

[2] Throughout I will use capital letters to indicate that a special, explicitly stipulated sense of a term is intended.

1.1 Introduction: Action-Guiding and Appraisive Approaches

The key concern of this thesis is to explore accounts of what is good for us that are specifically intended to serve as standards to guide our rational deliberations about what we should do and how we should live our lives.[3] Rather than offering a positive solution to the problems it identifies, the thesis will instead constitute a sustained criticism of two influential conceptions of what is good for us that aim to fulfil this aspiration. These accounts are given in terms of what we desire and what gives us pleasure, respectively; these features of ourselves *qua* individuals I will refer to collectively as our 'Subjective Values'. In this opening chapter, I aim to develop a framework within which these two 'Subjectivist' views will be subject to criticism – though the framework itself is intended to be neutral between and compatible with a range of specific views.[4]

In proceeding with this task, I shall make use of the literature from three existing approaches to the subject matter: two new, and one old.[5]

The first approach is a particular kind of specialised, theoretical enquiry into the good life conducted within analytic philosophy. Here theorists aim to find precise conditions that encapsulate what it is for a human life to go well in the abstract, as expressed in the quasi-technical concept of 'well-being'. This research programme has its intellectual roots in the late 19th Century, and continues to be pursued by many contemporary academic philosophers today.

Our second approach is a movement within ancient Graeco-Roman ethical thought: the practical enquiries into the good life begun by Socrates, Plato, and Aristotle, and which continue via debates amongst rival philosophical schools throughout the Hellenistic period amongst the Stoics, Epicureans, Cynics, Cyrenaics, and Skeptics. The shared starting point here is the problem of how human beings should make sense of

[3] In this enquiry I will not deal with the interesting question of *when* such rational, cognitively-guided actions occur, in either its normative or descriptive aspects – nor their relation to other kinds of human behaviour; rather, we bracket off this issue and focus entirely on the content of such deliberations when they do occur.

[4] See especially 1.4-1.6 below for discussion of implicit assumptions that might bias the enquiry in favour of one particular type of theory.

[5] In this chapter I will use the term 'approach' in a loose way to gather together sets of theories that have been offered within the same intellectual and historical context, and which largely share the same goals and conception of the subject matter; the three example approaches given below should make the term sufficiently clear for our purposes.

their lives as a whole, though participants to these debates come to a wide variety of conclusions about how best to do this. Some of these responses, such as ancient hedonism, are Subjectivist in the sense I will go on to criticise; many, however, instead centre on acquiring the virtues.

The genesis of our third approach is on roughly the same timescale as the first, but within academic economics and the social sciences rather than philosophy. Within the normative parts of these disciplines I shall be concerned with, theorists advance accounts of how to make choices in general, focusing now on discrete actions rather than our lives as a whole. What all such theories prescribe here is that we act so as to maximise our preferences; our key example is the Expected Utility Theory (herein EUT) of neoclassical microeconomics, a highly developed technical framework of rational action.

In this opening section, I address certain important differences between these three approaches, as well as between each of them and the framework of the present enquiry. To this end, it will be helpful to introduce a distinction between what I will call 'Appraisive' and 'Action-Guiding' approaches to the Human Good. An 'Appraisive' approach – as is taken by most work on well-being within contemporary analytic philosophy – is one that aims at standards of evaluation that are purely abstract and theoretical, with no import on what real-world agents should actually do. In contrast, like the present enquiry, both ancient philosophical enquiries into the good life and the relevant modern work on choice within economics and related subjects today are 'Action-Guiding' – meaning that they instead aim to develop conceptions of the good for the purpose of informing the choices of these agents, as part of a normative theory of deliberation. They therefore have strong ties to practice. We now explore this distinction further, whilst elucidating some other central features of our three key approaches.[6]

We begin with the modern philosophical approach. Many analytic philosophers working on well-being today – such as Feldman, Sumner, Rosati, Velleman, Sarch, Fletcher, Hooker, and others – aim primarily to develop and justify accounts of what it is for a human life to have gone well or badly, as assessed from a vantage point external to it. They seek the general features of particular human lives – themselves usually

[6]Given space constraints, I shall not attempt to write a history of the emergence of these three approaches, or discuss the long interval between the ancient and modern periods in which they develop – though for a warning about omissions of this type, see MacIntyre, 1998b, pp. 260-261.

described at a highly abstract level – which determine when they have thus gone well or badly for the person in question, as judged from this third-person perspective.[7] Such theorists typically expect the account to take the form of a precise logical analysis, given in terms of individually necessary and jointly sufficient conditions.[8]

The kind of conceptual knowledge these philosophers are after is purely theoretical in character, in the sense that the assessments made and general standards of evaluation for human lives offered are not taken to have direct implications for practice, and these alleged 'facts' about what well-being consists in are typically seen by these analytical theorists as strictly independent of our attempts at its incarnation in the practical world.[9] For instance, Feldman – an influential advocate of pleasure-based accounts of the good life within this tradition – writes that 'Hedonism (as I understand it) says nothing about what we should do, or what we should seek.'[10] Their enquiries therefore constitute an 'Appraisive' approach in the above sense, and this also suggests a convenient disciplinary boundary: philosophers of this ilk set themselves the task of discovering, elucidating and justifying the appropriate conditions governing whether a human life goes well or badly, but do not concern themselves either with investigating what causes enable such lives to come about, nor with enabling their readers to improve their own lives in practice.[11] Within moral philosophy more broadly, Sidgwick evinced this spirit early on:

> I have thought that the predominance in the minds of moralists of a desire to edify has impeded the real progress of ethical science: and that this would be benefited by an application to it of the same disinterested curiosity to which we chiefly

[7]Cf. Feldman, 2004, p. 13: 'I want to know, in the abstract, what features make a life good for the one who lives it'; see also Fletcher, 2016 for a concise and accessible introduction with extensive references to the recent literature.

[8]Feldman, 2004, p. 13.

[9]For criticism of this attitude, emphasising that the mode of justification for any claim within moral philosophy is always ultimately practical, see Collingwood, 1992 [1940], p. 402.

[10]Feldman, 2004, p. 31. Feldman indeed seems at times even to dispense with the ambition of actually applying his abstract criterion in practice, so long as there is a 'fact of the matter' about whether it does apply – see *ibid.*, p. 65.

[11]See e.g. Sumner, 1996, p. 16 for the view that philosophical theories of the nature of well-being must be purely 'formal' in this sense; see also Feldman, 2010, pp. 6-7. For criticism, see Angier, 2015, pp. 13-14, who argues that this dichotomy between the nature and causes of happiness is based on a misleading conception of its genesis via 'efficient causation': rather, he urges that the Aristotelian view that 'happiness is inextricably grounded in its conditions' (p. 23).

owe the great discoveries of physics.[12]

This text would go on to exert a great influence on the founders of analytic moral philosophy, and this break from prior traditions thus marks the agenda for much of what is to come afterwards.[13]

Turning back to antiquity, we find a much different philosophical approach to the study of the Human Good, a canonical example of which is Aristotle's *Nicomachean Ethics.* In Book I, Aristotle elucidates the practical *rationale* for his enquiry by invoking the model of an archer who is more likely to shoot successfully when he has a target (*skopós*) to aim at,[14] and this same metaphor for ethical theory informing practice is espoused by a number of other ancient Greek thinkers too.[15] What is illustrated by Aristotle's remarks here is true of ancient philosophising about the good life more generally: such enquiries are not seen as purely theoretical endeavours; rather, their key purpose is to inform the choices of human beings in the real world.[16] Thus, the approach is 'Action-Guiding' in the above sense.

In contrast to the abstract approach of most analytic philosophers of well-being, the ancients typically begin by considering the choices faced by a situated human being – who will have their own pre-existing commitments, talents, desires, preferences, tastes, and other personal characteristics – and ask what is best for them to aim at having, doing, or being under these circumstances. Although typically it is again the agent's life as a whole that is central,[17] the focus is now on their own embodied conception of the kind of life they aim to live and the expression of this ideal in their practical reasoning, rather than developing criteria

[12] Sidgwick, 1981 [1907], p. viii.

[13] On the divorce of theory from practice within contemporary academic moral philosophy, see MacIntyre, 2016, p. 71. By the time of Whewell's lectures on the history of ethics in the 19th Century, it could indeed be claimed that Aristotle's ethical remarks are 'not those of a moral theorist, but those of a man of the world' – Whewell, 1862, p. 27. For modern resistance to this attitude, see Collingwood, 1992 [1940], p. 402; Dewey, 1998, Vol. 1, p. 66; and MacIntyre, 1992, p. 3.

[14] Aristotle, *NE* I.1, 1094a22-25 [All references to Aristotle are from Jonathan Barnes' edited complete works; Aristotle, 1984]. See Annas, 1993, p. 34 for criticism of Aristotle's formulation here, via a distinction between *skopós* – the thing aimed at – and *telos* – the agent's attainment or achievement of that thing. This is a point observed by Arius, who is quoted in *ibid.*, p. 34.

[15] For instance, the Stoics: see Annas' discussion of Antipater of Tarsus: Annas, 1993, pp. 402-403.

[16] For an accessible introduction to philosophy in the key period following Aristotle, see Adamson, 2015; see also Cooper, 2012 for further discussion of the ancient schools I will discuss.

[17] Cf. Annas, 1993, pp. 27-29; p. 440; p. 443. For an exception here, see the discussion of Aristippus in Section 2.3 below.

for assessing their lives from an external vantage point. The theories of the good life on offer in the ancient literature are thus not conceived in terms of abstract accounts of a technical philosophical concept such as 'well-being', but rather as the basis for Action-Guiding responses to a broad practical problem that each human being faces: they tell us how to order our existing concerns, interests, ambitions and projects and give each their due place as we shape the unfolding story of our lives as a whole – and, often derivatively, how to act in the here-and-now.[18] Indeed, Aristotle goes on to say, quite explicitly, that 'the present inquiry does not aim at theoretical knowledge like the others (for we are inquiring not in order to know what excellence is, but in order to become good, since otherwise our enquiry would have been of no use)'.[19] This again contrasts starkly with Sidgwick, who counters that 'my immediate object – to invert Aristotle's phrase – is not Practice but Knowledge'.[20]

Returning to the modern world; a quite different Action-Guiding approach to the Human Good can be found within our third key approach, as exemplified by the formal framework of choice given by EUT.[21] Here the focus is again on guiding what ordinary agents actually do in practice, and the resulting conception of the Human Good – now given by the maximisation of their preferences – is put in service to just this end. However, the starting point is now no longer the agent's life as a whole; rather, the approach instead focuses on specific, local choice problems in the here-and-now – for instance, selecting between different bundles of consumer goods subject to fixed prices and an income constraint.

Comparing our three approaches, then, we see that the more comprehensive enquiries of ancient philosophy have been resolved into a study of rational action within economics which ignores the broader context of agents' lives as a whole, and abstract enquiry into the good life within philosophy from which no practical consequences for how to act are drawn. A core goal of the present thesis is to offer a remedy to this fragmentation.

In this next section, I will introduce some further features of the

[18] On ancient philosophy as a way of life, see again Cooper, 2012, Chapter 1.

[19] Aristotle, *NE* II.1, 1103b27-29; cf. *NE*, X.9 1179b1-4. On Aristotle's warnings against 'taking refuge in theory' (in Ross's apt translation) at the expense of commitment to practice, see Dunne, 1993, pp. 319-320.

[20] Sidgwick, 1981 [1907], viii.

[21] Here we will interpret this framework as a normative guide to action, rather than a descriptive theory that merely aims to explain or predict what real-world agents will actually do; see Section 2.1 below for detailed discussion.

Action-Guiding framework of enquiry that will be developed here; but first, we pause to consider a potential objection to what has been said so far. For readers of an analytical bent may respond that this contrast between the two types of approach – Appraisive and Action-Guiding – cannot be of any great theoretical importance for the serious philosophical study of the Human Good itself. 'Yes,' they may retort, 'when contrasting the two types of approaches in question it is plain that there is a change in style, focus, and aspirations. However, at the level of rational justification these cannot ultimately bear on what the correct account of the Human Good turns out to be, since this question must be settled in a way that is strictly independent of any practical use to which we will later put it.' Hence, even when pursuing an Action-Guiding approach that asks what human beings should take as their ultimate ends in practice, whether in an ancient-philosophic or a modern-economic guise, it might be claimed that any adequate account of the *content* of these objectives must be vindicated by just the same standards that inform a theoretically-focused Appraisive approach.

This objection fits well with the contemporary separation of the conceptual and practical tasks of moral philosophy – and indeed, some analytic philosophers of well-being have explicitly advocated that we first focus on setting up our abstract criterion governing what the Human Good is, only later turning to the practical task of how to achieve this in practice – or else handing this over to a different group of thinkers altogether. Feldman thus contends that 'we cannot do a fully responsible job of answering these practical questions until we have a somewhat clearer answer to the question I do mean to ask' – that is, to find the correct abstract criterion for the Appraisal of human lives.[22] Nevertheless, I shall argue that the core assumptions and attitudes underlying the objection are misguided – and indeed may be a product of the specialised, purely conceptual approach that is dominant within the contemporary philosophy of well-being itself, and its consequent divorce from moral practice. Economists and other social scientists, meanwhile, have long recognised the importance of the distinction between merely evaluative criteria and active goals – as expressed in Goodhart's law, which states that 'When a measure becomes a target, it ceases to be a good measure.'[23]

[22]Feldman, 2004, p. 13; cf. Feldman, 2010, p. 8: 'the philosophical project has priority over the empirical project'; see also Moore, 1993 [1903], p. 21 for the same strategy regarding goodness in general.

[23]As famously put by the social anthropologist Marilyn Strathern (paraphrasing

In particular, I shall later argue that the core concept of the Human Good within economics has been accepted partly because it has proved to be conducive to practical success in their particular enquiries into human choice, rather than purely because of any intrinsic, *a priori* plausibility.[24] And more generally, from the next section onwards we shall see that in fact the spirit in which the topic is broached does impact substantively on what the content of an adequate theory of the Human Good can be.[25]

1.2 Preliminary Features of The Enquiry

In this thesis, we are interested in Action-Guiding accounts of the Human Good that can be applied in practice by ordinary moral agents in the real world, as they respond to the ongoing problem of working out what to do and how to live. This section draws attention to a number of features that our framework of enquiry must have if this aim is to be realised, and describes how it will therefore differ from other approaches – both Action-Guiding and Appraisive – that are currently dominant.

1.2.1 Situated and Global Perspectives

The first distinction I would like to introduce is between the situated perspective that must be adopted by any practically-applicable Action-Guiding enquiry into the Human Good, and the global view of an agent's life taken by Appraisive approaches within contemporary analytic philosophy. From the latter, global perspective, an agent's life is assessed from a privileged standpoint that is external to it; in taking a situated perspective, in contrast, we aim to understand the agent's choices from their own internal point of view. Consequently, several additional constraints arise that theorising must respect to be adequate to informing practice.[26]

Firstly, real-life decisions about what to do are always made from the local physical, social, and cultural circumstances of one particular individual, with their own peculiar history – all of which constrains the

Keith Hoskin) in her critique of assessment practices in education; see Strathern, 2016; The same idea is also known as 'Campbell's law' within the social sciences more broadly.

[24] A point also made forcefully in Ghoshal, 2005 – see especially p. 20.

[25] For defence of the same view, see Annas, 1993, p. 42; Dewey, 1998, Vol. 1, p. 109; and Prinzing, 2021, p. 158: 'concept(s) of well-being should also be able to feature prominently in ordinary decision-making. This places a few constraints on the descriptive content of the concept.'

[26] For more on this distinction, see our Pfeifer, Bongard, and Berry, 2011, pp. 54-55.

options available to them.[27] Indeed, when thinking about the direction of their lives as a whole, then at any one moment they are not choosing between different lives *per se*, but rather between competing ways to live out the remainder of their current life as it has unfolded hitherto.[28]

A second, related consequence of the situated perspective is a changed relationship to time. Typically what is sought on an Appraisive approach is a single, timeless assessment of a human life as a complete entity; but on an Action-Guiding approach we are confronted not just with one, but a vast multiplicity of opportunities for action at each stage of our ongoing lives. Each such choice has its own unique information base, as we acquire new beliefs over time and lose old ones as memories deteriorate. Moreover, as we develop as moral agents our values, tastes, desires, and preferences will also change too, and our conception of how our lives ought to go in future may change with them. At each stage of our lives we bear a complex relationship to these future selves, since who we will later become is both partially unknown to us and also dependent upon what we now choose to do. And again, the nature of our future choices also depends on our present actions too.

A third aspect of the situated viewpoint is uncertainty. The real world in which we must act is highly complex, and agents only have access to limited information about their environment, which is itself constantly changing, and continues to affect us even if we do nothing.[29] This uncertainty is compounded by the unforeseen consequences of both our own actions – including acquiring further information about the environment – and those of others. Hence, the future in general is not foreseeable in detail.[30] In contrast, when adopting a global viewpoint, contemporary Appraisive theorists typically assume that all the facts of the agent's life may be known with certainty so that assessment can take place. This fits well with the analytic philosophy of well-being's typical focus on abstract, fictitious scenarios where facts about the case at hand may simply be stipulated, just as they are within pure mathematics.[31]

[27]Cf. Dunne, 1993, p. 14 (discussing Gadamer).

[28]Cf. Annas, 1993, p. 93: 'I already *have* a life ... I am already embedded in particular contexts of society, culture, gender, education, and so on.'

[29]Cf. Dunne, 1993, p. 268.

[30]Cf. Pfeifer, Bongard, and Berry, 2011, pp. 25-26 on 'complete agents' that must act in an uncertain world.

[31]See e.g. Bramble's discussion of Nozick's 'experience machine' thought experiment, where the implications of uncertainty are avoided by switching to a global view: 'We can ignore the question of whether we would want or choose to plug in, and instead ask directly *whether it seems best for someone* – either oneself, a loved

We now give some examples of how these points can impact substantively upon our enquiry. In contemporary philosophy, one important type of Appraisive theory comprises 'desire-satisfaction' accounts, which will be discussed extensively in Chapter 2. Though the technical details vary, on all such views what explains how well an individual's life has gone is the extent to which their desires are satisfied. However, as we shall see, theorists advocating this type of theory typically do not work with the agent's actual *de facto* desires, but rather focus on what their desires *would* be like after having been appropriately laundered through certain idealising conditions: 'Thus it is common for desire accounts to stipulate that the only desires that count are those which are sufficiently rational or considered or informed, or otherwise 'corrected'.'[32] For example, one common tactic is to look at what the agent's desires would be like *if* they had 'full information'. However, if these idealisations of the agent's desires are to be of any use for guiding their actual choices then they must be available to the agent themselves – and this puts restrictions on the modifications that an Action-Guiding theory can legitimately make. In particular, it is unintelligible to advise agents to exclude certain types of desire from consideration if these can only be identified in light of information that they do not presently possess.

Another repercussion of this change in perspective can be seen through a comparison with Appraisive assessments that draw on desires that occur at many different points distributed throughout the agent's lifetime. From a situated viewpoint, at any particular moment the details of our future desires may not be known in advance (as just noted, one key reason for this is that they may be within our power to affect through our present actions). Such future desires are therefore not encountered simply as fixed items of reference which we passively take into account in deciding what to do; hence, further modifications to this type of account are also necessary if they are to serve as Action-Guiding theories.

---

one, or a complete stranger – to plug in (or to have plugged in) *in a case where the machine, as a matter of fact, does not malfunction, the premises are not overrun by fundamentalist zealots, the scientists do not turn evil, etc.*' – Bramble, 2016a, p. 139 [emphasis in original]. On pure mathematics, see Lockhart, 2009, p. 25: 'The edges are perfect because I want them to be— that is the sort of object I prefer to think about.'

[32]Sumner, 1996, p. 130; some theorists only rule existing desires out of consideration rather than introducing new ones, whereas others introduce new, hypothetical desires, or even focus on the desires of the agent's 'ideal counterpart'. See Section 2.2 below for detailed discussion.

1.2.2 Choices Under Uncertainty and Outcomes

In addition to facing uncertainty about the effects of our actions, there are invariably also some aspects of our lives that are not within our control at all. Such conditions are typically an important part of a third-person Appraisive account of well-being.[33] However, with an Action-Guiding approach we are specifically investigating the implications of agents' embodied conceptions of the good life for how they make decisions. In particular, the focus here will be epistemological: rather than investigating possible features of human lives whose exhibition makes them valuable, I will instead consider certain normative theories about *how* embodied agents identify goods as objects of choice within their situated practical deliberations (that is, that they should do so by following their Subjective Values). Hence, we will only be concerned with those features of our lives that can actually be affected by such choices.[34] We are not interested in what we might merely wish to happen to us, but rather, how best to direct our energies given the real opportunities we have available; for us 'Things without all remedy, Should be without regard'.[35] This restriction is not as strong as the Stoic injunction that we care only for those goods that we can *guarantee* for ourselves – which on the Stoic view encompasses only certain features of our inner lives: namely, developing our characters and acquiring the virtues.[36] For as a consequence of the aforementioned uncertainty inherent to a situated view, we shall now see that actions whose success is merely probabilistic may need to be considered too.

Another key consequence of this uncertainty is that agents cannot always compare the range of actions currently open to them in terms of their actual induced outcomes, since in some cases these will be unknown

[33]Cf. Cohen's criticism of Sen's capability approach because of an alleged lack of attention to passive benefits derived from goods: Cohen, 1993, p. 23; Williams, meanwhile, distinguishes 'importance' as such from 'deliberative priority' – Williams, 2006 [1985], p. 183.

[34]See Annas, 2011, p. 92 on *'the circumstances of a life' versus 'the living of a life'* as one's mode of dealing with these circumstances, with doing so skilfully being the core focus of ancient ethics; see also Prinzing, 2011, p. 158 for advocation of 'mutability' as a *desideratum* for agents' conceptions of well-being.

[35]*Macbeth*, Act 3, Scene 2. In economics, standard choice theory distinguishes between 'choice variables' and 'externalities', with the former being available for agents to optimise over, whilst the latter affects their utility despite being outside their control; however, if a factor is under *no one's* control then it is ignored and left out of the model. See also Robbins, 1935 on the conditions that must be met to constitute a worthwhile economic problem.

[36]See Epictetus, 2008, p. 6 (I.1), who cites 'the power of positive and negative impulse, of desire and aversion' as within our control.

at the moment of choice. To be of use here, our enquiry must therefore attend to what it is to make good decisions from within such a position of limited information.[37] As we have said, contemporary philosophers of well-being typically concentrate on the evaluation of complete lives whose features are assumed to be known with certainty or simply stipulated in advance; they therefore have no use for the theory of probability. However, within EUT and related theories, it is essential. Here the individual actions focused on (e.g., purchasing a certain amount of insurance) may be thought of as choices between corresponding 'lotteries' that lead to particular concrete outcomes with certain associated probabilities (typically these outcomes are quantitative, and often given in terms of money). A normative guide to action is then arrived at by assigning possible outcomes a utility score and then using probability theory to maximise over expected utility.[38]

As well as enabling economists to address a much wider range of scenarios, attention to cases involving uncertainty has also provided a powerful justification for their underlying conception of the Human Good. As we shall see in Chapter 2, in 1947 von Neumann and Morgenstern proved mathematically that *any* agent whose preferences over *lotteries* satisfy certain reasonable-sounding conditions will indeed behave *as if* they are maximising the expected value of a utility function over *outcomes*.[39] By taking a situated perspective that includes choice over uncertainty, then, the available modes of justification are extended in a way that has substantively impacted their enquiries.

Although von Neumann and Morgenstern's achievements are impressive here, however, we should also note at the outset that mainstream economists' formal models of choice under uncertainty embody only one highly specific and peculiarly modern way of addressing our central task

[37]For instance, taking a risky gamble may still be considered a poor *choice* even if it happens to lead to a good *outcome.* In economic parlance, we choose *ex ante* and not *ex post*; for the economic importance of this distinction, see e.g. Motta, 2004, p. 65. For an illustrative contrast between philosophical and economic perspectives here, see the interchange between Broome and Jones-Lee on *ex ante* vs *ex post* approaches to cost-benefit analysis: Broome, 1978; Jones-Lee, 1979; and Broome, 1979; see also Sumner, 1996, pp. 129-132, who insists that an Appraisive theory must focus on *ex post* evaluation, and criticises the prospectivity of desire theories in this respect.

[38]See Hansson, 2018, Section 5.1 for criticism of this standard division of labour, wherein axiology first provides assessments of definite outcomes and decision theory then combines these with probabilities to arrive at optimal actions, since the 'moral aspects of risk-taking' are thereby left out of account.

[39]Neumann and Morgenstern, 2004 [1953], pp. 617-628; see also Section 2.1 below for detailed discussion.

– and one that is not without its problems. To take one example, human beings are limited not only with respect to our knowledge of the environment, but also in our capacity to perform the calculations necessary to use the information we do have in an optimal way. Until quite recently, most economists have largely ignored this point – yet as behavioural economists now argue, allowing for considerations like this is important for both normative and descriptive enquiry into choice, in order to make contact with what real-world agents can actually do, rather than addressing only the mythical *homo economicus*. What results is the study of what Herbert Simon has called 'bounded rationality', and he explains his approach to the description of human choice behaviour as follows:

> Broadly stated, the task is to replace the global rationality of economic man with the kind of rational behavior that is compatible with the access to information and the computational capacities that are actually possessed by organisms, including man, in the kinds of environments in which such organisms exist.[40]

Since we are seeking an Action-Guiding theory that can likewise be applied to human beings in the real world, such considerations must guide the present enquiry too.

1.2.3 Self-Effacingness

We now consider another issue that arises in moral philosophy more generally: whether it is permissible for a theory to be 'self-effacing'. In her insightful discourse on the virtues, Julia Annas defines a moral theory to be self-effacing if it 'tells us to achieve the aim of the theory not by aiming at it but precisely by *not* aiming at it, by not doing what it bids and by not becoming the kind of person it recommends, but doing, and aiming at being, something entirely different.'[41]

[40] Simon, 1955, p. 99. For other early work in this vein, see also Coase, 1937 on 'transaction costs' that had previously been neglected; and Simon, 1947 for discussion of the more limited, situated, satisficing 'administrative man'. For an overview of some later work, see Kahneman, 2003; and for a classic and influential response on behalf of the mainstream tradition, arguing for the value of seeking highly simplified models that nevertheless make accurate predictions, see Friedman, 1953, p. 14.

[41] Annas, 2011, p. 155; see also Annas, 2007, pp. 210-212; Parfit, 1984, pp. 23-24, who defends this as a coherent possibility; and Stocker, 1976 for criticism of the splitting apart of reasons and motives in modern philosophy.

The charge of being self-effacing has perhaps most often been levelled at utilitarianism.[42] The classical proponents of the theory – Bentham, Mill, and Sidgwick – understood the Human Good in terms of pleasure, but the so-called 'hedonic paradox' suggests that by aiming 'directly' at pleasure we are often less likely to achieve it.[43] However, although Bentham did recommend that a hedonistic conception of welfare be the dominant Action-Guiding ideal in just this way – writing that 'Nature has placed mankind under the governance of two sovereign masters, *pain* and *pleasure*. It is for them alone to point out what we ought to do, as well as to determine what we shall do'[44] – later theorists including Mill and Sidgwick were more keenly aware of the issue. Though Mill seems to oscillate on the practical role of the core utilitarian doctrine,[45] Sidgwick responds to the issue by positioning it as an abstract standard of rightness that applies to actions but that need not lend itself to a direct operational role in the ordinary practical deliberations of plain persons: 'it is not necessary that the end which gives the criterion of rightness should always be the end at which we consciously aim'.[46]

In a well-known critique, Bernard Williams complains that when utilitarianism is thus allowed to become increasingly indirect, then ..

> .. there seems nothing to stop, and a lot to encourage, a movement by which it retires to the totally transcendent standpoint from which all it demands is that the world should be ordered for the best, and that those dispositions and habits of thought should exist in the world which are for the best, leaving it entirely open whether those are themselves of a distinctively utilitarian kind or not.[47]

This raises the issue of what the purpose or goal of such theorising

[42] See Annas, 2011, pp. 157-159.

[43] See Section 2.4 for a critical discussion of the argument.

[44] Bentham, 1970 [1789], p. 11.

[45] See Mill, 1998 [1863], pp. 64-65 for a passage emphasising the principle of utility as an abstract criterion of rightness as distinct from motive; in contrast, for utility as an active guiding principle that can be used to settle rival claims of the heuristic 'secondary principles' actually employed by plain persons, see *ibid.*, pp. 71-72.

[46] Sidgwick, 1981 [1907], p. 413; see also *ibid.* p. 490, where he entertains the view that 'that the vulgar should keep aloof from his system as a whole'. However, the preface to the first edition instead indicates a core focus on 'how conclusions are to be rationally reached in the familiar matter of our common daily life and actual practice' (*ibid.*, pp. vii-viii). See also the discussion of Mill in Section 1.7 below.

[47] Smart and Williams, 1973, pp. 134-135; see also Williams, 2006 [1985], pp. 106-108, which also mentions R. M. Hare's two-tier system.

can be, if it makes no direct contact with practical action.[48] Sidgwick sometimes attempts to assume the role of a disinterested natural scientist – but Williams, pointing to the 'important colonialist connections of utilitarianism', contends that what Sidgwick was really proposing was a 'Government House' morality that addressed itself to a political elite with its own coercive and possibly suspect purposes.[49] With some more recent 'indirect' utilitarian theorists, in contrast, he urges that the view now has no social location at all: it is 'transcendental to life, existing in a space quite outside the practice it is supposed to regulate or justify'.[50] Either way, and as with the Ancients, Williams can see no proper place for a moral theory other than being directly an account of what we should *do*.[51]

Returning now to our present enquiries into the Human Good; from an Appraisive perspective, it may be unreasonable to rule out self-effacing theories *a priori*, since the theoretical criteria which furnish philosophers with a means to judge whether a life goes well may be quite different from the principles that should actually guide plain persons if they are to effectively achieve these same ideals in practice; it seems these are indeed distinct and separable questions. In contrast, with our goal of seeking an Action-Guiding theory things are quite different. It is clear that self-effacing theories *must* be impermissible here, since the primary purpose of the conception of the good life sought is specifically to guide agents' embodied practical reasoning. Insofar as this is our key focus, there is no possibility of a gap between an independently 'true' theory of the Human Good and what agents should aim at in practice.[52]

1.2.4 Moral Philosophy and Empirical Psychology

A further implication of aiming for an Action-Guiding theory is the extent to which our investigations are liable to spill over from moral philosophy into other domains of enquiry. Here our goal is not to give a logical

[48]Cf. Annas, 2011, p. 157: '*by whom* is the theory's aim to be achieved?'

[49]Williams, 2006 [1985], pp. 108-110; cf. Sidgwick, 1981 [1907], p. 490: 'it may be desirable that Common Sense should repudiate the doctrines which it is expedient to confine to an enlightened few.'

[50]Williams, 2006 [1985], p. 110; see also Sen and Williams, 1982, p. 15: 'utilitarianism needs *some* embodiment'.

[51]Cf. Annas, 2011, p. 158: 'We are owed an account of how we can still understand it as an *ethical* theory'.

[52]*Pace* Parfit, 1984, p. 24. Some possible problems with taking this attitude are addressed in Section 1.3 below.

analysis of a concept, but rather to say something useful about a problem everyday moral agents face in their lived experience – and as we shall see, making sense of how things are for us here can be illuminated by work in relevant empirical disciplines. Indeed, it may be best not to see our central task as based within a single, isolable area of enquiry at all.[53]

It is often remarked that the ancients draw no sharp distinction between moral and political philosophy,[54] but for Aristotle a full understanding of *eudaimonia* requires that we throw psychology into the mix too: 'clearly the student of politics must know somehow the facts about soul, as the man who is to heal the eyes must know about the whole body also'.[55] And well into the early modern period, enquiries into moral philosophy are generally thought of as inseparable from the search for a correlate account of human nature. By the time we reach Kant, however, a change of attitude has occurred. 'Anthropology' is recognised as essential for the *application* of moral theory: 'a metaphysics of morals cannot dispense with principles of application, and we shall often have to take as our object the particular *nature* of human beings, which is cognized only by experience'.[56] The practical aspects of moral motivation, meanwhile, are still a key concern.[57] However, the *content* of the fundamental abstract moral imperatives at the heart of his system are to be justified strictly *a priori*: 'moral principles are not based on what is peculiar to human nature'.[58] And later, with Sidgwick we meet the claim that the study of moral psychology 'no more properly belongs to Ethics than the corresponding questions as to the cognition of Space belong to Geometry'.[59] The Appraisive approach to well-being

[53]Cf. MacIntyre, 2011 [2007], xviii; for an introduction to the Appraisive study of well-being within social science from the perspective of philosophy of science, arguing for the mutual interdependence of responsible empirical science and useful normative philosophising, see Alexandrova, 2017; for further calls for a more inclusive, interdisciplinary approach, see again Prinzing, 2021; and Bishop, 2015.

[54]MacIntyre, 2011 [2007], p. 161.

[55]Aristotle, *NE* I.13, 1102a17-20.

[56]Kant, *Metaphysics of Morals,* 6:217 [all references to Kant are from the Cambridge complete works]. See also Kant, *Groundwork*, 4:388-399; 4:412; for the importance of practicability as such, see Kant, *On the Common Saying: That May Be Correct in Theory, But it is of No Use in Practice*, 8:273-313; and Louden, 1992, pp. 100-103 for discussion of Kant's view.

[57]Kant, *Groundwork*, 4:449-450.

[58]Kant, *Groundwork*, 4:410 [footnote]; see also *ibid.*, 4:425; 4:442; and especially 4:388-390: 'That which mixes these pure principles with empirical ones does not even deserve the name of philosophy'; and *Metaphysics of Morals*, 6:215.

[59]Sidgwick, 1981 [1907], p. viii (Introduction to the First Edition) – though in the third edition of 1884, Sidgwick indicated in a footnote that this statement requires 'slight modification'. Similar influential views were offered by other moral

within contemporary analytic philosophy has again followed his lead here: investigation into what it is for a life to go well now tends to be carefully insulated from empirical enquiry.[60]

The core goals of this project are normative rather than descriptive, and I will aim to carve out a special, partially autonomous domain for ethical enquiry, which as I conceive it has quite different goals to natural science. But nevertheless, in order to accomplish the key objectives of both of the two chapters that follow we will need to attend to relevant empirical studies, and in addition to modern and ancient philosophy and normative economics the enquiry will therefore draw on a number of other literatures.

Firstly, we will need to elaborate our proposed Subjectivist solutions to our core problem by exploring what an agent's life would be like if the corresponding modes of practical reasoning were to be consistently adopted – assuming this to be possible – and thus discover what powers of mind and character would be required for them to do so effectively.[61] To this end, in Chapter 2, the many recent empirical experiments conducted into decision-making in psychology and behavioural economics will provide insights into what an agent acting under the auspices of one of our purely subjectivist views of the Human Good would actually be like.[62] This will also lead us to discover what powers of mind and

philosophers around the middle of the 20th Century, including Hare and Stevenson. For a contrary call for ethicists to study 'philosophical psychology', see Anscombe, 2000 [1957], p. 78; and for further classic criticism, see Dewey, 2002 [1922], pp. 10-13; and pp. 295-296; see also Dewey, 1998, Vol. 1, pp. 109-111, and Dewey, 2002 [1922], p. 296 on the potential contribution of science to the theory of values. Finally, for a historical view, see MacIntyre, 1982, who dates the key division to the 1780s, in Kant and Thomas Reid, who were striving to avoid entangling their purified, rationalistic conceptions of duty and agency with a Humean, Subjectivist conception of practical reason based on the passions. This issue has also stimulated recent philosophical debate, which I cannot go into here.

[60]For instance, Feldman often seems to have little interest in the empirics of pleasure or pain; see Angier, 2015, p. 15 for a critical discussion – though see Feldman, 2010, Part 3. For further exceptions to this trend within the contemporary philosophy of well-being, see Tiberius, 2006, p. 493-505, which draws on empirical resources to argue for a life-satisfaction view; Crisp and Kringelbach, 2018, pp. 211-215, which discusses neuroscientific work on 'liking' as the basis of pleasurable experience; and Haybron, 2016 on operationalising 'eudaimonic' measures of well-being for psychological research.

[61]Cf. MacIntyre, 2011 [2007], p. 27: 'we have not yet fully understood the claims of any moral philosophy until we have spelled out what its social embodiment would be'; on this topic, see also MacIntyre, 1998b, p. 258.

[62]For some existing empirical research on plain persons' conceptions of well-being and the impact of these on outcomes, attending to 'hedonism and eudaimonism' and offering provisional support for the position that will be argued for here, see McMahan and Estes, 2011a and 2011b.

character would be required for human agents to thus implement these theories in practice.

Secondly, having set out what the social embodiment of our Subjectivist conceptions of practical reason would be like, in Chapter 3 we will turn to their evaluation. Rather than following the now-dominant but somewhat superficial method of comparing our accounts to what ordinary people presently feel inclined to *say* about what is good for them, I will instead focus on our initial development as rational agents in practice, and in particular on the prerequisites that must be in place for us to develop the powers necessary to mobilise the resources for deliberation that our two theories offer. Through drawing on the literatures of developmental psychology, social psychology, and medical and therapeutic ethics, I will argue that supportive others must take a central role in the development and maintenance of these core rational powers; our Subjectivist theories will then be thereby criticised as being too individualistic to be viable in practice.[63]

1.2.5 Precision, Completeness, and Generality

Lastly, a few preliminary remarks are in order about the form we should expect a successful Action-Guiding account of the Human Good to take, beginning with the degree of precision we will require.[64]

As indicated above, most contemporary philosophers of well-being working in the analytic tradition today aim to arrive at Appraisive criteria for well-being that are stated very precisely – typically, in the form of individually necessary and jointly sufficient conditions. Our Subjectivist theories then offer various reductive accounts of these conditions – some of which are highly technical.[65] Meanwhile, contemporary microeconomics also makes claims to have attained great precision, being presented in the guise of the formal mathematical theory. Economists have achieved this by abstracting what they have taken to be the key features of rational

[63]The reader may notice a lack of a systematic literature review, which would not be possible due to the extent of the ground covered; rather, I have sought to make use of the findings that seem to me to make the most sense, that are the subject of some consensus, and that have robust empirical support. See Churchland, 1989, p. x for defence of the same strategy for the type of interdisciplinary 'synthetic philosophy' I engage in – in the sense of this term given by Schliesser, 2019, following Herbert Spencer's coinage.

[64]See Sections 1.5-1.6 for a fuller discussion of the more general issue.

[65]See, for instance, Feldman's detailed discussions of the concept of pleasure; Feldman, 2004, Chapters 2-4.

agency, and then using this minimal framework as a basis for developing formal models of singular circumscribed choice problems, such as buying insurance.

Despite the aspirations thus held by two of our key approaches, however, as the enquiry progresses we will find that the variety and complexity of life in the real world entail that such precision will be difficult to achieve whilst holding on to the goals underlying the framework we are developing. For instance, from Walras onwards, economists' successful development of a powerful and elegant mathematical theory has required quite drastic simplifying and idealising assumptions – which as we have seen already have led to a neglect of important features of real-world human choices, and which contemporary behavioural economics is only just now beginning to correct.[66] Given the complexity of the general human problem of how to act and how to structure our lives as a whole, then, in characterising our two Subjectivist accounts we will instead come to settle on a fairly rough-and-ready account of their overall direction rather than hoping for precise technical implementation – the limitations of which will become clear in Chapter 2.[67] Likewise, whilst examining the limitations of our two Subjectivist views by comparing them to alternative modes of practice in Chapter 3, we will find it better to follow the ancients and settle for getting things right only roughly and in outline. The point is again well put by Aristotle:

> Our discussion will be adequate if it has as much clearness as the subject-matter admits of, for precision is not to be sought for alike in all discussions ... it is the mark of an educated man to look for precision in each class of things just so far as the nature of the subject admits: it is evidently equally foolish to accept probable reasoning from a mathematician and to demand from a rhetorician demonstrative proofs.[68]

The critical nature of the present enquiry also engenders another reason for us not to adopt a precise, technical approach to the subject matter. A more conventional way of proceeding would be to being by to elaborating precise definitions of the concepts most relevant to the argument, such as 'deliberation', 'rationality', 'pleasure', and 'desire'. Yet if we followed this

[66]See Bruni and Sugden, 2007 for a historical view.

[67]See especially Section 2.1, on decision theory.

[68]Aristotle, *NE* I.3, 1094b13-28; see Dunne, 1993, pp. 243-244 for a discussion of Aristotle's view here.

route, our later argument could be sidestepped by any theory which took a different view of their nature – and indeed, any attempt to rigorously define these notions is bound to be somewhat contentious. In addressing a wider range of Subjectivist theories, and in allowing the contributions of other types of view to remain visible, then, we will therefore find it best to leave these matters somewhat open for the time being – though some of these notions may later take on more a definite shape as we home in on specific Action-Guiding theories.

A related feature a theory might have is what I shall call 'comprehensiveness': that it renders a determinate answer for every case to which it may be applied. Within the contemporary philosophy of well-being, for instance, the Appraisive conditions encompassing well-being that are sought are often expected to be capable of being applied operationally, returning an unequivocal evaluation of any conceivable human life whatever when it is presented in abstract terms.[69] Likewise, EUT aspires to comprehensiveness in the context of rational action, insofar as any choice that can be formulated as a lottery over outcomes can in principle be settled by the maximising decision procedure offered by the framework.[70] This has in turn placed requirements on the underlying formal account of the Human Good too: agents' preferences must be 'complete' in order to be considered rational by standard theory, in the sense of yielding a comparison between *any* two possible outcomes.[71]

In contrast to this modern work, in her authoritative account of ancient ethical theory Annas demonstrates that again comprehensiveness was not an ambition that the Greek philosophers strove to achieve in their enquiries into the good life; instead, a great deal was left up to the situated practical judgement of the individual moral agent.[72] Moreover,

[69]See e.g. Feldman, 2004, p. 13: 'Ideally, I would like to find a principle that would yield a ranking of lives' – where this ranking is based on the systematic numerical aggregation of values assigned to discrete elements; see also Griffin, 1986, pp. 2-3, who advocates 'completeness': that a theory can yield 'an answer to any moral question that presents itself'.

[70]Things are complicated somewhat when we consider multiple agents acting together; e.g., certain games have no Nash Equilibria. On the expansion in the 1960s of the scope of economic choice models into such unfamiliar territory as whether it is rational to engage in criminal behaviour, following Robbins' influential topic-neutral definition in terms of the general allocation of scarce resources, see Backhouse and Medema, 2009, pp. 229-230.

[71]Cf. again Section 2.1 below for discussion.

[72]Annas, 1993, pp. 6-7; p. 443; cf. Aristotle, *NE* II.9, 1109b23: 'the decision rests with perception'. See also MacIntyre, 2016, p. 51 on Aquinas: 'in the life of practice there are no fully adequate generalizations to guide us, no set of rules sufficient to do the work for us.'

she further suggests that this aspect of the contemporary philosophical approach may be a mistake resulting from the influence of inappropriate hypothetico-deductive models drawn from the philosophy of natural science.[73] In Chapter 2, I will argue in some detail that we should again follow the ancients here, aiming only for the general guidance of action in broad terms rather than a deterministic algorithm for what to do in all cases, and that individual perception and judgement must continue to play a central role.[74] Moreover, rather than aspiring to an entirely transcultural investigation, I shall also primarily restrict myself to rational action within a particular cultural and historical setting within which the theories we will investigate have flourished.[75]

Another consequence of taking a more relaxed attitude to both precision and completeness is that we will be less interested in fanciful and unrealistic case studies that tend to populate the contemporary philosophical literature on well-being.[76] Such thought experiments are perhaps fair game when seeking an exact, rigorous, comprehensive statement of the 'logical essence' of a concept – but here our subject matter is instead the guidance of human action and life in the real world. And since real agents only require the resources to choose between the kinds of actions and lives that are actually available to them, their attitude toward such conceivable but purely fictitious scenarios will likely be that 'It isn't so, you know, so if you please we won't suppose it'.[77]

At the same time, the practical focus of the project will ensure its relevance to the problems of real life in all their complexity, rather than

---

[73] Annas, 1993, p. 7; p. 11; p. 442; again, Sidgwick has exerted a lasting influence here. On the enduring influence of mathematics and physics on philosophy, see Dewey, 1998, Vol. 1, p. 21.

[74] See Louden, 1992, pp. 102-103, wherein he ascribes this view to Kant, and *ibid.*, pp. 92-95 on decision procedures *versus* Action-Guiding accounts more broadly. See also Nussbaum, 2001 [1986], pp. 298-306 for the same view, emphasising flexibility and resourcefulness; Dunne, 1993, pp. 33-35 for an endorsement of this attitude toward comprehensiveness in the context of Newman's thought, based on the Aristotelian idea that *phronesis* is always essential in any complex situation; and *ibid.*, p. 15 for a similar discussion in the context of Gadamer's views.

[75] See the end of Section 1.3 below for further discussion of these limitations of scope.

[76] For instance, Rawls' grass-counter (Rawls, 1971, pp. 432-433) and Nozick's experience machine (Nozick, 1974, pp. 42-43); see also O'Connor, 2012, who argues that this approach to moral philosophy not only lacks practical value but is also actively harmful because it encourages a heavily impoverished view of human agency. Unfamiliar and unrealistic cases are also reported to generate 'imaginative resistance' and invite biases in our intuitive judgements; see De Brigard, 2010, Weijers, 2013, and Weijers, 2014.

[77] Geach, 1977, p. 64 – paraphrasing Dickens' Tommy Traddles.

resting content with an artificial, purely formal hollowness:

> Compelling attention to details, to particulars, it safeguards one from seclusion in universals; one is obliged, as William James was always saying, to get down from noble aloofness into the muddy stream of concrete things.[78]

The above observations about the variability of human life may, however, point to another issue: that of over-generalisation. For so far, I have made it seem as though the central practical problem of how to act and live which all human beings face – a problem whose characterisation will occupy us until the end of the chapter – is always one and the same; yet perhaps instead human beings are actually each confronted with quite different problems, with quite different attendant choices and possibilities, by virtue of the particular circumstances of their lives – social, political, cultural, historical, and also personal:

> People will speak of a savage as 'confronted by the eternal problem of obtaining food'. But what really confronts him is the problem, quite transitory like all things human, of spearing this fish, or digging up this root, or finding blackberries in this wood.[79]

We may therefore have not just one general problem but as many specific and local ones as there are individual human beings. Worst, still: what it is best for me to do, how I should live, and whom I should strive to become may all depend upon who I now am and what my particular abilities and latent talents are – and if we extend this list to 'preferences' or even 'tastes' then a Subjectivist account comes to seem quite appealing. My brother and I may face the same external circumstances and enjoy the same material opportunities, but that does not mean the same choice of path will suit us equally well: 'This – is just *my* way:– where is yours? Thus I answer those who asked of me 'the way'. For *the* way – does not exist!'[80] But then what role is left for philosophy?

Though the point is well taken, and a concrete, one-size-fits-all model of the good life as sought by some of the ancients may indeed be a chimaera, we shall see that there are other possible tasks for philosophers to accomplish here. For instance, MacIntyre suggests one view that is

[78] Dewey, 1998, Vol. 1, p. 30.
[79] Collingwood, 2013 [1939], pp. 32-33.
[80] Nietzsche, 2005 [1892], p. 169.

consistent with acknowledgement of this diversity: that philosophy might not aim to deliver a single homogenous answer for how humans are to live that applies in all circumstances, but instead focus on what all adequate answers must have in common.[81] Moreover, despite their situatedness and particularity, ordinary moral agents find themselves to at least some extent engaged in a more general enquiry of the kind attempted by philosophy here whenever they come to critically reflect on the answers and assumptions about their own good that they have found for themselves and relied on hitherto.[82] Room must be left for individual differences, and this is another reason not to aim for completeness; yet I shall argue that we share enough common heritage to make helpful generalisations that do apply to humanity as a whole.[83] The methodological underpinnings of the enquiry will be discussed further in the next section.

1.3 Guiding Practical Reason and Normative Constructivism

In a description of his experiences at archaeological digs, the philosopher and historian R. G. Collingwood explains that any competently-planned excavation of a site will always set out to answer some specific and determinate set of questions concerning an aspect of the society whose former dwelling it marks. It does not do to simply turn up, have a look around, and see what happens to be uncovered.[84] Likewise, in his philosophical work, Collingwood also advocated a move away from conventional propositional logic to an alternative 'logic of question and answer' wherein both the meaning and veracity of a proposition always depend on the question to which it is intended to be a response.[85] He elaborates:

> a body of knowledge consists not of 'propositions', 'statements', 'judgements', or whatever name logicians use in order to designate assertive acts of thought ... but of these together

[81] MacIntyre, 2011 [2007], p. 253. See also Dewey, 1998, Vol. 1, p. 25: 'Search for a single, inclusive good is doomed to failure'.

[82] Cf. MacIntyre, 1992, pp. 3-4; MacIntyre, 2009b [1999], p. 157.

[83] Though again, see Section 1.3 on the cultural particularity of the present critical argument, based on the special shape the problem takes on within the present social order, as well as certain assumptions we may adopt because they are also made by the theories I am challenging.

[84] Collingwood, 2013 [1939], pp. 24-25; p. 122.

[85] Collingwood, 2013 [1939], pp. 29-43.

> with the questions they are meant to answer ... a logic in which the answers are attended to and the questions neglected is a false logic.[86]

Though I offer no such general logical theory, the attitude underlying the present critical argument will be similar. Rather than asking whether our two Subjectivist theories are in some sense 'true' accounts of the Human Good, in and of themselves, what underlies my argument here is a logic of *Problem and Solution*. Our core concern, as we have said, is with a problem that we as human beings all face, which comprises two connected aspects – how we should rationally deliberate about our actions within concrete situations in the here-and-now, and how we should rationally decide what our lives as a whole should be like.[87] After characterising this problem further, we will then consider how our two Subjectivist theories can be brought to bear as potential solutions. And for us it is *only* their success in this capacity that is relevant to their evaluation – just as the evaluation of a medicine is based solely on its capacity for mitigating the corresponding disease.[88]

The project is thus teleological in character – though the teleological aspects are not located within the subject matter under investigation, whether natural or otherwise, but rather built into the aims of the enquiry itself: it is not metaphysical, but consists only in being clear about the purpose of what we are doing. The theory of the Human Good we seek will find its point and purpose in guiding human action – and as we shall see, this assumption will be appealed to throughout the argument in characterising what any adequate solution to our core problem must be like.

At this point, it is natural to ask: How exactly does such a conception of the Human Good fit into a broader account of practical reason? As a first approximation, to consider a particular item good in an Action-Guiding sense is to regard it as *choiceworthy* in our practical deliberations – where such an item may be an object, an outcome, an action, or an

[86] Collingwood, 2013 [1939], pp. 30-31; see also Somerville, 1989 for discussion.

[87] See Section 1.4 below for more on these dual aspects and the relation between them; again, the relation of these undertakings in their developed form to other types of human behaviour will not be considered in the present work – though in Chapter 3 we shall be greatly interested in the genesis of our capacity to engage in them.

[88] That is, ignoring side effects. Cf. Anderson, 2018 [2005], Section 2.5 on Dewey: 'The form of a contextual standard of value is: it solves the problem encountered in this situation'; and Dewey, 2002 [1922], p. 199; see also Griffin, 1993, p. 135 and Scanlon, 1993, p. 185, for similar views.

entire way of life.[89] This very rough conception of the role of theory will be expanded on later in the chapter, once an important distinction that will be drawn in the next section is in place; we now pause to consider an objection to what has been said so far.

Our enquiry has now received its *telos*; yet if what I claimed earlier about the importance of this goal for how our enquiry will turn out is correct, one might reasonably ask: 'Is an opening move like this legitimate?' Rather than being a contribution to the philosophical study of well-being as it is currently conducted, am I not just changing the subject by addressing an entirely different question? In response to this objection, I offer a few thoughts of a pragmatic nature in support of this unorthodox move.

Firstly, let us for a moment step back from the assumptions that have become dominant with contemporary philosophy today, and ask: 'What do we actually want a theory of the Human Good *for*?' If we regard this need as being posed in broad social terms rather than the narrow context of academic philosophical enquiry, then I believe it is Action-Guiding approaches that have the most to offer here. For the kind of abstract and abstruse theorising that the Appraisive theorist offers can be engaged in only in quite peculiar circumstances, whereas the question of what to do and how to live receive at least an implicit answer in every lived human life.[90] It would therefore be highly surprising if these questions were thought not to be worthy of attention for their own sake.

A second possible response here is to argue that the capacity for judgement and the conceptual resources brought to bear within the current analytic philosophy of well-being themselves ultimately derive from our experience with practical deliberation about action; that 'We are all of us agents before we are theorists, and it is only because we are agents that we have subject matter about which to ask those questions that take us into theory.'[91] It would then only be our experience – *qua* human beings

[89] Cf. Collingwood, 1992 [1940], pp. 419-422; and Hurka, 1987 for an analysis of 'good' in terms of what one ought to pursue. Williams argues that we do not always choose under the 'guise of the good', and understand that 'the merits of the thing in question may go beyond [our] own interests or power of response' – Williams, 2006 [1985], p. 125; see also *ibid.*, p. 58. However, such choices would then simply be outside of the scope of the present enquiry, on Action-Guiding theories of the good in particular.

[90] Cf. MacIntyre, 2011 [2007], p. 138; MacIntyre, 1992, pp. 3-4; Annas, 1993, pp. 27-29.

[91] MacIntyre, 2016, p. 72

– with the practical problem at the centre of Action-Guiding enquiry that enables us – *qua* philosophers – to engage in such Appraisive enquiry at all, and the oft-unquestioned 'intuitions' about the good life that form such an important part of the methodology of the contemporary philosophical approach would ultimately derive from our practical experience of choosing and of acting upon our choices. If this were accepted, then the practical problem I will focus on would thus have a certain priority over the Appraisive project.[92]

Thirdly, what I am doing is in many ways a return to an older mode of enquiry; if we do change the subject, then, we only change it back to one with a more established history. Indeed, there is no shortage of historical philosophers who have also aspired to a broadly Action-Guiding approach – for this one must only venture outside the analytic literature and its immediate progenitors: as well as the ancients, we might point to Augustine, Aquinas, Petrarch, Montaigne, Hume, Adam Smith, the existentialists, and many others within the Western canon alone.[93]

We now move on to discussing an important consequence of the present methodological framework: that is, its potentially revisionary nature. For in following the logic of Problem and Solution sketched above, I am not trying to *explain our existing* conceptions of the Human Good naturalistically through appealing to a falsifiable hypothesis about *their* purpose or developmental origins.[94] Indeed, the target conception of what is good for us is not something given in advance of our enquiry; 'Our job is not to describe an idea already in existence independently of our search'.[95] Instead, the goal is to use the resources provided by our Subjectivist accounts to attempt to *construct* a concept of the Human Good that is suited to helping with the particular practical problem which I have identified as worth focusing on for its own sake – a paradigm that I will call 'Normative Constructivism'.[96] Yet this may require a departure from the ideals we currently hold: if they turn out not to be

[92]*Contra* Feldman, 2010, p. 8 (quoted in Section 1.1 above).

[93]Cf. MacIntyre, 2011 [2007], p. 27.

[94]Cf. Craig, 1990 for an enquiry of this other sort in the context of our concept of knowledge, based on its role in flagging approved informants; see also Pettit, 2018 for an attempt to likewise account for morality naturalistically, via a 'rational reconstruction' of its origins which emphasises the importance of ethical integrity for achieving personhood.

[95]Griffin, 1986, p. 1; cf. MacIntyre, 2016, p. 50 on some important ends not being specifiable *a priori*.

[96]For an overview of recent philosophical work that also takes this attitude, under the label of 'conceptual engineering' – much of which draws inspiration from Carnap's notion of an 'explication' – see Cappelen, Plunkett, and Burgess, 2020.

suitable in this context, it is now these conceptions themselves that must be modified or even cast aside in pursuit of more adequate ideals.[97]

The investigation will thus require us to reject at the outset Sumner's 'Criterion of descriptive adequacy', which asserts that an account of well-being must respect 'our' everyday usage of and intuitions about the concept, so that we can assess 'candidate conceptions for their fit with our ordinary experience of welfare and our ordinary judgements concerning it'.[98] In contrast, on the approach pursued here, theories of well-being 'are to be assessed solely for their *normative adequacy*', so that 'the only criterion of adequacy for a candidate conception will be its ability to play its designated role within the framework in question'.[99]

This point is especially important because although the project will offer sustained criticism of our two Subjectivist accounts of the Human Good, at the same time I do accept that such theories may best encapsulate current everyday thought and feeling on the matter.[100] What such views fit especially well with in the present social order is a widespread type of individualism, dating from dramatic social changes in perhaps the late 18th Century onwards, that emphasises individual freedom of expression and self-determination.[101] A correlate image of moral agency has now also become prominent, and contemporary views of the self also tend to be highly individualistic: atomistic, independent, and self-directing.[102] And in consequence, it is now widely believed that the resources we need to determine the best way for us to live do indeed reside internally within ourselves, *qua* individuals. Such a conviction was indeed already being expressed by Mill, who writes that 'If a person possesses any tolerable amount of common sense and experience, his own mode of laying out his existence is the best, not because it is the best in itself, but because

[97]Again, in contrast to Craig's naturalistic approach: 'it is not the idea to construct an imaginary concept, but to illuminate the one we actually have, though it be vague or even inconsistent' – Craig, 1990, p. 2.

[98]Sumner, 1996, p. 8.

[99]*Ibid.*, p. 8. See also Griffin, 1986, pp. 40-41 for the present attitude in the somewhat narrower context of well-being: 'instead of starting with a notion that we might later fit into morality, we start with morality and ask which notion fits it'.

[100]Though for an experimental philosophy argument for the view that the received notion has Aristotelian elements, see Braddock, 2010.

[101]See Taylor, 1992, pp. 25-29 for a brief sketch of the emergence of such a view through Descartes' epistemological and Locke's political atomism, Rousseau's appeal to the 'voice of nature' beyond society, and in the writings of Herder; for the view that a Subjectivist understanding of practical reason is perpetuated by the broader structure of our present society, see MacIntyre, 1987.

[102]For the view that conceptions of the good life are bound up with corresponding views of the self, see Taylor, 1989, p. 105.

it is his own mode.'[103] And it thus becomes imperative that our ideals for how to live are not be imposed by society, nor by outside forces in general: 'Not only should I not fit my life to the demands of external conformity; I can't even find the model to live by outside myself. I can only find it within.'[104]

Given these features of the background culture, it is quite natural that many contemporary Western philosophers are also drawn to adopting the Subjectivist view that 'individuals are the ultimate authorities concerning their own welfare'.[105] And since the primary mode of argument within analytic philosophy is to appeal to shared 'intuitions' whose social and historical roots are too often left unexamined, and the immediate conceptual resources of its protagonists are shaped within the advanced capitalist societies of the West, it is not surprising that widespread acceptance of Subjectivist accounts of well-being should follow in academic circles too.[106]

Despite the dominance of Subjectivism in the present social order, however, many other societies have not shared these inherited assumptions about what a good human life can be. Compare now the values described above to the conception of the good life in Homer's description of the archaic Greek heroic age: to discharge one's social responsibilities well, to show loyalty and attain success in the contests of life, and above all to exhibit the virtue of bravery (think also of ancient Sparta here too).[107] Again, in ancient Athens, when male human beings were firstly

[103]Mill, 1989 [1859], p. 67; see also *ibid.*, Chapter 4 for development and defence of this liberal view. The relation of this position to his official hedonism about well-being is unclear. In politics, compare the saying – sometimes attributed to prime minister Sir Henry Campbell-Bannerman – that 'self-government is better than good government'; in economics, see Mishan, 1969, p. 172: 'economists are generally agreed – as a canon of faith, as a political tenet, or as an act of expediency – to accept the dictum that each person knows his own interest best.'

[104]Taylor, 1992, p. 29. For a summary of the empirical literature on how exposure to individualistic *versus* collectivist cultures influences behaviour, see Sapolsky, 2017, pp. 273-282; see especially p. 275 describing a study showing a preference – as measured by stress hormone levels induced in recall – of Americans for influencing others over being influenced, whilst in East Asians, this preference is reversed.

[105]Sumner, 1996, p. 171; cf. Harsanyi, 1977, p. 645; for argument for the contrary view, see MacIntyre, 2009b [1999], p. 71.

[106]Cf. MacIntyre, 2011 [2007], p. 2: 'the techniques of analytical philosophy are essentially descriptive and descriptive of the language of the present at that'; see also MacIntyre, 2016, p. 102 on capitalism and consumerism promoting Subjectivism, and MacIntyre, 1998a, pp. 58-60 on ideology as theory that privileges widespread but ultimately mutable social practice. Behavioural economists, meanwhile, take themselves to have shown that many of our 'ordinary judgements' about what is good for us are in fact irrational; see here Chapter 2 below.

[107]See Siedentop, 2014, p. 38 on Republican Rome and Sparta as concerned with

citizens rather than consumers, the picture of the ideal life looked very different: here it was the exercise of the moral and intellectual virtues through civic participation which was central.[108] And studies of other cultures will of course reveal further disagreement about which kinds of life are best.[109]

What this diversity in outlook reveals is that our current concepts could have been other than they are: although Subjectivism may seem natural within the contemporary social order we now inhabit, beyond this other – and possibly higher – ideals are possible.[110] It will therefore not do to blithely assume that these other perspectives are inferior to the conceptions that we currently possess. To merely recapitulate our contemporary standpoint is thus of limited value: a rational vindication of Subjectivist views must go beyond merely reporting presently-prevailing values or linguistic practices, and a comparative approach is advisable if we are to properly evaluate whether they afford suitable responses to the central problems of action under consideration.[111]

In addition to these observations about cultural variability, later we will also note that our present conceptions of the Human Good are actually themselves at odds even with our own situated practice in important respects.[112] In particular, in the final chapter I shall argue that their underlying individualist assumptions render them too extreme to be adopted consistently, and thus they are inevitably incompatible with some of the empirics of actual human action. As well as drawing on our own inner resources, we also have a central need to draw on the knowledge and wisdom of those from whom we need to learn as we develop as moral agents: to a great extent, we can become who we are only through the

---

honour; on Spartan values, see also Burckhardt, 2002, pp. 27-28.

[108]Though ordinary Greeks may have had more hedonistic tendencies than the philosophers whose views now tend to dominate our picture of Greek culture; see again Burckhardt, 2002 for a more balanced view, and Siedentop, 2014, p. 38 on ancient Athens, Corinth and Imperial Rome as more pleasure-seeking cultures.

[109]For a further example of cultural variation today, see Sawatzky *et al.,* 2011 on the Inuit view that the land they inhabit is a 'direct determinant' of well-being; see also MacIntyre, 2011 [2007], p. 255 for discussion of the issue.

[110]Paraphrasing Nietzsche, 2002 [1986], p 90; see also Wittgenstein, 2003 [1953], p. 195^{e} on overcoming the belief that 'certain concepts are absolutely the correct ones'; for the view that well-being is not straightforwardly a 'natural kind', see Prinzing, 2021, p. 155-156.

[111]Cf. Prinzing, 2021, p. 154 for a different line of objection to the standard descriptive project: 'if there is such a thing as *the ordinary concept* of well-being, it is a reflection of the ordinary (i.e., pre-scientific, often parochial) understanding of human beings'.

[112]Though of course, here we are not taking *de facto* contemporary practice as having ultimate authority either.

influence of supportive others. Hence, although our cultural assumptions about radical self-determination and our received system of values may explain the intuitive appeal of our Subjectivist views, no real support is to be found for them here.

Although gaining clarity on how we already think and talk may at times be useful, then, this cannot constitute the whole of our enquiry if we are to achieve our goals.[113] Our first aim is responding to an ongoing practical problem; as such, our having 'intuitions' that the concept should be rendered in one way rather than another cannot be considered as conclusive evidence here, and there is no binding requirement to defer to 'our' received ways of thinking about the Human Good.[114] Though they bear the weight of inertia, our existing conceptions cannot be an 'ultimate authority': on the contrary, we may even stand in dire need of rescuing from them.[115] Our enquiries into the good may therefore result in us finding that we need to change the word, whereas contemporary philosophy has merely described it.[116]

At this stage, a *caveat* must be introduced to what has been said so far in this section. Although we will not take society's existing *opinions* about the Human Good as binding, this does not preclude us from taking into account the existing social *conditions* in which we must now live and act. Hence, a number of features of the present social order will be tacitly assumed as partly constitutive of the problem we are addressing here – for instance, our individual freedom of expression and of lifestyle choice, without which a lot of what follows would be unintelligible. Such limitations are somewhat inevitable, however, since aspiring to an entirely trans-cultural standard in an enquiry of this nature is futile.[117] Imagine, for instance, a modern philosopher, attempting to address human beings as such *qua* neutral individuals under the rigid hierarchical structure of feudalism, or plying modern ideals of self-expression within traditional

[113]Cf. Raz, 1986, p. 15 in the context of political freedom: 'What we need is not a definition nor mere conceptual clarity. Useful as these always are they will not solve our problems. What we require are moral principles and arguments to support them.'

[114]Cf. Griffin, 1986, p. 2; and MacIntyre, 2011 [2007], p. 83: 'The introduction of the word 'intuition' by a moral philosopher is always a signal that something has gone badly wrong with an argument'.

[115]For the view that this is the attitude with which Socrates initiated moral philosophy itself, see Dewey, 1998, Vol. 2, p. 45.

[116]Paraphrasing Marx. Wittgenstein is the most illustrious defender of the contrary quietist view: 'Philosophy may in no way interfere with the actual use of language; it can in the end only describe it.' – Wittgenstein, 2003 [1953], p. 43^{e}.

[117]See MacIntyre, 1987, p. 3 on the cultural particularity of the present philosophical subject matter; see also see Dewey, 1998, Vol. 1, p. 5; p. 34 for further discussion.

nomadic, agrarian, or pastoral societies wherein it is assumed that 'The only ambition of the son is to be like the father'.[118] In particular, although I will be critical of individualism as a moral outlook, and although parts of the argument may generalise to human beings as such, it is ultimately still the perspective of an individual moral agent living under present conditions that I will take up here.[119]

None of this should be taken to imply that it is invalid to adopt a broader socio-political perspective wherein large-scale features of our society are not considered immovable – or indeed to deny that in fact 'institutions are the real stuff of ethics'.[120] But although our choice of focus here should not be taken to denigrate the importance of these more radically revisionary approaches, these are not my present concern within this particular enquiry: 'I personally hold society to be first in every way and individual's current involvements to be second; this report deals only with matters that are second.'[121]

The above remarks also make plain another aspect of the paradigm adopted here: that the locus of decision-making is an individual human being that can independently deliberate and make choices without requiring direct support from others, in keeping with the modern, atomistic view of the self that I have claimed lends plausibility to the Subjectivist view. Yet this conventional notion of agency has itself been criticised in the recent psychology literature, and for rather longer within academic sociology and continental philosophy.[122] Moreover, these criticisms have often stressed precisely the same social considerations that I will appeal to in carrying out the present critique of Subjectivism about *how* such individual deliberation should proceed and the normative standards it should employ.[123]

[118]Bronowski, 1974, Chapter 2 (discussing the nomadic Bakhtiari).

[119]See Alexandrova, 2017, xxxix-xvi for defence of 'mid-level' theorising about well-being that applies to particular groups in particular circumstances; and Collingwood, 2013, pp. 59-65 for a sustained denial of the view that philosophical problems are 'even in the loosest sense of the word, eternal.'

[120]Knight, 1996, p. 8.

[121]Goffman, 1974, pp. 13-14; see also Bowles, 2004, pp. 47-49, on institutional arrangements in game theory (such as allocations of property rights) being represented as either fixed rules of games (exogenous) or equilibriums of larger games (endogenous).

[122]E.g., in the literature on 'distributed agency'; see Enfield and Kockelman, 2017 for a presentation of recent work from a variety of disciplines; see also MacIntyre, 2016, p. 72 for Aristotelian criticism of an exclusive focus on the individual first-person perspective, as opposed to a collective view of action aimed at common goods.

[123]See Taylor, 1989, p. 40 for the view that people from India feel incapable of making important choices in the absence of key family members.

These more radical attacks on individualism thus give the notion of agency itself an essentially social dimension, and this threatens to cast doubt on our underlying framework of individual choice as a whole. Yet regardless of the validity of this challenge in general, we will nevertheless be entitled to retain the conventional view of agency within the present, critical work. For these background assumptions are also adopted by the proponents of the Subjectivist views that it is the sole aim of the thesis to criticise.[124] Indeed, once we have explored Subjectivism as an Action-Guiding view in the next chapter, in the final one we shall go even further and focus specifically on a particular conception of deliberation that is tailored to the precise requirements for deploying the view in practice; that is, practical reason as the Subjectivist views it. Finally, an advantage of this *reductio ad absurdum* approach is that we are thereby taking on Subjectivist views on their own terms, whereas these rival approaches may make assumptions that are so different that their criticisms do not make contact, as there is too much disagreement on fundamental matters for productive engagement to occur.

To summarise the key points of this section, then: the critical argument offered by this thesis will not proceed by maintaining that our two Subjectivist theories do not reflect our current thoughts or values – nor indeed that they harbour any intrinsic logical inconsistency, since both conceptions have been worked out in painstaking detail under the contemporary Appraisive approach. Rather than issues of conceptual analysis, I am instead concerned with seeking a theory that will find application in practice in a psychologically realistic setting, as an ideal for individual human beings to aim at in their practical deliberations. It is in *this* arena that an Action-Guiding account of the Human Good must make its claims.[125]

1.4 Timescales

In our A.I. book *Designing Intelligence*,[126] we argue that a complete explanation of the behaviour of any organism or artificial system should

[124]Though see Sections 3.5-3.6 for defence of the view that rational agency must be in a sense *sustained* by others.

[125]Cf. MacIntyre, 2013, pp. 27-30; Collingwood, 1992 [1940], p. 402. See, conversely, Hills, 2007 on evaluating conceptions of practical reason by allying them with a theory of value.

[126]Pfeifer, Bongard and Berry, 2011, pp. 22-23.

incorporate three different 'timescales': nested temporal perspectives which have their origins in biology. The three timescales are: the here-and-now, the developmental, and the evolutionary.[127]

To see the timescales in action, imagine a lion cub playfully pouncing on his sibling, who intuitively responds by moving his paws to block the faux attack; he then continues to wrestle with his brother and tries to pin him down. In the here-and-now, the proximal cause of our lions' play may be the pleasure it induces and the instinctual desire they feel to engage in it.[128] And if a lion could talk – and if, *contra* Wittgenstein, we could understand him – then this is what they might say if we were to ask about their motivations: 'I just *like* pouncing on things.' Yet their activities can also be understood as an expression of a pre-existing relationship through which such pleasure responses and desires are themselves cultivated. Indeed, from a developmental perspective, it is in part through engaging in play that our young lions build up a repertoire of social and physical skills that enables them to flourish later in life: through learning what they need to learn they become able to hunt successfully and to defend the pride from outside threats, or perhaps to ward off challengers for a dominant role within it. Finally, from a much broader temporal perspective, evolutionary processes are ultimately responsible for equipping the cubs with the capacity for these instincts, desires, and affective reactions – largely due to these social and developmental benefits of play.[129]

As this example makes plain, one reason why it is necessary to consider different timescales is that the organising principles of behaviour are themselves different across each level: ontogenetic mechanisms of growth and development are quite distinct from how the nervous system operates in the short term, and both are again dissimilar to the replication, mutation, and selection processes of natural selection.[130] At the same time, the different timescales are linked and often not cleanly separable: for example, if you hit your thumb with a hammer, causing pain and hence a

[127]The same basic scheme was proposed by Waddington, 1957, pp. 5-7. See Sapolsky, 2017, for a more expansive discussion of timescales and corresponding levels of explanation of behaviour; running via, respectively, the nervous system, sensory systems, endocrine system, learning and development, acculturation, the ecological shaping of culture itself, and ultimately evolution; a summary is presented on pp. 6-7.

[128]Sapolsky, 2017, p. 205

[129]See Sapolsky, 2017, p. 204; for more on the developmental implications of play, see Berghänel *et al.*, 2015; see also MacIntyre, 2009b [1999], p. 85 for philosophical discussion.

[130]Pfeifer, Bongard and Berry, 2012, p. 23.

defensive reaction in the here-and-now, this also teaches you to be more careful with the hammer in future, and if you continue working with the hammer and other tools then over time your prowess as a handyman will develop. Obviously, this development also affects the 'here and now' of future situations. And evolution affects the process at all levels, not only enabling our capacity for learning, but also the opposable thumbs, hand morphology, and hand-eye coordination that enables our species to use tools such as hammers at all.

In addition to explaining behaviour, the idea of timescales can also be applied to our core focus here: the search for a tractable model of practical reason centred on an Action-Guiding theory of the Human Good. I have indeed already suggested that rational deliberation is again heterogeneous in nature: that we can reason both about what to do in the here-and-now, and about how best to make sense of our lives as a whole.[131] This key distinction between the kinds of normative guidance a human rational agent needs corresponds to the following two temporal perspectives:

1. *Action Selection*
   This perspective concerns human beings' use of practical reason to guide their choices and decide what to do in the here-and-now, in response to the specific circumstances at hand.

2. *Life Planning*
   This perspective concerns the agent's life as a whole: their answer to the question, 'What should my life be like?', and their implementation of this ideal in practice.[132]

As an example of these two timescales in action, think first of carefully choosing a main course from a menu of options whilst dining at a restaurant, based on our current conception of what kinds of food are best for us to eat here-and-now, given our existing preferences and constitution. This is the perspective of Action Selection. From the perspective of Life Planning, consider instead attending to our evolving relationship to food

[131] Cf. Scanlon, 1998, p. 126, who draws much the same distinction that I do here. Scanlon is skeptical about the role of the concept of well-being for rational deliberation, but only within certain constraints that I shall later set aside; see Section 1.6 below. Again, in this project I also largely bracket off empirical questions about the importance of cognitive-guided behaviour in comparison to other species of behaviour, focusing solely on what happens when we do engage in rational deliberation.

[132] On similarities between the modern idea of a Life Plan and the ancient notion of *telos* or final end, see Annas, 1993, p. 38.

in the long term, such as if we set about educating our skills and palates as a means to pursuing a career as a professional chef; alternatively, we may aim to undergo a quite different education in culinary taste by coming to appreciate and enjoy – and thus acquire the implicit tendency to select – food that is good for our health.[133]

Life Planning, the province of reflection on the good *life* proper, includes deliberation about its external features: the professional or social circumstances we strive to position ourselves in, what material resources we control, or our relationships with others. But I will also take it to include internal aspects of our lives, such as our character: indeed, as we have already seen, with the classical Stoics it is only this latter type of feature of how lives go that should be considered truly valuable, and for them externals are at best mere 'preferred indifferents'.[134] The Life Planning perspective can also only arise for certain kinds of agents: in order to consciously act with a clear view to our long-term interest, we must be capable of planning out large stretches of our future behaviour and action. It therefore assumes a special relationship to time often thought to be unique to human beings.[135]

Paying attention to these two timescales will be especially important for the present enquiry because from different temporal perspectives we may have a quite different relation to our Subjective Values, as we saw with the culinary example given above. I cannot, here-and-now, simply *decide* which desires to have or what I will find pleasurable on the basis of thought alone, so from the perspective of Action Selection these items appear relatively fixed.[136] Yet in the context of Life Planning the impotence of reason is alleviated here: over the course of a lifetime, we *do* have the ability to train, educate and order our Subjective Values – including our preferences – according to our cognitive judgements of what is best for us, in the course of moving towards the best possible life we can envision.[137] This idea will be developed further as we apply

[133]For a study of the varying importance of immediate affect-driven *versus* long-term cognitively-driven processes for food choice, see Shiv and Fedorikhin, 2002. As another example, consider a scientist deliberating about how to interpret the results of a particular experiment, or thinking about the direction of their career as a whole.

[134]Long and Sedley, 1987, pp. 354-359; see also Rosati, 1995 on how our direction in life impacts our Subjective Values.

[135]See MacIntyre, 2016, p. 3 on animals' (and juvenile humans') lack of capacity to distance themselves from their immediate felt needs and desires; see also MacIntyre, 2009b [1999], p. 60 and Hearne, 1987, pp. 163-165.

[136]MacIntyre, 2009b [1999], p. 115.

[137]MacIntyre, 2009b [1999], p. 116; p. 122.

our timescales to our two core Action-Guiding approaches: contemporary economics, and ancient ethics.

We begin with EUT and related approaches. As suggested earlier, these accounts not only make Action Selection central, but in fact omit the perspective of Life Planning entirely.[138] Within this tradition, rationality is understood solely in terms of making discrete choices that maximise our utility – or expected utility, where uncertainty is involved – where 'utility' means the satisfaction of whatever *de facto* preferences we have at the moment of choice. Since these preferences are taken as fixed prior to and independently of the models of choice the theory provides, our Subjective Values are not available for deliberate modification within mainstream economics.[139] In particular, attention to only the first of our two timescales here precludes any notion of rationally-guided moral development through the education of the passions – even though in ancient ethical thought this process is central.[140] As we shall see in detail later on, this neglect thus severely limits human possibilities; moreover, for the present enquiry, it also entails that the attraction of certain rival accounts of the Human Good – that is, those which give a central role to the cultivation of character and development of the virtues over time – becomes invisible from the outset.

As we have seen already, Ancient ethical theory tends to instead begin with Life Planning: consideration of how an agent can make sense of their lives as a whole, which may include attending to what kind of person they should aim to be.[141] Because the acting moral agent thus takes their own self-development as the fundamental concern, some contemporary critics have found the approach to be narcissistic at best, or even egoistic and self-centred.[142] However, the *content* of the theory typically involves

---

[138]Though some normative economic work takes a longer-term perspective that includes the individual choices made by many successive generations; for a classic contribution to intergenerational welfare economics, see Ramsey, 1928.

[139]In economic parlance, preferences are typically 'exogenous'. There are some recent exceptions to this: see e.g. Boissonnet, Ghersengorin, and Gleyze, 2022, and see Bowles, 2010 for a discussion and review of the literature.

[140]See Annas, 1993, p. 38; p. 43 for the argument that ethical enquiry must go beyond the here-and-now perspective: 'particular choices take place within, and have to be understood in a context of, larger patterns in one's life.' Another area of economics does make room for a distinction of this type: in producer theory we distinguish between short-run and long-run cost functions, where in the latter case more aspects of a firm's structure may be considered 'choice variables' that are available for optimisation.

[141]Annas, 1993, pp. 27-29; p. 440; p. 443.

[142]See Swanton, 2014.

the development of other-regarding virtues as a central aspect[143] – for instance, the Stoic school's core imperatives concern the internal aspects of Life Planning and urge us to recognise virtue and good character as the sole Human Good: to tread the path of the *prokopton*.[144] Although it is our lives as a whole that are the initial focus, consideration of these virtues then requires us to choose in other-regarding ways at the level of concrete actions, and Action Selection is thus closely to Life Planning via this strict adherence to virtuous conduct.[145] Aristotle therefore sees 'right action' as an essential component of *eudaimonia*, and consequently devotes many pages to discussing concrete actions; Epictetus, meanwhile, implores us to 'Firstly, tell yourself what you want to be, then act your part accordingly.'[146]

Although the way in which we choose to act in the here-and-now thus forms no small part of our character, we should also take care that the ancient prioritisation of the Life Planning perspective not be overstated. In particular, we should not assume that every choice we make, however casual, must derive directly from a broader conception of how our lives as a whole are unfolding:

> One need not always be thinking of the last end, whenever one desires or does something ... while walking along the road one needs not to be thinking of the end at every step.[147]

This question of the priority of the two timescales will be important to the enquiry – yet since not all schools of thought agree here, we will be wise to remain neutral for the time being, developing the framework of enquiry without prejudicing the matter at this stage and thus distorting our later accounts of these rival views.[148] Rather than attempting to settle the matter *a priori*, then, in what follows I shall therefore treat the relation between our dual contexts of Action Selection and Life Planning

[143]See Annas, 2011, pp. 154-156; Annas, 2007, and Annas, 1993, p. 44 for similar responses to the narcissism objection.

[144]For a summary, see my Berry, 2019a.

[145]See Hursthouse, 1999, pp. 35-37 for a discussion of 'v-rules': restrictions at the level of Action Selection that consideration of the virtues enjoins.

[146]Epictetus 2008, p. 168 (III.23).

[147]Aquinas, *Summa Theologiae*, IaIIae, q. 1, a. 6; however, for Aquinas such choices are in a more fundamental sense *grounded* in such a conception, in 'virtue of the first intention' – *ibid.* Compare with Angier, 2015, p. 16: 'As Aristotle would be the first to say, even though, at a formal level, *eudaimonia* always and necessarily constitutes my ultimate reason for action, what figures in my actual practical deliberations are reasons grounded in concrete, particular goods-to-be-achieved.'

[148]Cf. the attitude to similar partisan issues advocated in Sections 1.5 and 1.6 below.

as an internal component of particular Action-Guiding theories, such as those we will examine in depth in the next chapter. Nevertheless, at this juncture it will be worth raising some preliminary considerations of a general nature that are pertinent here by supplementing the above remarks about normative theory with some observations about contemporary practice.

Some particularly assiduous agents may strive to make all of their here-and-now choices with a view to their impact on their lives as a whole. Yet not everyone approaches their deliberations like this; think again of the different attitudes diners take when ordering in a restaurant. Moreover, in everyday life, our behaviour may often be guided by some conception of a normal day, week, month, or even year, and this preexisting structure means that if we so desire we can for long periods of time avoid explicitly thinking about the general direction of our lives at all; we may be carried along by habit and the social and institutional scaffolding that surrounds us until we are confronted with quite special circumstances.[149] Yet sometimes we may naturally feel invited to go beyond this here-and-now reasoning and reflect on our lives as a whole, bringing into play our vision of how we intend them to move forward from here.[150] Consider three types of instance.

Firstly, most agents will at some point encounter urgent, formative choices which are not settled by precedent but which will inevitably have a large impact on what follows: consider an inherently long-term decision, such as buying a house or immigrating to a new country, where the mode of deliberation is to some extent imposed by the external features of the choice context. Such 'hinge' choices may take many other forms: whether to continue or abandon a long-term project such as writing a novel, what career options to pursue, which side to choose in an argument where friendships are at stake, whether to adopt a child, or whether to accept a proposal of marriage.

In a second type of instance, longer-term oversight may become needed due to our lives or character having gradually veered off in an unwanted direction, such as a descent into licentiousness: here we must now proactively select a time to intervene in our own habitual behaviour. Such

[149] Cf. MacIntyre, 2016, pp. 6-7.

[150] Cf. Annas, 1993, p. 28 for more on this type of reflective thought and its genesis; cf. MacIntyre, 2011 [2007], pp. 239-241 for the claim that to fully understand and explain a segment of human behaviour we must consider how that individual episode fits into the agent's life as a whole, and Vallacher and Wegner, 1987 for an empirical perspective on these levels of identification.

cases also reveal that the purview of practical reason comprises not only specific contexts of choice between fixed options, but also actively taking the initiative and embarking on spontaneous self-interventions, as we strive to do 'the right action at the right time'.[151]

Thirdly, on still other types of occasions, in contrast, specific episodes are the driving factor for broader reflection: for instance, many addicts have reported on the awakening experience of 'hitting rock bottom'.[152]

We now consider three reasons that urge us to follow the ancients in emphasising the longer-term Life Planning perspective more generally. Firstly, the overall direction of our lives largely determines the here-and-now contexts of choice that we will come to face. It may therefore be advisable to start here, like an artist who first composes and draws in the large-scale features of a landscape before working out the local details. Secondly, we have seen that the range of responses we will consider in a given situation and our characteristic modes of Action Selection in general will both depend on our character: the question of *what* we are to do is thus conditioned by our answer to the question of *who* we should become.[153] This in turn is itself a central part of what our lives are like, and so comes under the Life Planning paradigm, which as we have noted can include internal as well as external considerations. Thirdly, as suggested earlier, certain modes of deliberation can only be effective over the long term, when we can successfully alter more of the relevant variables – such as our projects and relationships, and especially our Subjective Values, which have no small effects on our local, individual actions in the here-and-now, but are relatively rigid from this narrower perspective.

There are, however, also competing reasons that pull in the direction of prioritising Action Selection: for instance, the compounded uncertainty of distant future outcomes over a lifetime *versus* the comparative certainty of the here-and-now.[154] Moreover, to be effective the results of deliberation within Life Planning must be implemented in concrete

[151]Seth, 2012, p. 37; this is a point sometimes neglected by economists, who focus on static, imposed choices between fixed options.

[152]Cunningham *et al.*, 1994.

[153]See also Williams, 2006 [1985], p. 185, on the kind of person we are affecting the nature and frequency of our deliberations.

[154]For an empirical perspective, see Schunk and Usher, 2012, p. 18; this consideration was also urged by the Cyrenaic philosopher Aristippus the Elder of Cyrene, whom we shall meet in Chapter 2. Another consideration is that developmentally speaking, the capacity for Action Selection appears before that of Life Planning; see Chapter 3.

actions; otherwise we may end up spending a great deal of time planning and imagining our own future without this making contact with what we actually do, and thus find that 'Life is what happens to you while you're busy making other plans.'[155] There is therefore a trade-off here which may require delicate judgement to resolve.[156]

These preliminary observations suggest that to be fully adequate to its role, an Action-Guiding theory must offer guidance for both kinds of rational deliberation – a thesis I will defend later on in Chapter 2. We should be cognisant of this when considering the role a successful Action-Guiding theory of the Human Good should play and the structural form it should take; topics which we address in the next two sections.

1.5 Role of Theory: Action Selection

In this section, we reflect on the role that a conception of the Human Good might take within a broader account of Action Selection.

When we rationally deliberate about what to do in the here-and-now, we are comparing the *reasons* we have for pursuing each possible course of action open to us. This leads us immediately to the question: 'What is it to have a good reason?' MacIntyre offers a plausible answer here:

> To act for a good reason is to act for the sake of achieving some good or preventing or avoiding some evil.[157]

This broadly Aristotelian, goods-based perspective on reasons fits well with the empirical literature on human cognitively-guided action, and is, I think, quite generally applicable, doing justice to the nature of practical reasoning across different cultures and historical time.[158] And regardless of whether this is the only coherent sense in which one could have a reason for action, given the goal of the present project it is clearly the

[155]John Lennon, "Beautiful Boy"; see also Dewey, 2002 [1992], p. 197 on the 'vices of reflection'.

[156]See here the large literature on narrow *versus* broad bracketing in economics; broader bracketing may be helpful (see e.g. Camerer *et al.*, 1997 on cab drivers' working hours) or unhelpful (see e.g. Simonson, 1990 on 'diversification bias', discussed in Section 2.4 below).

[157]MacIntyre, 2016, p. 8; cf. MacIntyre, 2009b [1999], p. 24 and Quinn, 1993, p. 234.

[158]See Anscombe, 2000 [1957], p. 22, and Annas, 1993, p. 29: '... it is hard to understand the performance of a deliberated action unless the agent saw something good in the outcome he was trying to bring about'. As we shall see below, a 'good' here need not be something external; for instance, in the Aristotelian tradition it is often the performance of the action itself.

one we should focus on.[159] However, this of course takes us only so far as the further question: 'What is a good?'

Here we find a prime place for a theory of the Human Good to make a contribution, in providing us with an answer to this question. Yet we should bear in mind that the concept of the good employed in practical reason may extend beyond what is good for us in particular, as in cases of other-regarding action, and perhaps even beyond what is good *for* anyone.[160] Unsurprisingly, then, the normative theories we will consider embody different and incompatible conceptions of how they are to find application here, with one important difference between these rival accounts being the relative importance of our own good *versus* that of others in guiding our actions. Although the accounts of practical reason we will examine all allocate *some* guiding role to our notions of what is good for us, then, the details are quite different from one theory to the next. We now briefly consider some of these possibilities.

The Subjectivist modes of practical deliberation we will be primarily concerned with have in common the individualist view that each of us must only look within ourselves, to our own desires and affective responses, in order to determine how best to act, so that reasons are always 'internal' reasons that are reflected in our '*subjective motivational set*'.[161] Moreover, since such views do not attach intrinsic value to securing pleasure for or satisfying the desires of other people, they have an inherent tendency to privilege the good of the acting agent.

Mainstream economics in particular takes an extreme position here: in EUT, practical reason is *entirely* constituted by maximising our own personal utility. Utility in turn has been construed quite narrowly, and in many cases the agent's wealth is seen as an adequate substitute.[162] In neoclassical economics, then, agents only make choices based on con-

[159] See the recent literature for alternative 'process-based' views wherein reasons are understood as premisses in sound reasoning; e.g. Way, 2017. There are various other important distinctions here that I cannot enter into here; though on normative *versus* motivating reasons, see Cunningham, 2021.

[160] Cf. Sumner, 1996, p. 13: 'my own well-being is just one possible end of action among others for me'; for contrasting views on the latter issue, see Kraut, 2011, who argues that all goods must be good *for* particular agents; and Moore, 1993 [1903], pp 148-153, who argues that the concept 'good for' is incoherent.

[161] Williams, 1979, p. 18 [italics in the original]; Williams' view is also partly psychological as well as normative, since for him human reason is only able to select an action for which the requisite motivational power already exists.

[162] See Buchanan, 1989, p. 20: 'Individuals must be modelled as seeking to further their own narrow self-interest, narrowly defined in terms of measured net wealth position'.

sideration of their own personal good in a narrow sense;[163] and modern agency theory has extended this paradigm to situations where the agent in question may have otherwise been thought to be deliberating in a quite impersonal way – such as in running a large company, or holding public office.[164]

Generally speaking, for a viable theory of action, such broadly egoistic accounts tend to require support from the complementary framework of 'morality' in the modern sense: that is, an external system of rules that provide socially necessary boundaries to an agent's unrestricted pursuit of their desires.[165] However, since at least the *Theory of Moral Sentiments* economists have argued that such purely selfish actions on the part of individual agents can lead to socially desirable outcomes. This surprising notion reaches its height of development in the celebrated First Welfare Theorem, which arises within the modern theory of general equilibrium, as inspired by Walras and Marshall, and set out axiomatically by Arrow, Debreu, and McKenzie.[166] Here outcomes secured by trade under self-interested motivations are seen to be efficient or optimal, in a certain technical sense;[167] thus Hume's knave is guided aright by Smith's invisible hand.[168]

Noting a slew of empirical inadequacies with this traditional egoist approach, recent work in behavioural economics has instead taken a somewhat more enlightened view, allowing an agent's 'preferences' to depend on the material payoff of others. However, it still retains the same simple maximising conception of the relation between what is good for us – now in this more inclusive sense – and our practical reasoning.[169]

[163]See also Mill, 2000 [1984], pp. 97-98 for a defence of this selfishness assumption, though only in the narrow context of abstract economic theorising as understood at the time; see also Sen, 1977, pp. 317-318 for discussion of a similar move by Edgeworth.

[164]On the expansion of economics into non-market domains, see again Backhouse and Medema, 2009, pp. 229-230.

[165]As appears in, for instance, Hobbesian contractarianism; for criticism of 'Morality' in this sense, see MacIntyre, 2016, pp. 64-69.

[166]Ingrao and Israel, 1990; Debreu, 1959.

[167]Specifically, Walrasian Equilibrium Allocations are Pareto Efficient, and belong to the 'core' of the model economy, meaning that they cannot be disrupted by a coalition with a mutually profitable deviation.

[168]Cf. Hume, 1994, p. 113. These ideas have been repeatedly used to justify minimal government intervention, due to the apparent desirability of market outcomes; however, the power of the second welfare theorem – which states that *any* interior Pareto efficient outcome is a competitive equilibrium for *some* set of endowments and prices – becomes apparent only when we consider redistribution.

[169]See Section 2.1 for a discussion of key examples, with references to the literature; see also Sen, 1993, pp. 36-37, and see Sen, 1997, pp. 326-327 for a tripartite

Acting in consideration of others is thereby condoned, but only insofar as doing so has a positive impact upon our utility: that is, insofar as we have 'social preferences' that are themselves inherently other-regarding, so that 'individual preference is and remains sovereign in judgment of what is good and bad.'[170]

We remarked earlier that Ancient ethics focuses on reflection on our lives as a whole, and thus contrasts with EUT by making Action Selection largely subordinate to Life Planning. In addition, another important difference is that although the focus is on *our* lives, and accusations of narcissism have thus been raised, 'best' here is in fact now understood in a more open and all-things-considered sense, rather than as material or personal benefit for us in particular.[171] Similarly, when the ancients do attend to rational deliberation about concrete actions in the here-and-now, we typically find that our personal good is much less important in deciding what to do: here the key concept is instead that of what 'right action' requires.[172] Yet rather than being based on abstract moral rules, this latter notion is itself generally understood in terms of the good of all those whom our actions affect, and the shared goods inherent to the *polis*, thus still relating to the Human Good *per se* – so, from a broader perspective, it remains true that 'every action and choice, is thought to aim at some good'.[173]

In contrasting neoclassical economics with behavioural economics and ancient ethics, then, we see that the specific way the problem of Action Selection is conceptualised will itself depend partly upon the nature of the particular solution under consideration; moreover, rival understandings of the problem will of course also heavily condition the acceptability of corresponding solutions too.[174] Thus amongst these rival traditions of practical rationality there is disagreement not only in *what* is to

distinction between classical narrow self-interest, defined in terms of personal material gains; broad self-interest, which also includes feelings of sympathy for affected others; and our commitments, which may extend beyond our own welfare however broadly understood.

[170]MacIntyre, 1993, p. 10; cf. again MacIntyre, 2009b [1999], p. 71 for a denial that individuals are necessarily authoritative about their own good, and see Sections 1.7 and 2.1 below for further discussion.

[171]See the next section for further discussion.

[172]See e.g. Aristotle on acting according to 'right reason' – *NE* VI.2, 1139a22-25.

[173]Aristotle, *NE* I,1, 1094a1-2; cf. Annas, 1993, p. 29 for discussion.

[174]Our is thus a 'wicked' problem, which is another reason to avoid precise technical formulations. For a classic article on wicked problems in general, see Rittel and Weber, 1973; the notion of a 'problem' is of course itself a normative one.

be chosen, but also in *the manner in which* one is to choose.[175] This given, then, we must for now remain neutral between these rival options; otherwise we will not be able to make full use of the available resources, nor critically engage with these accounts on their own terms. Hence, although we have seen in a rough way what the central problem of Action Selection is and what work a theory of the Human Good would have to do in this context, if we are to avoid a limiting bias at this stage then any more detailed answer to the proper role of such an account must therefore be postponed. And we must also take note of any further such partisan issues as we continue expounding our basic framework of enquiry.

1.6 Form of Theory: Life Planning

In this section, I raise the issue of what form a theory of the Human Good should take in the context of Life Planning. As with Action Selection, we shall soon see that a number of rival views are possible here; given our commitment to making use of all the resources at our disposal within our three core approaches, then, for the time being we must again take care not to prejudice the enquiry at this early stage through an *a priori* restriction to theories with one particular type of structure or scope whilst neglecting other possibilities. However, as we shall also see shortly, such potentially-limiting assumptions are in fact adopted in much of the contemporary Appraisive philosophical literature on well-being – and often deliberately and explicitly.

Since our investigation will need to make heavy use of these conceptual enquiries into the good life, in this section I will proceed by highlighting and commenting upon two such collateral assumptions, whose uncritical acceptance would otherwise place unwanted constraints on our Action-guiding enquiry – in particular, making invisible just those shared goods that will be the focus of the critical argument of Chapter 3.

In perhaps the most influential contemporary defence of hedonism in the Appraisive sense, Fred Feldman opens his argument by expressing his intention to offer a 'substantive theory of the Good Life'; an account of when a human life is good in itself. But it is notable that he immediately and without comment passes to another formulation: his aim is also

[175] See again MacIntyre, 1987, p. 3 for the view that rival theories of practical reason do not characterise 'one and the same timeless and a-historical subject-matter', but rather 'rational action is structured very differently in different times and places'.

to 'give an account of the *amount* of welfare, or well-being, that an individual enjoys'.[176] This move to what is considered to be an equivalent statement indicates that Feldman is conceptualising well-being as a kind of 'thing', which our lives can thus contain a determinate 'amount' of.

Appealing to a plethora of empirical evidence, the cognitive linguist George Lakoff and his collaborators have argued that it is typical of human beings to understand complex ideas through 'cognitive metaphors': coherent and structured ways of understanding ideas in one domain (typically a more abstract one) in terms of ideas from another domain (typically more concrete and grounded in our embodied sensory-motor existence). They explain: 'Primarily on the basis of linguistic evidence, we have found that most of our ordinary conceptual system is metaphorical in nature', where 'The essence of metaphor is understanding and experiencing one kind of thing in terms of another.'[177] The suggestion here, then, is that Feldman is thinking of and reasoning about well-being in a substantive, essentialist way, as though it is a kind of 'commodity'; that is, a valuable 'thing' which we can have more or less of, and to which our lives stand as a kind of receptacle to be 'filled up' with to a greater or lesser extent.[178] This is an assertion that Feldman's more technical proposals later in the book seem to confirm;[179] and though the metaphysical basis of the assumption may differ, it is also one shared by many other contemporary analytic philosophers working on well-being today.[180]

If well-being is thus thought of as a valuable 'thing' of a determinate and uniform kind, it is natural for a theory of the good life to then take the form of an account of its nature. For instance, Feldman defends the view that it consists of mentalistic 'units' of pleasure; in his terminology, 'hedons'.[181] He further argues that well-being can be chunked into 'units', and each such 'unit' of well-being must be localised to a specific time; it is then alleged that one can in principle 'calculate' the value of a particular

[176] Feldman, 2004, p. 1 [my emphasis].

[177] Lakoff and Johnson, 2003 [1980], pp. 4-5.

[178] For an existing application of Lakoff's ideas to happiness in particular, discussing amongst other suggestions the metaphors of happiness as 'fluid in a container' (p. 33) and a 'valuable commodity' (p. 38), see Kövecses, 1991. The use of the somewhat-vague term 'commodity' is not intended in a strictly economic sense of a raw material that is tradable.

[179] Indeed, a more recent monograph is entitled '*What is This Thing Called Happiness?*' – Feldman, 2010.

[180] Cf. Angier, 2015, p. 6 [footnote 5]. For details of the official metaphysical basis of Feldman's view, see Feldman, 2000; Feldman, 1997, pp. 116-118.

[181] Feldman, 2004, p. 25.

life by summing up the 'amount' of this commodity that it 'contains' in each such episode:

> Ideally, I would like the theory to assign specific (perhaps numerical) values to those elements, and then to give a systematic way of aggregating those values so as to yield a value for the whole life.[182]

We now reflect on the import of this assumption when accounts of the good life are put to work in an Action-Guiding context, focusing on Life-Planning.

At the most general level, what a particular human rational agent needs to engage in practical reasoning of this kind is what I shall call a 'Life Plan': a possibly rough but explicit and articulate vision of the best kind of life they can hope to achieve, so that they may rationally evaluate their choices according to this standard.[183] And as above, *if* the good life is simply one containing large amounts of a valuable 'thing' called 'well-being', a natural role for philosophy again becomes that of informing the content of such Life Plans by identifying the nature of this commodity. However, in the broader literature there are a number of other routes by which theory aims to make an Action-Guiding contribution to practice here – so again, we must be cautious about this implicit assumption, since it threatens to obscure the attraction of accounts that approach the issue differently, with the unwelcome consequence that some form of Subjectivism may come to seem unavoidable.

A quite different conception of what philosophy might do here is to render a definitive picture of one specific sort of life; one that is intended to suit all human beings as such – as with, for example, the distinct practical and intellectual visions offered by Aristotle in the *Nicomachean Ethics*, or the ideals offered by some strands of Thomism. Alternatively, an account might conclude that an agent's ideals need to be highly tailored to who that agent is and what the circumstances of their lives are, and thus instead aspire only to an account of *how* agents are to formulate and develop their own individual, personalised Life Plan, and

[182]Feldman, 2004, p. 13; see Section 2.3 for a critical discussion of the alleged measurability of pleasure.

[183]For differences between this notion and the concept of well-being as it usually appears in contemporary analytic philosophy, see Scanlon, 1998, p. 130; see also *ibid.* and Raz, 2004, p. 281 on the limitations of this dominant conception for practical deliberation; Sumner, 1996, p. 1 for an opposing view; and Rawls, 1971, pp. 407-416 on the relation between a person's good and their life plan.

what resources they will need to draw on, rather than giving a worked-out solution that is intended to apply in every case.[184] Subjectivism may take this route by directing agents towards a Life Plan that is maximally filled up with a certain commodity-like ‘thing’; yet the advice on offer might also take other forms too – for instance, giving only a number of general substantive features that any worthwhile life must have;[185] or instead, offer a set of what I refer to elsewhere as ‘design principles’ – broad heuristics that can be applied in a flexible way.[186] It may indeed combine such positive suggestions with negative prohibitions; thus on the Stoic account the constructive advice offered is that our Life Plan should centre around a coherent moral goal that is both consistent with and promotes the development of the virtues. And although the exact content is left unspecified here, some career paths – it is not hard to think of examples in contemporary society – can be ruled out entirely.

In terms of their content, Ancient accounts of the good life more generally are also often based on the virtues, and as a result tend to conceptualise the Human Good quite differently to contemporary philosophers: they often take as their key concern *what we do* and *who we become*, rather than some quantifiable ‘thing’ we might ‘fill up’ our lives with, in the manner envisioned by some recent analytic theorists.[187] For instance, Aristotle sees *eudaimonia* as a kind of *activity*, through which we integrate various individual goods into a cohesive whole:

> Whereas contemporary theorists conceive of happiness as akin to a technical product that can be systematically isolated and measured, on Aristotle’s ethical theory it is nothing over and above virtuous activity.[188]

[184]See Annas, 1993, pp. 6-7 on this being the ambition of ancient ethical thought in general. See also Griffin, 1986, pp. 56-57, for discussion of the assumption that perfectionism in an Appraisive sense must focus on a single concrete ideal, and that ‘to the extent that the ideal is really manifested in them, they will all be, on a higher level of generality, the same sort of life’.

[185]This approach is followed by, for example, some versions of the capability approach: see the list in Nussbaum, 2011, p. 33; see also the list of essential goods suggested in MacIntyre, 2013, pp. 24-25.

[186]Pfeifer, Bongard, and Berry, 2011, p. 19.

[187]Cf. Dunne, 1993, p. 110: ‘the function of *phronesis* is not to maximize a ‘good’ that one already knows and can come to *have*, but rather – a much more difficult task – to discover a good that one must *become*.’ See also Annas, 1993, p. 38: ‘ancient ethics does not aim at the production of good states of affairs, and so is not tempted to think that rationality should take the form of maximising them.’

[188]Angier, 2015, p. 19; cf. *ibid.*, pp. 8-9; and Aristotle, *NE* I.9, 1099b26. See also Angier, 2020, p. 149 for related criticism of Subjectivist views, arguing that ‘happiness itself is both logically and ontologically conditioned by the practice of the

This then of course bears on the form of the theory of the good life he offers: Aristotle 'takes the *telos* of human life to be a *certain kind of life*.'[189]

MacIntyre's development of this Aristotelian position introduces another key claim that further impacts on the structure that any adequate theory of Life Planning can take: that is, that a human life is best conceived of as an enacted dramatic narrative. Such a life thereby falls into some particular genre – tragedy, comedy, adventure, romance; and although for MacIntyre the virtues remain of central importance, the nature of the story so enacted is also among the most important questions to ask when deciding how well a human life has gone; indeed, for him it is ultimately what furnishes a life with the possibility of being a success or failure at all.[190] And from an Action-Guiding perspective, it follows that 'I can only answer the question 'What am I to do?' if I can answer the prior question 'Of what story or stories do I find myself a part?'[191] Yet such narrative-based questions are not settled by ascertaining the presence or absence of a certain substance-like 'thing' – any more than one judges a work of art simply by checking the degree to which it possesses a determinate, homogenous, essence-like quality called 'aesthetic goodness'.[192]

We will not pursue issues of narrativity further in this project, which offers a quite different argument against two specific Subjectivist views that have become dominant in contemporary philosophy and the social sciences, and which are clearly incompatible with any conception of the Human Good that has a narrative element.[193] However, as with virtue-based ancient accounts, since a view like MacIntyre's constitutes an as-of-yet viable position, a framework of enquiry that excludes it from the outset will be insufficiently neutral for our purposes. Given these rival

---

virtues'. See also Louden, 1992, p. 106 for further discussion of Aristotle's view.

[189]MacIntyre, 2011 [2007], p. 204; cf. Aristotle, *NE* I.7, 1098a18-19. See also Annas, 1993, p. 37, and p. 38: 'Rather, what I aim at is my living in a certain way, my making the best use of goods, and acting in some ways rather than others.'

[190]MacIntyre, 2011 [2007], p. 253: 'the only criteria for success or failure in a human life as a whole are the criteria of success or failure in a narrated or to-be-narrated quest'; cf. *ibid.*, p. 168; p. 247; and MacIntyre, 1992.

[191]MacIntyre, 2011 [2007], p. 250.

[192]See also Moore on his 'principle of organic unity', which states that judgements of the good are not in general aggregative in the way Feldman and other 'additivists' postulate: '*The value of a whole must not be assumed to be the same as the sum of the values of its parts.*' – Moore, 1993 [1903], p. 79.

[193]Though as it happens I am quite receptive to the narrative view, and believe that the critical argument that follows is broadly consistent with it.

possibilities, at this preliminary stage of the present enquiry we should thus avoid assuming that the goal of Life Planning is to 'fill up' our lives with a determinate commodity called 'well-being' whose nature it is the job of a theory of the Human Good to specify, with Subjectivists identifying this 'thing' with the satisfaction of their Subjective Values. Indeed, as noted above, our characterisation of Subjectivist views will instead be epistemic in character: we will insist only that agents are advised to identify what their good is through reflecting on these Subjective Values.[194] We now consider a second core assumption found in the contemporary philosophical literature.

Another central aspect of Feldman's overall framework of enquiry is the careful insistence that assessments of well-being must encompass only the question of why some lives are good 'for those who live them'. On this view, well-being must always be conceived of as a kind of individual benefit that accrues to the agent in question.[195] After arguing that the primary task of a theory of well-being is thus to determine the conditions under which my life is good 'for me' in this particular sense, Feldman then separates off and sets aside the distinct issues of whether my life is 'morally' good, or good because it is beneficial to others, or aesthetically good, or again good in the sense of an exemplar or paradigm.[196]

As with the first assumption we considered, this move is also currently ubiquitous and reflective of most contemporary analytical Appraisive accounts of the good life more generally: well-being, in the quasi-technical sense investigated in this literature, is only about the goodness of a life in this private sense, and synonyms for 'well-being' here include 'self-interest' and 'prudential value'.[197] But although analytic philosophers may be entitled to thus carefully pick their special conceptual problems and then restrict attention to these without issue, when it comes to these general problems of practical guidance particular formulations of the core questions embody substantive values, since they now translate into determinate conclusions about how to live.[198] And although it is

[194] See Section 1.7 below for more on this point.

[195] Cf. the previous section's discussion of self- and other-regarding focus at the level of Action Selection.

[196] Feldman, 2004, p. 9; see also Sumner, 1996, pp. 20-25 and Griffin, 1986, p. 21 for the same move being made by perhaps the two other most influential recent analytical theorists in this area. See Hurka, 1987, p. 73 for a critical discussion of various senses of 'good for'.

[197] Crisp, 2021, Section 1.

[198] Feldman, 2004, p. 12, footnote; cf. Finnis, 2011 [1980], p. 15-17 on concepts implying substantive judgements.

our lives that are in focus when we deliberate about what kind of life to aim for, nothing we have said so far about this core practical problem has committed us to understanding this exclusively or even primarily in terms of a personal 'benefit' in this narrowly individualist – perhaps even faintly consumerist – sense. The general problem of Life Planning is simply the question of what kind of life we should aim to live: the most choiceworthy life and the one we have most reason to strive for.[199] What is wanted is thus simply a Life Plan that embodies the 'best' way to make sense of our lives as a whole, where 'best' is understood in the broadest, all-things-considered sense – and the different ways in which a life can be good that Feldman catalogues might then be merely special factors in this overall choice, insofar as they are separable or even cleanly distinguishable at all.[200] So, irrespective of the validity of this assumption when it forms part of these Appraisive enquiries, we do not yet have sufficient grounds for a parallel narrowing of our enquiry into Life Planning.[201]

Being cautious about this second assumption of most contemporary analytical work on well-being is again especially important for the present critical enquiry because the resulting restriction of our core question would implicitly encourage and support just those Subjectivist standpoints I will be concerned with – as well as mask certain limitations of these views which I shall aim to expose in the final chapter.[202] In contrast, the assumption of the Human Good as a kind of individual reward or benefit also again systematically undermines rival accounts that might otherwise be quite attractive alternatives, such as some of those offered by the ancients.[203] Bentham may indeed receive support

[199]Cf. Collingwood, 1992 [1940], p. 419; see also Scanlon, 1998, p. 112 and p. 131, who again distinguishes the choiceworthiness of a life from 'well-being' in the usual philosophers' sense.

[200]Feldman indeed suggests that ancient philosophy might instead have been interested in a 'truly excellent life' conceived as a conjunction of these various interpretations, and is careful *not* to advance a hedonistic theory in this arena; Feldman, 2004, p. 12.

[201]Cf. Williams, 2006 [1985], pp. 5-6, who argues that 'There is only one kind of question to be asked about what to do', and that it is not ambiguous; see also *ibid.*, p. 19: 'Socrates' question, then, means "how has one most reason to live?" ... no prior advantage is built into the question for one kind of reason over another.'

[202]Cf. Taylor, 1992, p. 82. See the next section for discussion of the relation between Subjectivism as I understand it and the intrinsic valuing of others; and again, for brief discussion of a more enlightened version of EUT that considers 'social preferences' relating to others, see the previous section and especially Section 2.1 below.

[203]The dominant contemporary philosophical approach does allow that this private benefit may *depend on* other-regarding considerations: 'we can allow that my well-being depends on yours without introducing the confusing notion that my well-being

from many contemporary readers today in his forceful dismissal of the traditional view that acquiring the virtues could be part of the good life: 'What benefit, in any shape could be derived from impregnating the memory with such nonsense? What instruction from a self-contradictory proposition?'[204] Yet the restriction of our enquiry into the kind of life we should live to a narrow focus on well-being *qua* individual benefit may seem natural only from the perspective of our highly individualistic culture; one which seeks to downplay our dependence on others and our inseparability from our social roles. On a traditional Aristotelian view, for instance, things appear quite differently:

> For what education in the virtues teaches me is that my good as a man is one and the same as the good of those others with whom I am bound up in human community. There is no way of my pursuing my good which is necessarily antagonistic to you pursuing yours because *the* good is neither mine peculiarly nor yours peculiarly – goods are not private property.[205]

And the differences in assumptions and conceptual frameworks here have led generations of modern commentators to accuse Aristotle of conflating the more narrow contemporary notions of prudential and perfectionist value in his enquiries into the human *telos*; to claim that he illicitly changes the subject 'from the good (welfare) of a human being to the goodness (perfection) of a human being'.[206] So again, in order to make the best use of the available resources, and so that possible rivals to Subjectivism remain visible, we must avoid tacitly subjecting ourselves to this individualistic assumption of self-interest from the outset; rather, prior to presenting our contending theories of practical deliberation in detail, in the context of Life Planning we must for now remain 'entirely noncommittal ... about the kinds of consideration to be applied to the question'.[207]

---

is constituted by yours' – Crisp, 2021, Section 1.

[204]Bentham, 1834, p. 301; Bentham is likely thinking of other-regarding Christian virtues such as compassion, abstinence and piety here; Mill, meanwhile, is far more optimistic about the virtues – see Mill, 1998 [1863], p. 82, quoted in the following section.

[205]MacIntyre, 2011 [2007], p. 266; cf. Annas, 1993, p. 9.

[206]Sumner, 1996, p. 79-80; see Toner, 2006 for defence of Aristotle; and Wilkes, 1978 for further discussion.

[207]Williams, 2006 [1985], p. 5.

1.7 Intrinsic and Instrumental Value

In this section, we draw an important distinction between an item being intrinsically valuable for an agent *versus* being only instrumentally valuable – that is, a thing's being valuable for its own sake *versus* its only being valuable because of its connection to some further item.[208] We then discuss how this distinction applies to the Subjectivist views under consideration. Here we will again depart somewhat from the current philosophical literature, being specifically interested only in a particular conception of this distinction as it arises within the Action-Guiding framework we are developing.[209]

Within the Appraisive study of well-being, the term 'intrinsic good' is generally used to denote those features of human lives that directly contribute to explaining why they go well rather than badly, from the point of view of the objective, external assessment that an account of well-being is taken to offer. These directly good-making features contrast with those considered merely 'instrumentally valuable' – that is, valuable only insofar as they facilitate the attainment of those intrinsic goods that well-being comprises. Moreover, for assessments of well-being the latter are superfluous: on an Appraisive hedonist view, for instance, it does not much matter *how* an agent's life came to have a certain amount of pleasure – merely that in actuality their life did contain this amount.[210]

Within the present enquiry, in contrast, we are following the logic of Problem and Solution sketched above, and so for us the distinction between intrinsic and instrumental value is of importance only insofar as its application results in determinant practical consequences for what we are to do or how we are to live. Rather than being concerned with goods in their capacity for providing a rational justification for philosophers' external assessments of well-being, we are instead interested in them only as objects of choice that may be afforded a more or less privileged status within everyday moral agents' practical deliberations, in either their Action Selection or Life Planning modes. Our focus here is the mechanics of real-life human practical reason and its intentional structure: 'It does not matter to us what analysis can be applied to an action; what matters is to find out what analysis the agent applies to it in his own consciousness

[208] For the time being, we allow such an entity to be either a state of affairs, a fact, an object, a person, a property or trope, an activity, or anything else.

[209] See Zimmerman and Bradley, 2019 for discussion of the general philosophical issue.

[210] Fletcher, 2016; cf. Section 1.1 above.

of doing it, or having done it, or being about to do it.'[211] But what is this structure? We begin with the context of Action Selection, before moving on to Life Planning.

In his insightful treatise on moral psychology, Collingwood explains that two actions are connected as means and end if they can be viewed as part of a single larger action; that in doing the first we are conscious of doing only part of a more complex action that will later reach completion with the second. This distinction is central to his exposition of utility, which – despite Mill's protests – for him can only mean 'usefulness', and which he sets alongside 'goodness' and 'rightness' as motivating ethical reasons.[212] He elaborates:

> We have clearly seen that people become aware of that special form of goodness which is called utility as soon as they learn to perform a certain operation of thought, namely the operation of dividing a single act, the act I am now doing, into what I am immediately doing and what I am thereby mediately doing: in other words into means and ends.[213]

So, a person may be ..

> .. conscious of what he is mediately doing as good in itself, or good as an end, (that is to say, he becomes conscious of choosing it for its own sake), and of what he is immediately doing as useful, or good as means (that is to say, he becomes conscious of choosing it for the sake of the thing to which it is related as means to end).[214]

Following this account, and bearing in mind the discussion of reasons and goods in Section 1.5, I will say that a normative theory of Action Selection deems an entity to be an 'Intrinsic Good' (or 'Intrinsically Valuable') when it considers it worthy to pursue as an end in itself.[215] As such, it is capable of providing an agent with a reason for undergoing a purposive action that is complete in itself (though possibly not a *sufficient* reason,

[211] Collingwood, 1992 [1940], p. 461.
[212] Cf. Mill, 1998 [1963], pp. 54-55.
[213] Collingwood, 1992 [1940], p. 452; cf. Anscombe, 2000 [1957], p. 22; p. 36; p. 79.
[214] Collingwood, 1992 [1940], p. 452.
[215] This definition thus makes no mention of such value being based on 'intrinsic properties'; cf. the worry expressed in Korsgaard, 1983, who suggests that we use the terms 'final value' and 'instrumental value' here to avoid conflating this distinction with that of the kinds of properties – relational or non-relational – by virtue of which something has value.

since the achievement of greater Intrinsic Good may also be possible).[216] And the same appellation will apply when choosing between fixed options rather than undergoing a spontaneous action: the achievement of an Intrinsic Good provides us with a reason for selecting one object of choice rather than another in this sort of case too.[217]

In contrast, a normative theory of Action Selection will consider an item to be a merely 'Instrumental Good' (or merely 'Instrumentally Valuable') when it allows only that it justifies a mediate action that culminates in the achievement of some Intrinsic Good, to which it is related as a means; or else, when it indicates that it should be chosen from a selection of options only because of its connection with the achievement of such a good. For instance, an account of Action Selection may recommend that we seek an antidote to a deadly poison as a means to the Intrinsically Valuable end of survival, say, though only so long as the poison still threatens our lives and the antidote remains effective: the antidote is considered contingently valuable as a means to the end of good health, but once we are cured we no longer have reason to ingest it, since 'without the disease there is no need for the drug'.[218]

We now consider how this distinction applies within the present critical enquiry. As we have said, according to the Action-Guiding theories we will consider, the agent must look within themselves – to their inner Subjective Values – to ascertain what is good and best for them. Within a purely Subjectivist account of Action Selection, then, it is only the satisfaction of these Subjective Values that can be considered Intrinsically Good in the above sense. Any other item – be it material success, the achievement of a career goal, physical health, or the quality of our moral character – can only be Instrumentally Good at best. For what our Subjectivist Theories recommend when applied in this context is that such items should be objects of choice only insofar as they facilitate the satisfaction of our desires or the provision of pleasure. And crucially for our later enquiries, this extends to outcomes involving other people, too: whether this be the good of another person as an end in itself, or any aspect of our particular relationships with them. For outcomes such as these may be considered valuable to pursue only insofar as their

[216] Cf. Dunne, 1993, pp. 101-102 on action for its own sake.

[217] Again, this is the case that economists have primarily focused on; see Section 2.1 below.

[218] Plato, *Lysis*, 220d [All references to Plato are from John M. Cooper's complete works].

achievement serves Subjectivist ends.[219]

In some instances, it may be the case that the connection between the satisfaction of our Subjective Values and some particular extra-Subjective item is particularly robust: the well-being of a beloved spouse that we feel we cannot live without, for instance. But nevertheless, where this is the case, the inferior Instrumental status of the good in question may still be revealed counterfactually: if we consider a scenario where the usual causal link becomes severed, the theory must then recommend that the good in question should no longer be actively sought, nor used as a basis for our choices.[220]

In a well-known passage, Mill contravenes this stipulation by instead giving a strikingly central role to the extra-hedonistic good of virtue within the deliberations of the ordinary moral agent:

> does the utilitarian doctrine deny that people desire virtue, or maintain that virtue is not a thing to be desired? The very reverse. It maintains not only that virtue is to be desired, but also that it is to be desired disinterestedly, for itself. Whatever may be the opinion of utilitarian moralists as to the original conditions by which virtue is made virtue ... they not only place virtue at the very head of things which are good as a means to the ultimate end, but they also recognize as a psychological fact the possibility of *its being, to the individual, a good in itself, without looking to any end beyond it*; and hold, that the mind is not in a right state, not in a state conformable to Utility, not in the state most conducive to the general happiness, unless it does love virtue in this manner.[221]

A crucial difference between our approaches now becomes apparent: although recognising pleasure as the sole 'ultimate end' here, Mill establishes this claim via a mode of justification that does not – indeed, for him, *should* not – find any counterpart in the reasoning of plain persons.[222] Because his framework of enquiry licenses this tactic, his hedonistic theory can indeed support a robust and inviolable role for virtue within the deliberations of ordinary moral agents, who are to thus regard

[219]See Sections 1.5 above and 2.1 below on 'social preferences' within EUT.

[220]Cf. Anscombe, 2000 [1957], p. 36. However, later I will argue that practical reason is essential to the *pursuit* of the Subjectivist's goal; see Section 3.1 below.

[221]Mill, 1998 [1963], p. 82 [my emphasis]; see also Driver, 2014 on Mill's view.

[222]For criticism of the self-effacing nature of utilitarianism here, see Williams, 2006 [1985], pp. 107-108, and Section 1.2.3 above.

it as an Intrinsic Good in the above sense; even though *qua* enlightened moral philosopher his supporting logical justification for this concession ultimately appeals to considerations of utility alone. However, within our Action-Guiding framework this line of argument is not permissible, since here an account of the practical reasoning of everyday agents is itself the primary content of the theory, and such self-effacing moves contribute nothing of use here: 'a wheel that can be turned though nothing else moves with it, is not part of the mechanism.'[223]

Within the present framework, then, any pure hedonist in the context of Action Selection must advocate that plain persons uphold the virtues only under certain contingent conditions, and thus be willing under a range of circumstances to systematically sacrifice virtue in pursuit of greater amounts of pleasure. For otherwise, their practical reasoning would not have a distinctively hedonistic character, which requires them to value other items only insofar as they are contingently connected with pleasure. And this point continues to hold if – as Mill claims, and *contra* Bentham – the contingent conditions linking virtue and pleasure happen to be widely prevalent in practice. But in fact there may be serious obstacles to understanding genuine possession of the virtues in this way:

> For it turns out to be the case that – and this is in part at least one more empirical factual claim – although the virtues are just those qualities which tend to lead to the achievement of a certain class of goods, nonetheless unless we practice them irrespective of whether in any particular set of contingent circumstances they will produce those goods or not, we cannot possess them at all.[224]

Indeed, for MacIntyre, as for Augustine, Aquinas, and Aristotle, the virtues are an 'Internal Means' to the highest kind of human life – where a means is 'Internal' to a given end when the end itself cannot be adequately characterised independently of the means. Hence, the virtues have a more robust connection with their associated end than would be possible if they were of only Instrumental Value, receiving their rational justification *both* as a means to *and* as a constituent part of *eu-*

[223]Wittgenstein, 2003 [1953], p. 271. See also Section 1.2.3 above, and Mill, 1998 [1963], p. 63, on deliberately cultivating 'the conscious ability to do without happiness'.

[224]MacIntyre, 2011 [2007], p. 230; cf. *ibid.*, pp. 317-318; and MacIntyre, 2009b [1999], pp. 111-112.

daimonia.[225] The importance of this further distinction will also become apparent in Chapter 3, though here our focus will be our relationships with others rather than the virtues *per se*.

One limitation of the stipulation imposed here is that some forms of satisfaction of our Subjective Values may only be available by allocating Intrinsic Value to certain other goods at the level of Action Selection: certain pleasures that might only be obtained through a genuine, committed devotion to the church, for instance. Moreover, we can easily imagine agents to whom this type of pleasure is especially important, or who harbour desires of this nature that are particularly pressing. This point brings us now to the broader perspective of Life Planning, since we will see that deliberation at this level will to some extent render such goods available for the Subjectivist's use.

The precise nature of how the Instrumental *versus* Intrinsic Goods distinction emerges in the context of Life Planning will depend on the structure of the theory of the good life in question. For instance, if the theory takes the form of a list of essential 'ingredients' for the best kind of life, which it insists must always be part of any adequate Life Plan under a full range of possible contingencies, then these items will thereby be the Intrinsic Goods it recognises. Those items considered merely Instrumental Goods, meanwhile, may enter into an agent's Life Plan only when they happen to be causally necessary for pursuing these Intrinsic Goods: where they instead conflict with the achievement of these goods, or perhaps even merely fail to contribute to this end whilst consuming our limited time and energy, they must be extirpated from our pursuits entirely.

The Subjectivist theories I will be concerned with in this thesis have this sort of structure, though each gives a very short list of just one item. So, for the hedonist, the agent's life must first and foremost be an optimally pleasurable one. In filling in the details about what contingent priorities other than this they should adopt, an agent might be left to work backwards from here, taking into account their particular circumstances to develop a coherent Life Plan based on this sole principle. Yet a hedonist might also go beyond this minimal characterisation and supply further resources to aid here – such as offering empirical generalisations about what pleases us.[226] In the limit, the contribution of theory here

[225]MacIntyre, 2011 [2007], pp. 214-215; p. 174; cf. Aristotle, *NE* I.7, 1097b2-5; 1098a15-17; VI.5, 1140b6-7; and Angier, 2015, p. 9.

[226]For discussion of the empirical literature on this topic, see Section 2.4 below.

may even extend to a more-or-less determinate way of life: consider Epicureanism as a comprehensive philosophy of living, for instance.[227]

We have seen that the province of Life Planning encompasses internal aspects of our character, and hence our modes of decision-making in the here-and-now.[228] Hence, some of the principles for living arrived at – whether by the individual agent themselves, or having been explicitly proffered by an Action-Guiding theory – may advocate that we allocate Intrinsic Value to extra-Subjective sources of the good in our concrete choices in particular situations. And this route is open to any version of Subjectivism that applies directly at the level of Life Planning, whilst regarding Action Selection as subordinate.[229]

In the next chapter, we shall indeed aim to develop such versions of Subjectivism in some detail.[230] However, we shall also see that even here some limitations in the status that may be afforded to extra-Subjective goods will remain. For instance, although we may seek 'The happiness which belongs to the peculiarly to the way of life of the cloister'[231] with some seriousness, still Subjectivism will permit us to keep to this path only insofar as this *does* in fact result in a pleasurable life for us. An agent cannot commit permanently to such a life once-and-for all, on pains of abandoning Subjectivism as an ongoing mode of practical rationality. But with human life being as varied and unpredictable as it is, there is again always the possibility that the counterfactual comes into effect: although other entities may be intrinsic to certain particular kinds of pleasure, they cannot be intrinsic to our being pleased as such – and if the monkish life does cease to please us, Subjectivism must now enjoin us to abandon it. Even on this broader and more inclusive approach, then, it is only the satisfaction of our Subjective Values that is Intrinsically Good, with the pursuit of other goods remaining only ever conditional upon this core objective. And again, the key point for us later will be the limitations that this places on the nature of the relationships with others we may enjoy under these constraints on what our practical reasoning can be.

Returning now to Aristotle's enquiries into the good life in the *Nicomachean Ethics*, here we see a further complication in the relation between

[227]See Section 2.3 for a brief overview.
[228]See Section 1.4 above.
[229]On this point, see Section 1.4 above.
[230]See in particular Sections 2.2 and 2.4.
[231]MacIntyre, 2011 [2007], p. 177.

Intrinsic and Instrumental Goods as it takes shape within the broader temporal perspective of Life Planning. In common with our Subjectivist views, and like many other ancient thinkers, Aristotle takes it for granted that here an agent's practical deliberations must centre on one single 'Final End', 'with reference to which he will then do all his acts'; otherwise 'our desire would be empty and vain', so that for him 'not to have one's life organized in view of some end is a mark of much folly'.[232] Thus, for Aristotle, any adequate Action-Guiding conception of the good life must be organised around a single core goal: 'a single final end is what is required to make sense of a single life as a whole'.[233] And he of course goes on to formally identify this complete end as *eudaimonia*, stating that it is the ongoing achievement of this Final End (or *telos*) that constitutes living well and being well, before turning to the substantive question of what *eudaimonia* actually consists in. Yet it is not the case that all goods that are merely Instrumental with respect to *eudaimonia* are of equal deliberative status. Rather than a simple dyadic structure, what Aristotle instead describes is an overlapping hierarchy of 'for the sake of' relationships which makes plain the relative importance of the corresponding activities through which the Human Good may be achieved.[234] And Aristotle further argues that such complex nested structures are characteristic not only of human lives, but also of collective human endeavours in general, such as the sciences and crafts of all kinds.[235] For instance, the art of bridle-making not only facilitates but is subordinate to that of horse-riding, since if the practice of horse-riding were universally abandoned it would follow that human beings would have no need for bridle-making and that this activity should be abandoned also, since it would then become devoid of purpose.

Aristotle's observations about the deliberative structure of Life Planning in practice show that one and the same good is often considered both as an end relative to some associated means, and also as a means relative to a further end. Further, his arguments also suggest that given the complexity of human lives as a whole, what agents need in a Life Planning context is a structured list of priorities.[236] But this in

[232]Aristotle, *EE* I.1, 1214b10-11.

[233]Annas, 1993, pp. 31-33; for further defence of this view, see Annas, 2011; and Aquinas, *Summa Theologiae*, IaIIae, q. 1, a. 5: 'the will of an individual man must be fixed on one last end'; see also Anscombe, 2000 [1957], p. 21; pp. 33-34 for discussion.

[234]See Annas, 1993, p. 31 for discussion.

[235]Aristotle, *NE* I,1, 1094a3-16.

[236]Cf. again Williams on 'deliberative priorities'; Williams, 2006 [1985], p. 183.

turn yields a further requirement for the vindication of the Subjectivist accounts we will be concerned with in this thesis. As well as being the only absolutely ineliminable feature of any adequate conception of an ideal life,[237] the central Subjective element – whether it be desire or pleasure – must also be capable of acting as our Final End: that is, that all our existing projects and goals (which it does not advocate the abandonment of) can be understood as subordinate to and ultimately contributing to it as we develop our Life Plans, so that it functions as a focal point for organising and directing our lives as a whole.[238] And the other, lesser goods to be thus ordered and evaluated must of course include our social relationships too. But if it could instead be shown that our Subjective Values are in fact unsuitable in this regard, and that from this broader temporal perspective they must on at least some occasions be thought of as subordinate to other goods that extend beyond them, then our purely Subjectivist theories would stand defeated as comprehensive Action-Guiding accounts of the Human Good in the context of Life Planning. Yet this is precisely the claim I shall defend in Chapter 3.

1.8 Conclusion

In this introductory chapter, we have developed a framework of enquiry that is tailored to the goal of the overall project: to investigate theories of the Human Good that can serve as Action-Guiding ideals for practical deliberation in the real world. It is, I believe, a strength of this framework that it is flexible enough to be neutral between the rival theories offered by the various moral traditions we shall consider, whilst simultaneously bringing out their differences as concretely manifested in different conceptions of practical reason. In arriving at this point, we have also identified in a rough way the tasks that such a theory must perform, as well as introduced a range of further distinctions and concepts that we shall make use of later on.

In the next chapter, we discuss in detail our two Subjectivist accounts of Action Selection and Life Planning. We will also take note of certain theoretical and practical abilities that a rational agent must have in order

[237] Though again, practical agency is essential to the successful *pursuit* of the good; see Section 1.3 below.

[238] Cf. Hardie, 1965 for a classic critical article on *eudaimonia* as an 'inclusive' end.

to benefit from the contribution of these theories here; the development of these abilities and their background prerequisites will then be a major theme of the final, critical chapter.

Chapter 2

Two Subjectivist Accounts

> *All rational behaviour is aimed at maximizing utility – satisfaction, happiness, benefit or pleasure from certain outcomes. Given this definition of utility, any rational person would want to maximize their happiness. This seems straightforward enough and to want to do otherwise would seem foolishly self-defeating.*
>
> – Corr and Plagnol[239]

This chapter sets out our two subjectivist conceptions of the Human Good – desire-based, and hedonist – and considers what these have to offer as ideals that guide our practical reasoning: both as we decide how to act in the here-and-now, and in thinking about the direction of our lives as a whole. The key idea is that agents must look within themselves, to their inner Subjective Values, to determine what it is best to do.

In fully exploring this basic notion, we will need to draw on insights from all three of our key approaches: the careful distinction-making of analytic philosophy; the formal framework for rational choice and correlate empirical studies provided by economics and psychology; and the more flexible and comprehensive – but following advances in our understanding of human psychology, now in some ways dated – standpoint of ancient ethics. Along the way, I will raise some limitations of these Subjectivist accounts, though these will be preliminary and suggestive rather than conclusive. I will also defend a claim that is important for the later main argument against them: that for both types of theory, a certain capacity for practical judgement on the part of the agent would be required to render them viable as a practical guide to life.

[239]Corr and Plagnol, 2019, p. 71.

2.1 Decision Theory

The first type of Subjectivist theory of the Human Good I will discuss takes the fulfilment of our desires as providing the standard of what is good and best for us. In an Action-Guiding context, such theories hold that when an agent deliberates about what to do or what kind of life to live, they should look inwards to their existing desires for counsel. Our Subjective Values are thereby afforded a role that is primarily *epistemic* here, informing the acting agent about what the Human Good is – rather than bearing the metaphysical role of *making* something good in an abstract sense, as takes focus in the Appraisive enquiries of contemporary philosophy.[240] In this section, we begin at the level of Action Selection.

Rather than advocating a specific theory of desire, which is bound to be somewhat controversial, the overall argument will allow theories to appeal to desires in any possible sense.[241] However, in this section we begin with the more specific conative notion of 'utility', and spend some time examining an influential conception of practical reason in which it is at home: EUT (Expected Utility Theory). Dating from the end of the 19th Century, this developed mathematical framework has now become the mainstream approach to understanding individual choice within academic microeconomics, and enjoys widespread influence within the social sciences more broadly.[242]

Happily, we might think, the problem of what it is to make concrete choices in the here-and-now in pursuit of a desire-based conception of the Human Good is one that economists have worked out a solution to in great detail. For making choices on the basis of maximising utility – or expected utility, where there is uncertainty present – is what contemporary microeconomics is all about. And the resulting framework has reached such a pitch that we might hope a technical, algorithmic solution to our central problem of Action Selection has already been achieved. However, we should be cautious here: although we are specifically interested in the contribution that the framework can make to developing a normat-

[240]This particular sort of view is also opposed to alternative formulations where it is moving to the *state* of having one's desires satisfied that is good – though possibly because these desires have changed in the meantime; however, the arguments below are readily adaptable to this version of the account too.

[241]See Section 1.2.5 above; see also Schroeder, 2015 for some prominent philosophical theories of desire. Some analytic philosophers term this broad category 'conative' theories; cf. Kraut, 2007, pp. 94-95.

[242]See Mas-Colell, Whinston, and Green, 1995 for one of the most comprehensive graduate-level presentations of the mainstream approach.

ive theory of real-world deliberation, many practitioners understand the goals of EUT quite differently: either to describe and understand the mechanics of choice in practice, or else simply to model and successfully predict overt choice behaviour itself.[243] So, let us first look more closely at this rich tradition of enquiry on its own terms.

As part of their initiation into the mainstream approach, from undergraduate level onwards, students of economics learn to use EUT to solve a variety of standard problems by calculating the choice that maximises an agent's expected utility. The approach has been used to model a huge range of individual choice problems, and with the inclusion of techniques from game theory we can also model strategic interaction between more than one agent: we now have standard models of whether to accept a risky gamble, how much education to obtain, how many hours to work and how many to reserve for leisure, how to select a portfolio of bonds and risky assets, whether to purchase an object of unknown quality, how much to bid in an auction, how to design an optimal auction, and how to engage in bargaining and voting behaviour, amongst many others. Perhaps the most basic of these is the consumer choice problem of deciding which items to purchase from a fixed menu of goods, given certain prices, a utility function representing the desirability of outcomes, and a monetary budget. For concreteness, a typical example might run as follows:

> Matthew has £50 to spend at the pub. A beer costs £4, and a pizza costs £6. His utility function is $\mathscr{U}(x, y) = x^{\frac{16}{25}} y^{\frac{9}{25}}$, where x is the number of beers and y is the number of pizzas he consumes. What is his optimal combination of purchases?

Using standard techniques,[244] it is possible to show that the mathematically optimal solution is $(x, y) = (8, 3)$; since these are also integers, we thus conclude that Matthew should purchase 8 beers and 3 pizzas.

In real life, however, things are not straightforward as they may seem in the artificial context of the exam hall. For if one actually goes up to Matthew in the pub, or to any other ordinary person, and attempts to advise them in this way, one soon finds that although information about

[243] For discussion of these three interpretations of EUT, see Bermúdez, 2009, Chapter 2 and especially p. 43; see also Gul and Pesendorfer, 2008, Section 1 on the limited aims of standard economic analysis.

[244] For instance, one may rearrange the budget equation and substitute it into the utility function and then set the first derivative equal to zero – or else set up a Lagrangian. In practice, for such a simple problem it is perhaps most common to use a tangency condition.

prices and budget may be forthcoming, they will not simply report their utility function to us, and indeed are unlikely to say anything useful when asked about it. This brings us immediately to the central question: 'What is *utility*?' Yet here we soon find that different schools of economic thought have interpreted this key notion in different and incompatible ways.

In making sense of this complex debate, it is useful to distinguish between 'substantive' (or 'realist') and 'constructivist' (or 'formalist') conceptions of utility. According to substantive approaches, talk of utility refers to an 'independently specifiable quantity' that is a real feature of the agent in question; hence, 'To say that an agent maximizes expected utility is to say that the agent chooses the course of action that has the largest mathematical expectation of that quantity.'[245] According to constructivist approaches, in contrast, a utility function is merely a convenient way to represent an agent's 'preferences'. The constructivist approach is the dominant interpretation amongst mainstream economists today,[246] following the 'ordinalist revolution' of the 1930s, as orchestrated by Fisher, Pareto, Hicks, Allen and others; so, we first examine utility in this sense, beginning with the more fundamental concept of preferences.[247]

In orthodox consumer theory, agents' preferences are understood quite minimally, representing only how they would rank-order the various available options within a particular choice context. Taken collectively, the totality of these options is termed the 'choice set' of a problem. In the typical case, this set consists of 'bundles' that contain numerical amounts of each of a fixed array of divisible consumption goods. Bundles can thus be represented as vectors whose components are non-negative real numbers; economists often work with just two goods, since insights from this simple case are usually easy to generalise. The core notion developed here is that of a 'weak preference relation' of a particular agent, or 'consumer'. Mathematically speaking, this can be thought of as a purely formal object: specifically, a binary relation that may hold for pairs of options within the choice set. When applied to actual agents, this relation expresses the idea that a consumer 'weakly prefers' one option within a particular pair to another, or else is indifferent between them;

[245] Bermúdez, 2009, p. 47.
[246] Broome, 1999, p. 28
[247] For a short overview of this history, see Morgan, 2003, pp. 278-281.

we discuss the interpretation of this intuitive notion further below.[248]

Preferences over bundles must satisfy two axioms to be counted as 'rational' according to standard theory. The first is completeness: that for any pair of bundles $\boldsymbol{x}$ and $\boldsymbol{y}$, either $\boldsymbol{x}$ is weakly preferred to $\boldsymbol{y}$, or conversely.[249] The second is transitivity: for any three bundles $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}$, if $\boldsymbol{x}$ is weakly preferred to $\boldsymbol{y}$, and $\boldsymbol{y}$ is weakly preferred to $\boldsymbol{z}$, then $\boldsymbol{x}$ is always weakly preferred to $\boldsymbol{z}$ also. These axioms sound quite plausible as necessary conditions – though counterexamples to both as a requirement for reasonableness have been presented.[250] However, the resulting conception of rationality is extremely permissive, in that these assumptions appear to make no substantive restrictions on the *content* of agents' preferences, even though we would not normally say that an agent was 'rational' if their preferences were pathological or extremely eccentric.[251]

If further technical conditions are satisfied, an agent's preferences can be 'represented' by a 'utility function' that assigns numerical values to bundles of goods – where 'representation' means that more preferred bundles are assigned higher numbers. If the choice set is representable as a set of vectors and is convex, then a property called 'continuity' suffices: very roughly, that numerically similar bundles are placed nearby in the agents' ordering.[252] The 'Cobb-Douglas' form of utility function used in the example above is a common choice in research contexts, and has a number of desirable mathematical properties:

$$U(\boldsymbol{x}) = U(x_1, x_2) = x_1^{\alpha} x_2^{1-\alpha}$$

We now have a formal model of utility, which reflects how it is generally understood on undergraduate and graduate-level microeconomics courses today. However, if the framework is to be applied to real-world agents, then it must also be possible to ascertain what an agent's preferences

[248]See Hansson and Grüne-Yanoff, 2022 for further interpretive issues and a much more detailed discussion.

[249]Including the case where both relations hold – that is, where the agent is 'indifferent' between the two bundles.

[250]For instance, an agent may have no explicit preferences over options they have never considered; some economists therefore only require that the preference relation may be *extended* to a complete one – see Ok, 2002. For arguments against requiring transitivity, see Temkin, 1996; Temkin, 2012; and Quinn's 'self torturer': Quinn, 1993, pp. 198-209.

[251]Cf. again Rawls' grass-counter; see below for skepticism of this apparent neutrality, and see the next section for discussion of implicit restrictions on the content of agents' desires.

[252]Debreu, 1954.

actually are in particular cases – and this again requires a more detailed interpretation of the idea of 'weak preference' itself.

Perhaps the dominant interpretation of the underlying weak preference relation is given by 'revealed preference theory'. Popularised by Samuelson around the middle of the 20th Century, this approach focuses on the observation of real-world purchasing behaviour.[253] According to this view, a consumer weakly prefers one bundle to another just when they *might* (or actually *did*) choose it given the option of both.[254] This understanding is purely behavioural and is independent of any underlying internal processes or assumptions about the inner inclinations or feelings such a choice might reflect; thus, 'every psychological analysis is eliminated'.[255] This was seen as an advantage, since following the behaviourist school of Watson and Skinner in psychology, in the first decades of the 20th Century, economists increasingly converged on the view that inner experiences are hard to study scientifically and thus to be avoided as a foundation for their enquiries into consumer choice.[256]

When constructivism about utility is combined with this behaviourism about preferences, the result may be termed 'operationalism' about utility. Here utility is no more than a representation of an agent's (hypothetical or actual) choice behaviour.[257] Utility functions in this sense are widely used in many empirical applications and can be successfully fitted to particular agents in practice. In a laboratory setting, methods of reading off agents' preferences in practice are known as 'elicitation' techniques; these engage them in simple, accessible, but somewhat artificial tasks that are intended to reveal their broader attitudes.[258] And in modern 'micro-founded' macroeconomics, vast data sets and sophisticated statistical techniques are used to equip agents with utility functions that enable the prediction of future behaviour – sometimes on

[253]Samuelson, 1938. For a clear recent statement and defence of a minimal revealed preference approach, see Vredenburgh, 2020; see also Thoma, 2021.

[254]We say 'might' here because 'weakly' signifies that we include the case where they are indifferent and thus could also choose the other option; if their preference were 'strict' they would always choose it.

[255]Pareto [1900]; quoted in Bruni and Sugden, 2007. Other early arguments against earlier hedonistic interpretations of utility were given by Fisher, Hicks, and Allen around the 1930s; see e.g. Hicks and Allen, 1934.

[256]Meanwhile, positivism became a similarly dominant force within the philosophy of science – see Dietrich and List, 2016; for a 'rational reconstruction' of the history in economics, see Muramatsu, 2009.

[257]See Clarke, 2016 on the operationalist view of preferences; for further discussion, see again Bermúdez, 2009, Chapter 3.

[258]See e.g. the procedure described in Gneezy and Potters, 1997.

the scale of entire populations.

Despite the practical success of this ordinalist framework, a key limitation remains: it does not enable us to speak coherently about 'utility' in contexts where risk is present. For utility functions as defined so far only ensure that bundles are ranked in the correct order, and as such do not encode information about *how much* agents prefer some outcomes to others. An agent's 'utility function' is thus not uniquely specified – and indeed, *any* strictly increasing transformation will yield a new utility function that also 'represents' the same ordinal preferences just as adequately. For this reason, ordinal utility is insufficient when considering situations of choice under uncertainty, where we cannot choose final outcomes directly but only probabilistically, as here we must balance the attractiveness of different outcomes against their associated likelihood of occurring. And the same issue arises when modelling economic choice contexts that extend over time, since negotiating the trade-off between material well-being across the different periods in question also necessitates weighing the values of different outcomes in a quantitative way.[259]

In extending the EUT approach to risky choice, John von Neumann and Oskar Morgenstern instead considered an agent choosing from a set of 'lotteries'.[260] A lottery is here a scenario whose outcome A may take each of a set of possible values $(a_1, ..., a_n)$; in economic applications, these outcomes (or 'prizes') are generally represented as real numbers, and often thought of simply as monetary payouts. Outcomes then occur with particular 'objective chance' probabilities $(p_1, ..., p_n)$ that are known by the agent.[261] In 1947, they published a proof of what is now called the 'von Neumann–Morgenstern Representation Theorem': that if an agent's preferences over these *lotteries* themselves satisfy certain axioms, then we can express their preferences over the ensuing *outcomes* by means of a *cardinal* 'Bernoulli' utility function u, now unique up to a *linear* transformation, in such a way that the expected value of this Bernoulli utility for a given *lottery* correctly represents the agent's preferences over *lotteries.* Moreover, this expected utility score – the corresponding 'von Neumann–Morgenstern utility' of the lottery itself – ranks lotteries on an *interval* scale – that is, in such a way that the magnitudes of the differences in scores between pairs of lotteries can be meaningfully

[259] For a classic discussion of ordinal and cardinal utility, see Alchian, 1953.

[260] Neumann and Morgenstern, 2004 [1953]; their approach was inspired by Ramsey, 1931 [1926].

[261] See Resnik, 1987, pp. 88-91 for an overview.

compared.[262]

The axioms on preferences over lotteries required for the proof are completeness, transitivity, and continuity (introduced above), as well as a fourth axiom of 'independence': very roughly, that for lotteries that share some prizes and associated probabilities in common, preferences are determined by comparing only the parts that do *not* coincide.[263] We shall return to the issue of the EUT axioms later.

Although this result marked a big step forward for decision theory, we unfortunately still do not yet have an account of real-world Action Selection that meets the requirements laid down in Chapter 1. A first problem concerns a technical issue: the neoclassical approach as outlined so far assumes that outcomes and their associated 'objective chance' probabilities are known to the agent; yet this will rarely be the case in real life, outside of the casino and other artificial settings.[264] This given, it is natural to instead work with 'subjective probabilities' that represent agents' *de facto* beliefs about the likelihood of different states of the world.[265] The resulting collection of normative frameworks for choice in the real world is called 'Subjective Expected Utility Theory'. Yet determining what these 'subjective probabilities' actually mean and what fixes the values they take again requires further interpretative work.

In an influential treatment, Savage distinguishes three basic components: *acts* (i.e. what the agent does), *states* (i.e. ways that the world may be like, which agents have beliefs about but cannot affect through their actions) and *outcomes* (i.e. possibilities that agents have non-instrumental preferences over). An agent's chosen act then combines with the state of the world to determine an outcome.[266] He then proved that if an agent's (instrumental, ordinal) preferences over possible *actions* satisfy certain axioms, then one can assign both a *unique* subjective probability function P over *states* and a *cardinal* utility function u over *outcomes* in such a way that their initial preferences over actions are

[262] See von Neumann and Morgenstern, 2004 [1953], pp. 617-628; the proof was first included in an appendix to the 2nd edition of 1947.

[263] See Steele and Stefánsson, 2020, Section 2.3 for further details. There are also two further technical assumptions about the set of lotteries here; that they are closed under probability mixtures – that is, new lotteries can be constructed where the prizes are other lotteries – and that such 'compound' lotteries are treated the same as simplified equivalents with numerical prizes.

[264] I.e., agents typically face not precisely-quantified 'risk' but rather 'uncertainty'; a distinction famously drawn in Knight, 1921.

[265] Or in some alternative formulations, the beliefs they *should* have, given their evidence.

[266] Savage, 1954.

represented by the expected value of u, as calculated according to P.[267]

Savage's framework solves this particular technical problem in the context it is intended to be applied. However, at the same time, it is clear that his approach cannot give us what we are looking for here. For in Savage's framework, our preferences over outcomes and subjective probabilities are constructed *from* our given preferences over actions. Hence, they cannot also *inform* these same choices of actions in the way we are hoping for in the present enquiry.

This issue with co-opting Savage's framework for our present philosophical purposes is in fact an instance of a broader problem that extends to any operationalist understanding of utility more generally – making plain the force of the above warning that many EUT theorists do *not* intend to supply a theory of deliberation of the kind we are presently seeking.[268] For in any such operationalist understanding of utility, our choice behaviour determines what our preferences are, and one then constructs a utility function which these choices implicitly maximise.[269] Hence, if at the same time we attempt to utilise EUT as a normative Action-Guiding account of *how to make* our choices, whilst continuing to understand preferences as gleaned only from choice behaviour, then clearly we are threatened with a circularity. Indeed, the only substantive advice that an agent could derive from this framework would be to ensure that their choices are consistent with the basic EUT axioms, and hence 'rationalisable' by some stable set of preferences.[270] But what this amounts to is only the imperative that their present choices must be formally consistent with whatever other choices they happen to have made in the past – in terms of their content, agents are given almost free reign.[271] When an operationalist version of decision theory is put to work in the context of Action Selection as presently conceived, then, the resulting guidance offered will be too minimal to be of much use: 'an 'ought' that it is not so much as possible to flout is not really an 'ought'

[267] The axioms are not without controversy; see Steele and Stefánsson, 2020, Section 3.1.

[268] See again Bermúdez, 2009, p. 43 for the view that decision theory cannot simultaneously be 'an action-guiding theory of deliberation; a normative theory of assessment; and a tool for explaining and predicting action' – though his line of reasoning is somewhat different to that presented here.

[269] See again Broome, 1999, p. 28.

[270] Sen, 1977, p. 325.

[271] See also Broome, 1991, who stresses that decision theory itself is a purely formal mathematical framework; on a technical level, see Varian, 1982 for conditions on observable 'choice functions' being derivable from a rational weak preference relation, given some minimal background assumptions.

at all'.[272]

As well as failing to pin down a robust, substantive account of practical reason, this consistency requirement also introduces a further problem for our purposes. For it is also not even genuinely Subjectivist in the sense at issue in this thesis. We have seen that the restrictions imposed are in one sense too weak, in that they make no reference to the agent's actual Subjective Values – but in another sense, they are also too strong and may serve to override them. For our desires can and do change over time, and are less constrained by our history than a mere catalogue of past choices would be. Dine twice at the same restaurant with an operationalist choice theorist, and one can be expected to have one's rationality challenged; but Matt may just not be so hungry today, even though pizza is now on special offer at the pub.[273] In contrast, the Subjectivist must insist that if our inclinations do change, then what it is optimal for us to choose also changes with them. And for these dual reasons, an operationalist understanding of EUT must fail to faithfully express our Subjectivist theories of the Human Good at the level of Action Selection.

Despite its usefulness for the special problems economists address, then, from our point of view the sacrifices made in the name of rigour have been too great here, and for the present philosophical enquiry a different approach is needed that has a better fit with our requirements. Therefore, we now turn our attention to a different conception of decision theory: an EUT based on a substantivist rather than formalist interpretation of utility. Despite the neoclassical economists' persistent warnings against traversing this path, in a bid to develop a genuine theory of deliberation based on agents' actual Subjective Values we will now venture to treat utility as reflecting real psychological features of agents, but to which their actual choice behaviour is not necessarily faithful.[274] Moreover, it is worth pointing out that the difficulties with observing psychological elements that have been economists' key motive for excluding them are most pronounced from a third-person standpoint, such as is necessary for economists' particular work, as they aim to understand purchasing behaviour from an external, scientific perspective.[275] In contrast, these

[272]Wallace, 2020, Section 5; cf. also Mandler, 2001, p. 383.
[273]Bermúdez, 2009, pp. 50-51; see also the discussion of narrativity in Section 1.6 above.
[274]Cf. Dietrich and List, 2015.
[275]Cf. Scanlon, 1993, p. 195.

internal elements may be to some extent transparent to the agent themselves – and this is all that is required in the present enquiry, where we are taking a first-person perspective.[276]

In terms of the precise identity of the substantive psychological element to be maximised, the thought of some early marginalist economists – especially Edgeworth and Jevons – tended to follow Bentham's hedonism here.[277] And as we shall see later, Action-Guiding theories of choice based on felt experience are undergoing something of a revival in contemporary economics.[278] However, the most common modern versions of realism about utility today are still based on our conative rather than affective responses; in the remainder of this section and the next, then, we follow suit by understanding the psychological element in question in terms of the agent's desires.

Decision theories of this nature have now become central in various branches of the social sciences, and even enjoy some importance in the humanities – for instance, in versions of Rational Choice Theory utilised in both empirical sociology and history.[279] And though not enjoying huge influence amongst economists, a formulation of decision theory based on a concept of 'desirability' has been worked out rigorously by Richard Jeffrey.[280] Following Thomas Bayes, he expresses the core idea as follows:

> In the Bayesian model the agent's notions of the probabilities of the relevant circumstances and the desirabilities of the possible consequences are represented by sets of numbers combined to compute an *expected desirability* for each of the acts under consideration. The Bayesian principle for deliberation is then to *perform an act which has maximum expected*

[276] Though some economists have doubted whether we can have even introspective access to a real cardinal utility function; see Section 2.3 for further discussion of the measurability of utility.

[277] Jevons' definition of utility makes explicit reference to Bentham, and he attempts to treat economics as 'a calculus of pleasure and pain' – see Jevons, 1965 [1871], p. vi. Marshall's understanding, meanwhile, was centred on a needs-based conception of 'material well-being'; for a historical overview with further sources, see Morgan, 2009, p. 279; pp. 283-284.

[278] See Sections 2.3-2.4 below.

[279] See again D'Avray, 2010 for a critical view from a Weberian perspective. See also Lewis, 1988; Lewis, 1996 for philosophical discussion of versions of decision theory based on desires, and for more philosophical work on decision theory and substantivist conceptions of utility, see Hausman, 2011 and Bradley, 2017.

[280] Jeffrey, 1965; see also Bolker 1967 for proof of a representation theory for Jeffrey's approach, and Steele and Stefánsson, 2020, Section 3.2 for an overview of some technical details.

desirability.[281]

Though I shall make no attempt at formalisation, the account we will work with shares the same basic structure as the system Jeffrey describes here; however, following our present goals, it will be based on an agent's desires, rather than assessments of 'desirability' *per se*.[282] This brings us to the following conception of Action Selection:

1. Work out the outcomes that might result from each available action, given the constraints faced
2. Catalogue the strength of desire for each such possible outcome
3. Estimate how likely each outcome is for every action
4. Calculate the expected utility of each action as a weighted sum
5. Perform the action with the highest expected utility

In the interests of maintaining generality, we continue to leave the underlying account of 'desire' somewhat open here, assuming only that desires translate into cognitively-accessible cardinal utilities. Such, then, is the account; we now embark upon its evaluation.

The first thing to note is that despite the prevalence of this approach, an agent that was able to follow this recipe exactly in all of their choices would cut a rather strange figure. On the one hand, they would have to possess substantial intellectual and practical powers; not only able to solve often-complex mathematical problems, but also to incarnate the resulting solutions in their actions with great faithfulness. In choice problems that extend over time, for instance, the impulse to consume more than is optimal in earlier periods would be immaterial: such an agent would need to have unlimited power to delay gratification in pursuit of higher overall levels of utility.[283] On the other hand, it is clear that reason really is only a 'slave to the passions' here: though it is the proximal determinant of action, the content of its directives ultimately stem from the agent's desires, and its contribution is therefore merely

[281] Quoted in Bermúdez, 2009, p. 48.

[282] It is also worth noting that 'Desirability' features as a technical term of art in Jeffrey's presentation, as a function of propositions describing states of affairs; as such, it may differ from ordinary everyday usage.

[283] See the numerous follow-ups to the famous Stanford Marshmallow Study for the importance of this ability for many life outcomes – e.g. Mischel, Shoda and Rodiguez, 1989.

calculative, aiming solely to engineer their fulfilment. Meanwhile, the underlying preferences themselves are taken as stable, brute facts that are fixed exogenously, whose origins are of no account, and which lie entirely beyond the power of reason to modify deliberately.[284] This attitude may fit well with theorists' conception of economics as the study of 'human behavior as a relationship between *given* ends and scarce means',[285] and correlate aspirations to engage in objective scientific enquiry that is neutral on questions of value: *de gustibus non est disputandum.*[286] But it is somewhat odd that an agent with such a high level of intelligence and restraint would choose to emit such highly regimented behaviour in pursuit of satisfying *de facto* desires that may be no more than the expression of arbitrary whims.

Perhaps unsurprisingly, then, many empirical studies of what human choice behaviour is actually like have revealed radical departures from the EUT model of rational action. Much of this work was inspired by two classic experiments. In the first, Maurice Allais demonstrated that human subjects' responses to lotteries are more complex than could be accounted for by merely computing expected values: specifically, agents treat outcomes that are either impossible or occur with certainty as qualitatively different to those with numerically similarly probabilities, and in consequence violate the central 'independence' axiom of EUT.[287] And in the second, Daniel Ellsberg showed that real subjects exhibit 'ambiguity' aversion, preferring to bet on known *versus* unknown probabilities, thus causing difficulties for Subjective EUT approaches inspired by Savage.[288]

Following this early work, over the last half-century, the vast literature of *behavioural economics* has emerged.[289] This discipline aims to give

[284]Cf. Hausman, and McPherson, 2009, p. 5: 'A complete and transitive preference ranking is clearly a difficult intellectual achievement. It is the outcome of an unmodelled process of evaluation.'

[285]Robbins, 1935 [1932], my emphasis; for contrasting work on endogenous preferences, see again Bowles, 1998. "Endogenous Preferences".

[286]Stigler and Becker, 1977; for criticism of this aspiration to value-neutrality, see Sandel, 2013; especially pp. 122-125.

[287]Allais, 1953; for an early formulation of 'regret theory' intended to explain these findings, see Bell, 1982. The paradox is evaded by Kahneman and Tversky's 'prospect theory', wherein rather than using the raw probabilities when taking expectations, the agent first transforms them via an s-shaped non-linear 'probability weighting' function: see Kahneman and Tversky, 1979; and Tversky and Kahneman, 1992 for a more complex cumulative version that avoids certain technical issues with the original model.

[288]Ellsberg, 1961.

[289]See Corr and Plagnol, 2019 for an accessible introduction (though the authors appear not to have a working understanding of EUT; see p. 74). Some of the ideas occur much earlier; e.g. in Adam Smith – see here Boyd, 2020.

a more realistic view of how human beings make choices in practice, drawing on the latest insights from the social sciences, neuroscience, and especially psychology. Further violations of standard theory have thus been uncovered; for instance, in contrast to traditional assumptions about selfishness, agents often have 'social preferences' and appear to behave altruistically.[290] One notable example of social preferences occurs in an experimental set-up called the 'dictator game'. Here half of the participants (the 'proposers') are presented with a sum of money that they must decide how to share with a passive partner (the 'recipients'). Contrary to the predictions of neoclassical game theory, most proposers choose to send at least some money to their recipients rather than keeping it all for themselves, despite there being no further economic incentive to do so.[291] Following these empirical criticisms, and arguing that even such basic features of choice theory as risk aversion cannot be explained by the standard approach, two key practitioners have described EUT as an '*ex-hypothesis*.'[292]

In order to rescue EUT from these empirical criticisms, increasingly sophisticated non-standard functional forms for utility functions have been proposed. Consider, for instance, the 'ultimatum game': this is like the dictator game, where a proposer suggests a division of a fixed cash sum, but the recipient is now not merely passive and has the option to reject proposed divisions that they are unhappy with, leading to both players receiving nothing.[293] Assuming 'selfish' preferences based only on one's own monetary payoff, EUT dictates that rational responders should always accept any positive offer; proposers, meanwhile, should offer the lowest possible amount that will be accepted.[294] Yet in empirical studies, we see that real-life recipients often reject substantial offers if they are considerably less than an even split, and real-life proposers also tend to make far more equitable offers than classical assumptions would predict.[295] But economists have since responded by offering new formal

[290]See Bowles, 2004, pp. 10-11 for discussion, and see Section 1.5 above. The good of others, however, is still merely a means to the end of achieving higher levels of utility on the present account of Action Selection; cf. 1.7 above.

[291]Forsythe, Horowitz, Savin, and Sefton, 1994.

[292]Rabin and Thaler, 2001, pp. 229-230; the phrase refers to the Monty Python 'dead parrot' sketch.

[293]Forsythe, Horowitz, Savin, and Sefton, 1994.

[294]In technical terms, assuming money can be divided arbitrarily, the unique subgame perfect equilibrium is for the proposer to offer nothing and the responder to accept every offer.

[295]Güth, Schmittberger, and Schwarze, 1982; Forsythe, Horowitz, Savin, and Sefton, 1994, p. 348.

models, with more complex utility functions that now incorporate preferences for fairness over outcomes,[296] aversion to inequity (whether the discrepancy in monetary payoff is advantageous or disadvantageous),[297] and even intentions, to which choices are reported to be sensitive.[298]

Whilst neatly accommodating these particular experimental findings, however, a potential problem with this approach arises here. For even if EUT is intended only as a descriptive theory, then as the choice sets upon which these non-standard utility functions are defined are allowed to become more and more fine-grained, our models lose predictive and explanatory power. At the limit, we reach what some decision theorists have called 'comprehensive outcomes', wherein by attending to *all* features that the agent might plausibly care about – historical, social, and personal – any individual outcome is treated as an entirely unique entity, and preferences are once again simply a record of their actual choices, however heterogeneous these might be.[299] This is clearly not in the spirit of the original economic theory, and it has also been objected that EUT then becomes unfalsifiable.[300] Moreover, from our normative perspective, we are returned to the same dead end we encountered earlier: since utility again becomes merely a reiteration of the agent's actual *de facto* choices, it cannot serve as an independent standard that guides these same choices.

This reminder of the purposes of the present enquiry does, however, point to a second general response to these empirical criticisms of EUT. For although many economists aim to predict or explain actual human behaviour, our present focus is quite different: to offer guidelines that would characterise the behaviour of ideally rational agents. Thus, it might be argued that what is highlighted by such empirical studies is only a human failing in determining or carrying out the optimal choice, and this does not in itself impugn the normative results derived: 'Expected utility theory makes faulty predictions about people's decisions in many real-life choice situations ... however, this does not settle whether people *should* make decisions on the basis of expected utility considerations' –

[296]Rabin, 1999; Blount, 1995.
[297]Fehr and Schmidt, 1999; and Bolton and Ochenfels, 2000.
[298]Falk and Fischbacher, 2006.
[299]Steele, 2014, p. 199.
[300]Steele responds by distinguishing between the EUT framework as a whole – which is indeed unfalsifiable – and more or less fine-grained specific utility functions that are fitted to agents in particular contexts – which are not; she then urges the pragmatic nature of choosing an appropriate level of detail for such particular models: Steele, 2014, p. 201.

either because maximising the good is an independently plausible principle of rationality, or because of one of the representation theorems cited above.[301] The surprising revelations of behavioural economics would then be interpreted as merely error theories, growing up as they have done alongside the psychology literature on cognitive bias.

Although the response is to the point, and for us merely reporting observed behaviour inconsistent with EUT is indeed not enough to undermine its credentials, such empirical studies may yet cause problems of another sort even in the present normative context. For in this section we are specifically seeking an account of practical reason based on the acting agent's desires; but it may be that these desires themselves have certain content or exhibit a certain structure that prevents them from being expressed within the core EUT framework, so that EUT would again fail to be an adequate vehicle for mobilising our Subjectivist accounts of the Human Good at the level of Action Selection. To see how this might be so, we must turn our attention back to our formal maximising conception of rationality and the underlying axioms that characterise it.

So far, we have not explicitly imposed any restrictions on the content of the desires at the heart of the account of practical reasoning under consideration. And we also have seen that the underlying formal mathematical framework of EUT more generally has little to say about how the central notion of utility is to be interpreted.[302] Indeed, it is sometimes claimed that EUT can be applied in an entirely objective way that avoids making value judgements; that it is able to 'stand ready to serve any set of values which is otherwise (non-rationally) decided upon'.[303] As we have said, it therefore seems to fit well with some contemporary conceptions of what economists are supposed to do; for instance, those working as government advisors, whose official role is to present their findings as evaluatively neutral, objective facts to be used by policymakers.[304] According to this view, economics 'simply doesn't traffic in morality'.[305] However, despite this aspiration to neutrality, I shall

[301] Briggs, 2019, Introduction [my emphasis]. Savage's response to the Allais paradox was thus that the preferences induced in the experiment are irrational; see Savage, 1954, pp. 102-114. For criticism of behavioural economics along these lines, see Maialeh, 2019; on a recent 'normative turn' in economics, see Hands, 2011.

[302] See again Broome, 1999.

[303] Dunne, 1993, p. 188 (discussing Habermas); cf. *ibid.*, p. 277 for further criticism from an Aristotelian perspective.

[304] See again Robbins, 1935 [1932] for the view that economics is concerned only with means; see also Weber, 2012 [1904].

[305] Levitt and Dubner, 2006, p. 190.

claim that applying the EUT framework in the present context actually does entail that we implicitly impose certain substantive assumptions about agents' values.[306]

One such commitment results from the mathematical form taken by EUT.[307] Choice theorists are fond of saying that their definition of rationality is extremely minimal, requiring only very limited restrictions that ensure agents' choices are formally consistent with each other. However, if we are to have a practically applicable and genuinely Action-Guiding Subjectivist version of EUT, then the value of the outcomes that may result from an agent's actions must be representable in terms of cardinal, numerical utility, so that the mathematical operation of taking expectations can be performed. Moreover, the real power of the approach as it is deployed by economists in practice only appears when the outcomes of interest are themselves capable of numerical representation, as vector-valued bundles of goods. For only then can general features of the agent – such as risk-aversion – be investigated systematically. Indeed, in practice, actual economic analyses of particular situations usually do not begin with underlying preferences, but rather by equipping agents with a real-valued utility function that takes such vectors as inputs. If outcomes are represented merely qualitatively, in contrast, then utility scores must be assigned to them on an individual, piecemeal basis, and this soon becomes unworkable.

This mathematisation of the objects of value in effect imposes the assumption that everything an agent might care about can be both expressed and valued in quantitative terms.[308] It therefore fits well with contemporary Subjectivist views, wherein well-being is seen as a measurable commodity.[309] However, it also entails that EUT has difficulty handling intangible goods such as justice, fairness, autonomy,

[306] Mandler, 2001, p. 374. See Dunne, 1993, p. 7 for similar claims regarding 'objective-driven' instrumental theories of rationality in the context of education and beyond; *ibid.*, pp. 187-188 for discussion of the partisan nature of this 'disinfected', technical, instrumental reason, despite its 'seeming coincidence with the structure of rationality itself'; MacIntyre, 2011 [2007], pp. 88-89 for similar criticism of managerial effectiveness as a morally neutral value; and also my Berry, 2019b for a criticism of rationality given in terms of fulfilment of individual goals alone in the context of mathematical practice.

[307] For a historical viewpoint on the use of mathematics in economics, see Porter, 1994.

[308] See Islam, 2022 for discussion of the consequences and limitations of quantification in the social sciences; especially p. 198 on '*commodification*', or the 'quantification of life'; on Marshall's early warning about the mathematisation of economics, see Brue and Grant, 2012, p. 294; cf. Marshall, 1920, p. xxiii.

[309] Cf. Section 1.6 above.

freedom, and dignity, which cannot be modelled in this way without an unacceptable level of distortion.[310] Yet behavioural economics has shown that such goods frequently are objects of actual human desire.[311] Moreover, the prospects of representing the expression of virtues such as honesty, justice, and courage in numerical terms are also slim indeed, since these traits relate directly to action and are not in any way measurable commodities. And of course, the assignment of numerical utility only to individual agents entirely precludes recognition of those essentially shared and thus genuinely common goods that will become central in the next chapter.[312]

A further consequence of the mathematical formulation of EUT is that it entails that any two objects of value can be traded off against any other; that is, it implicitly treats all goods as commensurable (from Latin: *com* – 'together' and *mensurabilis*, from *mensurare* – 'to measure').[313] Hence, EUT cannot make room for the genuinely unconditional commitments and obligations of the kind that moral philosophers have often insisted that a real concern with fairness or justice would require, and which are also typically thought central to certain kinds of important human relationships.[314] Indeed, its restrictive nature is made plain here by the fact that for this reason some major philosophical views cannot be represented within EUT. For instance, as we have seen, on a traditional Stoic view, attending to virtue is of the first importance, and this overrides anything else the agent might strive for – though once virtue is secured, other goods (or 'preferred indifferents') may be pursued too. Such preferences are known as 'lexicographic' in the economics literature – after the way a dictionary is organised, whereby the first letter of a word is of overriding importance in determining the order of entries. However, lexicographic preferences violate standard assumptions on preferences, and in general *cannot* be represented by a real-valued utility function.[315]

[310]Sandel, 2013, p. 138 for criticism of economists' typical treatment of these notions: 'Are altruism, generosity, solidarity, and civic spirit like commodities that are depleted with use? Or are they more like muscles that develop and grow stronger with exercise?'

[311]See, for instance, the discussion of the 'ultimatum game' game above for empirical evidence on agents intrinsically caring about intangible goods, including fairness.

[312]Thus, in the context of policy evaluation, when one does often take a 'social planner's' perspective by aiming to maximise a 'social welfare function', this is always computed from agents' individual utilities.

[313]On decision-making in the face of apparently incommensurable sources of value, see Levi, 1986.

[314]Cf. MacIntyre, 2013, pp. 28-29.

[315]Specifically, such preferences violate the continuity axiom, so even the ordinal representation theorem cited above does not apply. If the choice set is finite, however,

Another deep commitment of EUT is that it is also inherently consequentialist, in that when evaluating actions only the resulting outcomes are taken to be possible objects of value; the nature of the agent's own contribution is therefore typically taken to be of no special import. The agent's role might in principle be included by enriching the choice set, explicitly stipulating their actions to be part of the 'outcome' in question.[316] Yet this move is not only rather artificial, going against the spirit of the theory, but also has inherent limitations. We can see this by comparison with the Aristotelian tradition, or with the many other theories in ancient ethics wherein virtuous action itself is the focus of *eudaimonia*, since neither acting consistently with the virtues nor acquiring the virtues as character traits can plausibly be described as a mere 'outcome' of a choice in the required sense; one reason for this being that the *manner in which* one chooses makes an important difference here. Conversely, since ancient ethics does not only aim at outcomes or the production of states of affairs, it has rarely come close to a maximising model of rationality.[317]

Finally, standard supplementary axioms on utility functions also embody normative assumptions that seem natural only within a quite specific cultural context. Consider in particular the axiom of monotonicity, which says that points of higher utility are always reachable by having increased quantities of any of the goods in question, and which on our substantive approach means that agents must always desire more and more of any good. Whilst not strictly essential to EUT – that is, not required for a minimal mathematical formulation – in practice, this feature of utility functions is employed in nearly all EUT models, and is a key condition in the central First Fundamental Theorem of Welfare Economics.[318] For some ancient Greek philosophers, however, this feature of an agent's desires would be seen not as a general and uncontroversial feature of rationality itself, but instead as pathological: a symptom of the human vice of *pleonexia,* or acquisitiveness for its own

---

we can find a utility function representation.

[316]See again Steele, 2014, p. 199.

[317]Annas, 1993, p. 38; see also Annas, 1993, p. 30: 'I want there to be a certain outcome, but I want something further as well: I want to *bring about* that outcome, not just to have it happen.'

[318]Though some statements use only the weaker assumption of 'local non-satiation', which says that points of higher utility are always available nearby. One type of exception occurs in models where central banks are assigned utility functions based on ideal levels for variables such as inflation and output.

sake.[319]

What attending to the underlying axioms thus reveals is that the technical assumptions required to translate a desire-based account of the Human Good into an Action-Guiding version of EUT are not neutral, uncontentious principles of instrumental rationality, since they indeed serve to preclude some major ethical standpoints from consideration entirely. Moreover, from the point of view of a desire-based theory of Action Selection, the empirical criticisms of EUT uncovered by behavioural economics cannot be explained away as the mere failure of agents to act rationally, since some of these show that our simple 'strength of desire' model that assigns numerical values to outcomes cannot do justice to the complex nature of actual human desires themselves.

In the next section, we shall see some further features of human desires which entail that fitting them into a rigid EUT framework must do violence to their actual structure. Specifically, here we shall explore how our desires may radically conflict with each other in a way that mere numerical weighting cannot adequately account for, and how our desires may change over developmental time in a way that we may choose to guide explicitly. This will also bring us to the complementary context of Life Planning, on which EUT has thus far said nothing. In summary, although the neo-classical microeconomic approach has been highly successful at modelling the choices over readily-quantified consumption goods that economists have traditionally been concerned with, the belief that all human desires can be faithfully rendered into an EUT framework represents a highly limited and implausible view of the human condition.

Our initial hopes for a purely technical solution to our core problem have thus for the moment stalled. In sections 2.3-2.4, we will aim to remedy this situation by considering the prospects of a Subjectivist, Action-Guiding version of EUT based on an understanding of utility as felt pleasure instead of desire. For it has been argued by some recent theorists that pleasure is measurable, and if this is so then a return to a Benthamite conception of utility might make available the sort of algorithmic account of practical reason that was sought in vain here. In the next section, we temporarily set aside the formal mathematical framework of decision theory, and consider Subjectivist desire-based accounts of practical reason more broadly.

[319] Mill's cultural assumptions led him to mistranslate this term to mean 'taking what belongs to another', whereas Nietzsche insightfully translates it as *'haben und mehrwollhaben'* – see MacIntyre, 2011 [2007], pp. 159-160.

2.2 Desire-Based Accounts In General

In the previous section, we raised some limitations of EUT as a means of mobilising desire-based conceptions of the Human Good. However, these objections do not apply to every kind of desire-based account of Practical Reason. For instance, in general a desire-based view might recognise and advise us to act upon desires specifically for such intangible goods as justice, commitments to unconditional relationships, or activity in accordance with the virtues, whilst not aspiring to represent the strength of these desires or their corresponding objects numerically; we need not aim for such precision here in order for the account to be practically useful.[320] Indeed, it is a dialectical strength of this type of view that its adherents seem to always be able to offer a desire-based interpretation for our pursuing anything we might have regard for: if such items are something we pursue, such a response would go, there must be a sense in which we 'want' them, and the desire-based theorist can argue that this is what really underlies our choice. Moreover, a similar interpretation can be given for any alleged normative constraint on our desires that we care to name.[321] Russell expresses the point well:

> All human activity is prompted by desire. There is a wholly fallacious theory advanced by some earnest moralists to the effect that it is possible to resist desire in the interests of duty and moral principle. I say this is fallacious, not because no man ever acts from a sense of duty, but because duty has no hold on him unless he desires to be dutiful.[322]

We therefore now turn to desire-based accounts of practical reason in general. Often called 'neo-Humean', contemporary philosophers who have worked out such theories include Donald Davidson, David Lewis, and David Gauthier.[323] We also continue to allow 'desires' in any sense of

[320]Cf. Section 1.2.5 above. For MacIntyre, it is the desire to act for the sake of the good *qua* good that is key; see MacIntyre, 2009 [1993], p. 70; p 87; and MacIntyre, 2016, p. 34; p. 37.

[321]For discussion of the general strategy, see Sen, 1977, pp. 322-323; see also Talbot, 2021 on 'Consequentialising' apparently deontological principles.

[322]Russell, 1999 [1950], p. 452; for the same view, see Kant, *Groundwork*, 4:460; and MacIntyre, 2009b [1999], p. 70.

[323]See Lewis, 1998; 1996 for defence of the Humean view that we are motivated only by our desires, based on a prior acceptance of decision theory; Davidson, 1963 on practical reasons as consisting of an internal pro-attitude towards a certain type of action together with a belief that the action in question is of this type, and arguing that reasons in this sense have a causal force; and Gauthier, 1987 on dispositions

the term to contribute to our deliberations here; in particular, the desires in question can in principle be *for* anything – including 'pro-social' desires relating directly to the well-being of others, and indeed even the desire to do good as such. But nevertheless, on this view, any normative practical reason for action that an agent might have must always be an 'internal' reason that is reflected in a corresponding desire, the satisfaction of which remains the only Intrinsic Good they can recognise.[324] In particular, we should act in the interests of others only insofar as we continue to have a desire to do so – a point which will become crucial in the following chapter.

In exchange for developing this broader and more flexible desire-based idiom for practical deliberation, however, we must also relinquish the powerful formal framework of choice provided by contemporary decision theory. We are then soon faced with the daunting task of how to weigh up competing desires. This becomes especially pressing when we realise that our existing desires may be radically inconsistent with one another; moreover, as mentioned in the previous section, these conflicts may have a more complex structure than the simple picture suggested by the version of EUT developed there can account for.[325] For instance, consider again a dietary choice; whether to order a burger and chips for lunch, or else a salad. An individual may find that in terms of these particular items themselves, the strength of felt desire for the first option is far greater; yet at the same time, they may also harbour a longer-term desire to lose weight, or else to improve their overall health. The conflict between this 'standing' desire and the desire for the burger and chips is unlike the conflict between the latter and a desire for the salad, since it does not attach to the outcome of this particular choice in a straightforward way, but rather belongs to a different sphere of deliberation entirely. When such discrepancies occur, then, the question arises: *which* of our desires should we listen to in deciding what to do?[326]

In exploring this question, we return to the insights afforded by the

---

to moral action as justified by their leading to the more effective satisfaction of our desires. See again MacIntyre, 1982 for historical discussion of the Humean conception of practical reason and its influence; MacIntyre, 1987 for further elaboration; and MacIntyre, 2009b [1999], p. 115 on Gauthier's contractarian view.

[324] In the sense of Section 1.7 above; cf. again Williams, 1979.

[325] For discussion of conflicting desires, see MacIntyre, 2016, p. 133; MacIntyre, 2011 [2007], p. 56; MacIntyre, 2008; and Rawls, 1971, pp. 554-560; see also Frankfurt, 1971 on first- and second-order desires, discussed below.

[326] Cf. Bermúdez, 2009, p. 61 for argument that resolving such complex conflicts of desire cannot just be a matter of formal logic; see also MacIntyre, 1993, p. 9-10.

Appraisive philosophy of well-being. Here we will see that in contrast to the permissiveness of EUT, most theorists have opted to place substantive limitations on the possible objects and nature of the desires whose satisfaction is good for us.[327] Once translated into an Action-Guiding context, these restrictions may then help us to resolve the problem of conflicting desires by directing us to satisfy some desires rather than others: that is, those desires which philosophical theory endorses as relevant to our well-being.[328]

In the *Meno*, Socrates contends that no one can desire what is bad for them. The argument runs roughly as follows: to desire bad things is to be miserable; no one desires to be miserable; and therefore no one desires bad things.[329] Yet from a modern perspective, the argument seems clearly fallacious, since one can desire the same entity under one intentional description but not another. And indeed, we know from experience – in advance of any such philosophical argument – that there *are* individuals who harbour desires for things that are bad for them, even according to their own standards: *pace* Socrates' second premise, an agent may even specifically desire their own ill-being – if they believe they deserve to be punished, say.[330] Moreover, as Plato himself recognises in this dialogue, we may form defective desires on the basis of false information.[331] Consequently, there is a broad consensus amongst contemporary analytic philosophers of well-being that the fulfilment of mere *de facto* desires is not an adequate basis for an Appraisive account of the good life, and some restrictions are needed.[332] We therefore now consider some of these philosophers' attempts to draw the required boundary around those desires whose fulfilment contributes to well-being.

Early on in the modern tradition, Sidgwick considers an account of the 'ultimate good on the whole for me' in terms of 'what I should practically desire if my desires were in harmony with reason',[333] and Rawls likewise

[327] Cf. Section 1.2.1 above.

[328] One issue with restricting or idealising desires in this way, however, is that it tends to undercut what is perhaps the primary appeal of desire-based Appraisive views within our individualist culture: that the evaluations are based on a standard that is at least implicitly endorsed by the agent themselves. For fuller discussions of further philosophical objections to the desire-based view as an Appraisive account of well-being, see Sumner, 1996, Chapter 5; Griffin, 1986, Chapter 2; and Crisp, 2021.

[329] Plato, *Meno*, 77c.

[330] Cf. Kraut, 2007, p. 103.

[331] See here the discussion of 'real interests' in Williams, 2006 [1985], p. 40.

[332] Though for accounts that aim to block these idealising moves and defend the resulting 'actualist' view against objections, see Heathwood, 2005 and Lemiare, 2020.

[333] Sidgwick, 1981 [1907], p. 112; as with Mill, the relation to his official hedonism

discusses what he describes as a utilitarian view given by 'the satisfaction of rational desire'.[334] Parfit, on the other hand, introduces another kind of constraint: this time on the objects of our desires. He distinguishes between unrestricted and success views. According to the unrestricted view, the fulfilment of *any* desire contributes to our well-being. He offers an example: suppose we meet a stranger who has what we believe to be a fatal disease. We desire for him to get better. Much later, unbeknownst to us, he does. Parfit finds it absurd that this later development should impact our well-being, and therefore takes it to be obvious that we should reject the unrestricted view. He then outlines the 'success' view, which he regards as more plausible: according to this position, only desires *about our own lives* are relevant to the evaluation.[335] And in recent years, a great number of further proposals for limiting the scope of desire have been offered: we now have Appraisive theories of well-being based on comprehensive desires, informed desires, second-order desires, real desires, true desires, existence-entailing desires, and desires that would survive a rigorous process of psychotherapy.

Returning now to our Action-Guiding enquiry; we have said that these proposed restrictions may allow us to resolve the conflicts alluded to, insofar as such an auxiliary criterion might furnish us with at least a partial theory of which of our possibly multifarious and often conflicting desires to follow and which to ignore. However, we have just seen that there are a large number of diverse views about which desires are relevant, each of which has a certain *prima facie* plausibility. So, we must now ask: Which such restrictions on the content of desires should we accept as binding? Yet in the face of so many rival accounts, one possibility that must be considered is that due to the complexity of human psychology and the range of desires we can have, any rigid *a priori* categorisation imposed by philosophical theory will always prove inadequate to the task. If such a canonical formal demarcation is not available – as the lack of consensus in the literature here seems to suggest – then perhaps agents must instead evaluate which desires to follow on a piecemeal basis, and we must therefore allocate a substantive role to the agent's own judgement in particular situations.

---

is again complex.

[334] Rawls, 1971, p. 30.

[335] Parfit, 1984, p. 494; He later advocates a further refinement to a 'global' version of this view, based only on 'someone's desires about some part of his life, considered as a whole, or about his whole life.' See also *ibid.*, p. 500 for a discussion of an 'ideal conditions theory' for when choices reveal preferences.

This brings us to a key overall conclusion of this chapter, one often emphasised by Aristotle: that both in concrete situations and in organising our lives as a whole, there is inevitably a need for agents to issue context-sensitive practical judgements that cannot be determined *a priori* at the level of explicit theory:[336]

> Practical intelligence is indispensable, because there are no algorithms for making the kind of judgments with which it is concerned; there are no algorithms because practical creativity is a matter of discriminating among the various forms of guidance to be derived from the past in a way for which there is no universal recipe.[337]

The question then arises of whether the desire-based view has the resources for providing normative standards that these context-sensitive evaluations and judgements can appeal to as we determine which of our diverse and rival desires to pursue. For if it cannot, it would appear that practical rationality must operate independently of desire here, with agents coming to recognise specific desires as more or less worthy of endorsement for reasons that are independent of their extant desires themselves. Yet to accept this position, wherein our preferences are evaluated according to external criteria, would be to depart from the standpoint of a neo-Humean account entirely.[338]

Setting this issue to one side for the moment, we now raise a second complication which we will have to accommodate: human agents in the real world are unlike those of mainstream economic theory in that their preferences change over time.[339]

In contrast to ancient ethics, which takes as a key theme the moral growth of the agent and the training and education of their desires over the long run, we have seen that a core aspect of EUT is its restriction to the context of Action Selection. In mainstream economics, then, this

---

[336]Cf. Dunne, 1993, p. 81 on the limits of technical rationality; see also *ibid.*, pp. 272-274 and Section 9.4 on the limitations of theoretical knowledge about universals for moral practice, and Aristotle's writings for the view that the most appropriate response is 'determined by reason [*logos*] and in the way in which the man of practical wisdom [*phronimos*] would determine it' – *NE*, II.6, 1107a1-2; quoted in *ibid.*, p. 312.

[337]MacIntyre, 1993 [1988], p. 7; see also MacIntyre, 2011 [2007], pp. 178-180; Nussbaum, 2001 for another account of practical reason emphasising particularism and attention to social context, and Collingwood, 2013, pp. 101-106 on the limitations of rule-guided behaviour.

[338]See here MacIntyre, 2008, p. 285.

[339]For a thorough discussion of the issue from a philosophical perspective, see Parfit, 1984, Ch. 8; see also Griffin, 1986, p. 16.

process of development is thereby left out of account. It is true that economists do sometimes consider problems that appear to extend beyond a here-and-now timescale; most obviously, within 'dynamic' choice problems, such as models of inter-temporal consumption. However, in the standard Samuelson model, the agent's 'instantaneous utility function' U is held fixed across each period; the impact of utility from subsequent periods is then reduced by multiplying by a subjective discount factor $\beta < 1$, representing impatience. Agents therefore aim to maximise the following quantity:

$$\sum_{t=0}^{\infty} \beta^t U(C_t)$$

Where C_t is consumption at time t. In the absence of uncertainty about future constraints such as income, such agents will never change their minds about what they intend to consume in future periods, and hence may as well be thought of as making a single choice about their future consumption path as the first period commences.[340] Similarly, although game theorists frequently discuss 'extended games' that also continue across more than one discrete time period, here agents are thought of as selecting a 'strategy' – a complete and contingent plan of action – *before* the game begins.[341] Indeed, as explained in the discussion of consistency in the last section, and *contra* those ancient philosophers who place great value on character development, agents who instead undergo 'preference reversals' over time are typically treated as irrational in the mainstream economics literature.[342]

Despite operationalist economists' insistence on consistency over time as a defining feature of rationality, however, in practice our desires do undergo changes even if we do nothing, and a desire-based theory of practical reason cannot ignore these changes if it is to be Subjectivist in the sense at issue in this enquiry.[343] Moreover, given that this is what human

[340] See Bowles, 2004, p. 8: 'The passage of time is represented simply by a discount rate; people do not learn or acquire new preferences over time'.

[341] Though see again Bowles, 2004 for an evolutionary approach that models time explicitly; see also Andreou, 2020 for discussion of some recent innovations.

[342] See e.g. Green, Fristoe and Myerson, 1994 for an example of preference reversal in the context of temporal discounting that challenges the empirical adequacy for real-world agents of the Samuelson view, which Samuelson himself raised some skeptical points about. On a technical level, economists have distinguished 'stationary', 'time-consistent', and 'time-invariant' preferences; see Halevy, 2015 for details.

[343] Cf. the previous section.

life is actually like, the *laissez-faire* attitude of passively accommodating and reacting to such changes as they happen to emerge is just one possible response we could choose here, as opposed to actively shaping our desires over time – and hence is a response that stands in needs of rational justification. For instance, in the intertemporal consumption problem just described; what if the agent was able to intentionally increase their subjective temporal discount factor β, investing energy in doing work on themselves over time in order to become more patient? Similarly, when our desires conflict in the here-and-now, the passive attitude of retaining all of these desires and allowing them to interact according to their own natural dynamics, based on their existing motivational strength and implicitly assumed importance, is again just one possible response here, and again one that must be chosen for an explicit reason if we are to be rational in this area of our lives.[344] Hence, any adequate Action-Guiding theory must provide some account of how we are to manage these conflicts and control this development. What the desire theorist must supply us with is both a way of rationally resolving individual conflicts of desire in the here-and-now, and a way for us to rationally guide the evolution of our desires over time. Moreover, this proposal must also be consistent with the Subjectivist epistemology wherein our it is our desires that inform us about the good.[345] What such standard of evaluation can they offer?

In responding to these challenges, a desire theorist may appeal to the complementary perspective of Life Planning. For a Life Plan may indeed include such internal elements as what our desires should be like, and from this longer-term Life Planning perspective our desires can be explicitly moulded by our own choices.[346] Through appealing to a concrete Life Plan, we may therefore be able to judge competing or incompatible desires as worthy or unworthy of endorsement, and to guide the development of our desires over time. This brings us to a more complete desire-based conception of practical reason. First, we determine what kind of life we are to aim for; then, we strive here-and-now to fulfil – and over developmental time, to nurture or form – only those desires that will bring this core vision to fruition. The next question at issue is then

[344]On 'warring' desires, see MacIntyre, 201, p. 37; on ordering our desires, see Rosati, 2006, pp. 33-36 and MacIntyre, 2016, p. 47.

[345]The strategy of achieving satisfaction by relinquishing our desires is therefore out of place here; cf. Rieff, 2008, p. 96: 'I doubt that Western men can be persuaded again to the Greek opinion that the secret of happiness is to have as few needs as possible'; for discussion of this strategy within Confucian thought, see Machek, 2021.

[346]Cf. Section 1.4 above.

whether a purely desire-based account is able to offer the resources we need to guide us in developing such a Life Plan, or whether substantive principles that go beyond and are independent of our desires must instead be brought in at this stage.

The natural way for this to work would of course be to allow our desires themselves to be the standard of choice between competing conceptions of how best to shape our lives as a whole, so that our Life Plan is itself a reflection of the kind of life we most desire.[347] But again, issues may arise if our desires about the kind of life we want are themselves multifarious or liable to conflict, thus specifying modes of living that are mutually inconsistent: which of them should we then listen to? For instance, we may recognise that given our current abilities, we cannot be professional tennis stars whilst also staying up late and drinking heavily with our friends every weekend, though we may have a strong desire to live each type of life.[348] As occurred within the context of Action Selection, attending to each and every desire relating to our lives as a whole may yield no single, coherent conception of what our Life Plan should be like. And of course, our desire for a certain kind of life may be subject to change over time – again, perhaps due to our explicit influence. Yet it also remains possible that there is a pre-existing hierarchical structure to our desires about what kind of life we want; some order within this chaos which may supply what we need here. In the context of Life Planning, perhaps we could then follow Bernard Williams' paraphrase of D. H. Lawrence's imperative to 'find your deepest impulse and follow that'; and similar ideas are also to be found in Nietzsche, for whom we must 'become who we are'.[349]

If these more fundamental desires about the kind of life we want are stable and coherent enough to pick out a particular way of life, then, they might thus provide us with the deliberative standard we are seeking, as we strive to live the kind of life and become the kind of person the resulting Life Plan prescribes. And since practical reason would ultimately be guided only by these other, more fundamental desires here, it would not need to operate outside of the scope of desire as such.

Such a view would find similarities with the conception of the relation between practical reason and desire found in Frankfurt, who attends

[347] Cf. again Parfit's view that desires about our lives as a whole should be made central to assessing well-being.

[348] See again Rosati, 2006; unfortunately, given *my* abilities, I can do neither.

[349] On Williams and Lawrence, see MacIntyre, 2016, p. 68; p. 150; p. 155.

carefully to the psychological structure of desire and famously introduces the concept of second-order desires: that is, desires relating directly to and evaluative of our given first-order desires.[350] So, on his account, we do indeed on occasion explicitly evaluate and sometimes discard our first-order desires in our pursuit of the Human Good as we conceive it.[351] However, for Frankfurt these evaluations are merely expressions of our further Subjective commitments, since 'it is only in virtue of what we actually care about that anything is important to us',[352] and there are in the end no principles to guide us other than the satisfaction of those Subjective Values with which we most deeply identify.[353] Frankfurt elaborates: 'There can be no rationally warranted criteria for establishing anything as inherently important ... our final ends are provided and legitimated by love ... The standards of volitional rationality and of practical reason are grounded only in ourselves ... only in what we cannot help caring about and cannot help considering important'.[354] Thus, reasons themselves are always based on pre-rational criteria, and it follows that 'Individuals will differ in the outcomes of their practical reasoning just as they do and because they do in their affective commitments'.[355]

One strategy for developing a critique of this more sophisticated desire-based view would be to double down on the attempt to separate reason from desire; to argue that – *pace* the expressivist account – in discharging its deliberative role here, rational thought must operate entirely outside of its purview. Yet although a key claim of this chapter is that there is a need for substantive judgement on the part of the agent as they deliberate about which of their Subjective Values they should endorse, the route to arguing for this stronger claim about the independence of this process from our desires themselves is a difficult one, reaching to the deepest issues in psychology and the philosophy of mind. Moreover, recent empirical and philosophical work has suggested that reason cannot in some sense be cleanly 'detached' from the emotions,

[350] Frankfurt, 1971; see also Sen, 1977, pp. 335-341 for a discussion of the extent to which EUT can accommodate second-order preferences.

[351] Cf. MacIntyre's view that rationality requires distancing ourselves from our desires – e.g., MacIntyre, 2009b [1999], p. 83.

[352] Frankfurt, 2006, p. 20.

[353] For a discussion of Frankfurt as an expressivist in this sense, see MacIntyre, 2016, pp. 44-48; see also Hume, 2000 [1739-1740], pp. 265-268 (*T* 2.3.3) for a classic statement of the view that reason cannot offer criticism of intrinsic preferences; again MacIntyre, 1982 for discussion of Hume's view; and Hansson, Sven Ove and Till Grüne-Yanoff, 2022, Section 8 on 'preference criticism'.

[354] Frankfurt, 2006, p. 22; p. 26; p. 33.

[355] MacIntyre, 2016, p. 45; cf. MacIntyre, 2009b [1999], p. 114.

with the traditional view of reason as an independent, stand-alone faculty being labelled 'Descartes' Error' by the influential neuroscientist Antonio Damasio.[356] Demarcating these aspects of the mind in the required way is therefore bound to be highly contentious, and this route will not be pursued further here.[357]

A second line of criticism would be to again point to recent work in behavioural economics – this time, in order to cast doubt on the proposed underlying stability being appealed to. For many practitioners of this discipline are highly skeptical about human beings having determinate preferences that are stable over extended periods of time, and some go as far as denying that this is true even at an individual instant; on some contemporary views, preferences are now instead thought of as being constructed in the moment, in response to completing specific elicitation tasks.[358] This research thus criticises an earlier wave of behavioural economics as based on a misguided view of a neo-classical 'Inner Rational Agent' that possesses definite preferences and whose attempts to make rational choices that fulfil these are thwarted only by impulsiveness and bounded rationality – of human beings as 'faulty Econs ... trapped inside and constrained by an outer psychological shell'.[359] Briefly, one core source of evidence appealed to here is studies showing that judgements and choices are extremely sensitive to framing effects, so that switching between different elicitation methods can therefore induce such 'preference reversals' as mutually inconsistent choices between policies. Two key practitioners conclude:

> the notion of preference that underlies modern decision theory is more problematic than economists normally assume because different methods of elicitation often give rise to systematically different orderings ... people do not possess a set of pre-defined preferences for every contingency. Rather, preferences are constructed in the process of making a choice

[356] Damasio, 2008; see also Dennett, 2017, p. 94-95.

[357] 'Preferences' are also sometimes understood as considered attitudes about what is most choiceworthy rather than felt desires; see also Bowles, 2004, p. 99 for an approach to preferences as 'reasons' understood in a very broad way, and see Sen, 1977, p. 329 for discussion of the issue. See again Lewis, 1988; 1996 for a converse argument moving from the premisses of decision theory to an attack on 'anti-Humean' positions that claim we are moved to action by our beliefs about the good rather than desires.

[358] See the recent 'JDM' (judgement and decision-making) literature – e.g. Lerner *et al.*, 2015, especially p. 814-817; and Stewart, 2009.

[359] Sugden, 2018, p. 82.

> or judgment ... the context and procedures involved in making choices or judgments influence the preferences that are implied by the elicited responses ... The data ... suggests that no optimization principles of any sort lie behind the simplest of human choices ... The research reviewed in this article is most compatible with the ... view of preference as a constructive, context-dependent process.[360]

This radically constructivist view of preferences in fact has a rather longer history within economics,[361] and the effects emphasised by behavioural economists here have also long been understood and exploited by those industries who profit from the deliberate inculcation of artificial desires that far outstrip need, and who have on occasion also attempted to deliberately promote desire-based conceptions of the Human Good amongst the general public.[362]

What these ideas suggest is that the coherent underlying regularity of a deep self that is requisite to the Williams/Lawrence/Nietzsche suggestion – 'an immutable, distinctive essence'[363] special to us *qua* individuals, existing largely independently of our social relationships, and enduring throughout our growth and moral development, may turn out in the end to be only a myth; the proposed individualistic approach to Life Planning in terms of merely 'working out what you *really* want' would then prove to be inadequate in important respects.[364]

The premisses of this second line of argument will be implicitly supported by what I have to say about development in the next chapter, which will present empirical research from developing psychology suggesting that prior to a process of socialisation and education into what to desire and how to regulate our desires, such a guiding 'deepest impulse' does not yet exist.[365] However, rather than pursuing this particular

[360]Tversky and Thaler, 1990, p. 202; p. 209; p. 210.

[361]For a classic early suggestion of this type of constructivism in the context of social comparison effects, see Galbraith, 1958.

[362]For a contemporary overview of the advertising industry from the standpoint of behavioural economics, see Corr and Plagnol, 2019, Chapter 7; at the start of the 20th Century, a disgraced Watson – a key founder of behaviourism – subsequently sought a career in advertising.

[363]Sapolsky, 2017, p. 225.

[364]See Sapolsky, 2017, p. 222 on the biology of environmental influences on behaviour, and especially the lasting influence of childhood experiences; see also here Dewey, 2002 [1992], pp. 84-85 on the social formation of individuality.

[365]See Bowlby, 1997 [1969], p. 349 on parental influence: "the behavioural organisation of a child of [twelve months] is much less stable than is that of the couple of which he is a partner"; *ibid.*, pp. 364-366 on initially labile attachment

strategy further, in this final chapter the argument will follow a third route; one which will be much easier going.[366] We therefore set this interesting and suggestive debate aside, carrying forward only the more modest conclusion that even *if* a purely desire-based conception of practical reason at either level proves viable, then any agent employing it would still have rather a lot of deliberative and interpretive work to do in ordering and choosing between the rival claims of their various desires – rather than being able to merely follow whichever of their desires over outcomes are strongest, as in the simple, algorithmic version of EUT given in the previous section.[367]

In successfully deploying a desire-based Action-Guiding account of any sort, then, human beings require powers of reason and judgement that are quite extensive. In the next section, we consider a different attempt at a technical solution to our core problem: an EUT based on our second Subjectivist theory, which understands what is good for us in terms of pleasure and pain.

2.3 Pleasure and Measurement

The second Subjectivist account we discuss is hedonism – from the ancient Greek *ἡδονή*, for pleasure.[368] Hedonism is a doctrine with a long and varied history, and comes in several forms. *Psychological hedonism* is a view concerning the causes of behaviour: that only pleasure, or pleasure together with pain, has the ultimate capacity to motivate.[369] In contrast, *evaluative hedonism* – the primary incarnation in contemporary philosophy, within the Appraisive study of well-being – provides an explanation of why some lives are better than others: that is, because

styles coalescing into a distinct personality style in the first year of life; and *ibid.*, p. 378 for a summary of attachment relationships' contribution to adult personality.

[366] See Section 3.1 for an overview.

[367] On rationality requiring us to justify why we act on some desires rather than others, see MacIntyre, 2009b [1999], pp. 68-69.

[368] A third type of Subjectivist view that I will not have space to cover here are 'life-satisfaction' views; see Sumner for a lucid philosophical treatment that argues for a sophisticated account of happiness as life satisfaction, and gives an Appraisive account of well-being as 'the happiness of an informed and autonomous subject' – Sumner, 1996, p. 172. However, the critical arguments of the next chapter are readily adaptable to this view too.

[369] As we have seen, Bentham was a key proponent of this view, and it still has its advocates today – see below. For simplicity, much of the time I will not mention pain explicitly.

they include a greater balance of pleasurable episodes.[370] However, the version of hedonism we shall consider over the following two sections instead aims to make a contribution in Action-Guiding contexts: both as a view about *normative practical reasons* for action in the here and now, and as an account informing *what our lives as a whole should be like.*[371] According to this view, our capacities for pleasure and pain should be taken as our key sources of knowledge about what is good and best for us: in practical deliberation, the ultimate standard of appeal is always to ask which choice will bring us the most pleasure. We begin by returning to our discussion of decision theory, now considering the hopes for a technical solution to our core problem that interprets utility in terms of felt pleasure rather than desirability.

We noted above that some behavioural economists have expressed dissatisfaction with the mainstream preference-based economic approach, arguing for a new direction.[372] These researchers have distinguished two core concepts of utility. The first is 'decision utility', which refers to the weight that options are afforded in agents' actual choices, as per the operationalist approach to preferences within the mainstream tradition considered above. The second is 'experienced utility', which instead refers to the hedonic impact of a particular outcome, and which they argue should now replace revealed preferences as the central normative concept within the theory of rational choice. This second conception of utility is also acknowledged to be a return to the older usage of the term popularised by Bentham, whom many of these researchers see as a forerunner.[373]

One driving force behind this movement is the combination of an implicit endorsement of hedonism as an Action-Guiding theory of the Human Good, together with a slew of empirical results suggesting that individuals tend to make systematic mistakes when attempting to aggregate hedonic feeling over time.[374] In much of this literature, agents are thought

[370]For contemporary proponents of this Appraisive view within philosophy, see for instance Feldman, 2004; 2010; Crisp, 2006; and Bramble, 2016b; for critical discussion, see Sumner, 1996, Chapter 4; Griffin, 1986, pp. 7-10; Kraut, 2007, pp. 120-122; and Parfit, 1984, pp. 493-494.

[371]See Rawls, 1971, p. 554-560 and MacIntyre, 2011 [2007], p. 77 for quite different critical discussions.

[372]Though this is also something of a return to an older paradigm, since some early marginalist economists had roughly the same idea.

[373]For an overview, see Kahneman, Wakker, and Sarin, 1997; and Kahneman and Thaler, 2006.

[374]See the next section for a brief review of the key literature; see also Gul and Pesendorfer, 2008, Section 6 for criticism from an operationalist perspective.

to ultimately want to maximise their experienced utility in practice, and this is also endorsed as rational. Yet it is argued that their attempts to do this are often flawed, leading to a mismatch between experienced and decision utility. For instance, agents' 'remembered utility' from past episodes is seen as a key driver of actual future choices, but often fails to faithfully reflect the actual affective features of their experiences.[375] Thus, an agent's 'remembering self' is seen as 'an error-prone witness to the actual wellbeing of her experiencing self',[376] and an important task for these researchers then arises in correcting these apparent biases.[377]

Such empirical results also threaten to undercut the classical justification for preference-based foundations of decision theory, wherein rational agents' well-being was assumed to be adequately reflected in their final choices, thus making appeals to their subjective experiences otiose. Moreover, we saw earlier that a key reason for the original move towards revealed preferences and away from felt experience was the behaviourist idea that subjective experiences are unfit objects for scientific study; yet this is a point which researchers arguing for a change of focus have also challenged – thus claiming that 'experienced utility is both measurable and empirically distinct from decision utility.'[378]

For our present purposes, if pleasure and pain are indeed measurable in the way that these researchers have claimed, then perhaps this will provide us with the standard that has thus far been found lacking: such an affective conception of utility could enable us to make clear-cut, objective choices between any options we might face, including choices between competing desires, and perhaps may even lead to an Action-Guiding version of EUT that can be implemented algorithmically. So, it will be worthwhile to spend some time discussing the empirics of measurement. However, before doing so, we must first investigate the underlying notion of pleasure itself. This concept is often left somewhat under-theorised in the empirical literature; it will therefore be illuminat-

[375] Kahneman, Wakker, and Sarin, 1997, p. 375: 'Psychological research has documented systematic errors in retrospective evaluations, which can induce a preference for dominated options'; see the next section for examples.

[376] Kahneman, 2011, p. 392; Kahneman is casting doubt on the view here, having previously endorsed it – e.g. in Kahneman, 1999, p. 4; p. 20.

[377] See here the literature on the 'libertarian paternalism' movement: e.g., Thaler and Sunstein, 2003; Thaler and Sunstein, 2008.

[378] Kahneman, Wakker and Sarin, 1997, p. 376. From a philosophical perspective, see Angier, 2015 for doubts about measuring happiness, as well as criticisms of contemporary notions of happiness because they lead to the view that happiness is measurable; see also Griffin, 1986, Part Two for further discussion.

ing to consider its historical development within philosophy instead.

On perhaps the most straightforward kind of view, implicit in much of the empirical work we will be interested in, pleasure is thought of as a simple sensation.[379] It is indeed then tempting to think of different episodes of pleasure being 'weighed' against each other. Versions of hedonism employing this simple conception are raised by a number of ancient authors, some of whom also entertain the idea that such comparisons may be possible. Plato, for instance, bequeaths us a memorable exchange in *Protagoras*, in which Socrates exclaims: 'how else does pleasure outweigh pain, except in relative excess or deficiency? Isn't it a matter (to use other terms) of larger and smaller, more or fewer, greater or lesser degree?'[380] But perhaps the best-known ancient view of pleasure as a simple sensation appears in the Action-Guiding version of hedonism advocated by the Cyrenaic philosopher Aristippus (c. 435 – c. 356 BCE).[381]

Aristippus commended a life of focus on immediate, particular pleasures that were available close by in the moment: 'for he said, the present only is in our power, not the past or future; the one being gone, the other uncertain whether ever it will come.'[382] Despite the underlying conception of pleasure being similar, this strategy thereby contrasts with the hedonistic approach to living considered in the *Protagoras*, which instead focused on maximisation over the long term. For Aristippus, all pleasures are also commensurate and do not much differ, though bodily pleasures are perhaps thought superior to mental ones – a view again inverted by Plato, whom later philosophers have tended to follow on this point.[383] The Cyrenaics as a school all agreed that pleasure in this sense was itself sufficient for the good life, against the broad contemporary consensus amongst other ancient philosophers that virtue was necessary. Moreover, in keeping with this idea, they were also proto-egoists, attaching only an Instrumental Value to friendship – an important point that was raised earlier, and which we shall return to

[379] For discussion, see Katz, 2016, Section 1.1.

[380] Plato, *Protagoras*, 356a; see *ibid.*, 351b-356c for further discussion of the hedonist view; and see *Gorgias*, 496c-e for discussion of a different conception of pleasure, now understood as the alleviation of pain, typically accomplished by sating one's appetites.

[381] Our key source for Aristippus' views is Diogenes Laërtius: see Laërtius, 1991 [herein *DL*] – especially *DL*, II, 86-97. This author is not to be entirely trusted, however – he also informs us that Epimenides once slept for 57 years; see *DL*, I, 109.

[382] Stanley, 1656, p. 5 [translating Claudius Ælianus, *Various Histories*, 14.6].

[383] *DL*, II, 90.

again later on.[384]

Following antiquity, interest in hedonism in the West subsided for a long time due to the growing influence exerted by the Platonic and Aristotelian inheritance and then under the asceticism of Christianity. It regained ground after the Enlightenment, partly due to the preparing influence of Hume and Locke's sensationalism, and then rose to dominance in the 19th Century as a central part of the utilitarian theory developed by Bentham and his followers.

Bentham is today perhaps best known as a proponent of utilitarianism as a moral theory, wherein morally correct action consists in maximising aggregate utility, understood in a hedonistic sense; the Cyrenaic position thereby becomes inverted, with egoism transformed into an impersonal benevolence.[385] However, Bentham's primary interest was legal reform, and his writings also span psychology, economics and law as well as philosophy.[386] In tension with his moral theory, on these matters he instead opted for a hedonistic psychology that is distinctly egoistic:

> On the occasion of every act he exercises, every human being is "inevitably" led to pursue that line of conduct which, according to his view of the case, taken by him at the moment, will be in the highest degree contributory to his own greatest happiness.[387]

The notion of pleasure underlying our choices here is once again a simple, raw sensation which is analogous to – indeed, perhaps the 'opposite' of – pain, and for Bentham the two are the most salient aspects of the whole spectrum of human experience. In line with the recent empirical work on 'experienced utility' mentioned above, on this sort of view the only relevant attributes to consider in weighing up different pleasures and pains are quantitative: Bentham distinguishes intensity, duration, purity, fecundity, and propinquity. Pleasure as thus understood seemed

---

[384] *DL*, II, 91; 97; cf. Section 1.7 above.

[385] Contemporary utilitarianism has since become detached from this particular view of well-being; see e.g. Singer, 1979 for a view tied to the fulfilment of preferences instead, although in 2014 Singer reverted to the hedonist version: see here de Lazari-Radek and Singer, 2014.

[386] See Dewey, 2002 [1992], pp. 212-213 for discussions of utilitarian moral theory and psychology and their relation to the economic thought prevalent at the time.

[387] Quoted in Sidgwick, 1981 [1907], p. 41. Like the senior Mill, Bentham often directs his tracts to a social planner who is to engineer society so that individuals' pursuit of their own utility – in line with this 'grand principle' of human nature – will lead to the achievement of socially desirable ends; however, this does not eliminate the tension – for how do we account for the motivation of the benevolent engineer?

to lend itself well to quantitive treatment, and blithe talk about the nominal measurement of utility ensued in philosophy, economics, and even politics. Yet as I shall now argue, even on this simple conception of pleasure there are reasons to doubt that precise measurements are possible.

The contemporary ambition to measure pleasure in a systematic way also has an early precedent in the writings of the political economist Francis Edgeworth, who imagined this feat to be achieved with a measuring instrument known as an 'hedonimeter' – an 'ideally perfect instrument' that would record data by 'continually registering the height of pleasure experienced by an individual', which would then be integrated over time.[388]

In contemporary philosophy, a similar thought lies behind Feldman's argument for the possibility – 'in principle' – of measuring pleasure and pain. He begins by claiming that episodes of pleasure may be assigned both a duration and an intensity; if the intensity varies over time, we focus instead on the 'average' intensity.[389] Multiplying these two statistics together then gives us the total 'amount' of pleasure contained in an episode. Feldman then assumes that there is a 'standard unit of measurement for these amounts', which he calls 'hedons' for pleasure, and likewise 'dolors' for pain.[390] However, the argument is too quick here and begs important questions. Even if individual moments can thus be totally rank-ordered in terms of their pleasurableness – which is for the moment merely an assumption – this only shows the existence of an *ordinal* scale of measurement, and does not imply that a *cardinal* or ratio scale is available, such as would be necessary for the suggested mathematical operations to be meaningful. So, let us see if we can draw on the recent empirical literature to develop the argument further.

A more detailed approach to this issue has been developed by Kahneman and his collaborators, who define the technical normative notion of the 'total utility' of an episode, which Kahneman elsewhere calls 'objective happiness'. This concept is also based on the idea of integrating an instantaneous 'experienced utility' function that measures intensity.[391]

[388]Edgeworth, 1879; see also Colander, 2007 for a discussion; and Moscati, 2019 for a detailed reconstruction of the history of measuring utility.

[389]The precise mathematical notion of 'average' in play here is not specified; in fact, the obvious choice for formalisation would be to start with the integral and divide it by the duration, making the following multiplication step otiose.

[390]Feldman, 2004, p. 25; see also Section 1.6 above on 'well-being as commodity'.

[391]Kahneman, 1999, p. 5; see Kahneman, Wakker, and Sarin, 1997, and Feldman,

The analysis begins with a record of an individual's self-reported hedonic state over time, as assessed on a numerical scale, with 0 corresponding to the absence of feeling. This they term a 'utility profile'. They assume that subjects are able to accurately distinguish different levels of affect in an ordinal manner, so that experienced moments are at least ordered correctly within utility profiles; but they do not assume this ordering reflects an interval scale, so that the difference 'between 7 and 6 need not be psychologically equivalent to the interval between 3 and 2'.[392] Hence, it does not yet make sense to speak of integrating these utility profiles over time. However, from here they introduce a method of rescaling the utility function, in accordance with the subject's own comparative trade-offs between intensity and duration:

> For example, suppose that the observer judges that one minute of pain at level 7 is as bad as two minutes of pain at level 6. According to the theory, this judgement implies that the original reports of pain should be rescaled, assigning level 7 a value that is twice as high as the value assigned to level 6.[393]

Given some further technical assumptions, the authors then rigorously prove that the rescaled utility function will now have the mathematical properties required for us to meaningfully define the 'total utility' of an episode as its integral over time, as measured on a ratio scale.

This approach is clearly an advance on Feldman's speculations, and makes plain the possible difficulties that may be encountered by philosophers' sometimes-naïve ideas about the measurement of affective states when attempts are made to implement them in practice. However, careful readers may have noticed an issue here: the rescaling of the ordinal utility function is based on the agents' preferences over utility profiles! So, we are back to preferences after all.[394] But within the present enquiry we had already encountered difficulties in preference-based accounts and were now looking for an independent standard that might supplement them – hence, Kahneman's approach will not provide what we need here.

2010, Chapter 3 for discussion.

[392]Kahneman, 1999, p. 5.

[393]Kahneman, 1999, p. 6.

[394]This is most clear in the formal statement of Theorem 1 in the appendix to Kahneman, Wakker, and Sarin, 1997, p. 400: 'For a binary relation $\precsim$ defined over the (utility) profiles ...' Kahneman also points out that the approach is similar to the popular concept of 'Quality-Adjusted Life Years' – but this is also clearly related to agents' preferences. The resulting view is thus a version of preference-hedonism, in the sense of Parfit, 1984, p. 493.

Moreover, as we have seen, the same experimental literature has also concluded that preferences may be systematically biased if considered as reflecting attempts to maximise experienced utility, and indeed elsewhere Kahneman urges that we 'reject preferences as the final criterion for the value of experiences.'[395]

More generally, despite Edgeworth's early confidence, to my knowledge no convincing technical account of precisely 'measuring' raw affective pleasure independently of preferences has yet been given, and there have likewise been persistent barriers to constructing a scientifically valid nominal measurement system for pain. Collingwood's remarks on the matter therefore still ring true:

> no one has ever put forward a calculation belonging to such a calculus involving figures to which he will commit himself and which he will submit to a fellow-expert for independent checking, as a man will commit himself to an evaluation of π or a calculation of the Earth's diameter or the moon's weight.[396]

Moreover, prior to this being achieved, it is yet not evident what further conceptual and technical complications might arise here; as Edgeworth himself realised, we cannot talk about measurement independently of the actual process and instruments used for measuring, and indeed we cannot be entirely clear *a priori* on the nature of what we are measuring until we actually attempt to do it in practice. What is then needed is a viable tradition of measurement that yields definite, standardised values that are consistent across observers, as we now have for measuring length, weight, and temperature. Indeed, even in these apparently prosaic cases, physics was for a long time hampered by ambiguities caused by a lack of a sufficiently clear measurement procedure.[397]

Moving beyond this agnosticism, there are also positive reasons to doubt that it will ever be possible to adequately characterise the quality

[395]Kahneman, 1999, p. 10. Indeed, the third axiom used to establish the theorem, 'monotonicity in total utility', seems to be violated in the cold water immersion study described in the next section; see here for an extended discussion of failures to both predict and remember our affective states.

[396]Collingwood, 1992 [1940], p. 409; Collingwood is responding to W. D. Ross here.

[397]For a classic presentation of the 'operationalist' view wherein the meaning of a concept is revealed by its means of measurement, opening with a discussion of length, see Bridgman, 1927; on how conceptions of temperature were influenced by progress measuring instruments, see McCaskey, 2020; on the measurement of sensation, see Tal, 2020, Section 3.3.

of a subjective experience with a single numerical value.[398] It has been claimed that support for this project may been drawn from an older tradition of psychophysics, and the existence of correlations between verbal reports and objective physiological indicators such as facial expressions.[399] Yet even the notion that experience in general occurs via a singular mental 'channel' is controversial, as opposed to some kind of distributed representation wherein sensations are experienced in parallel across multiple parts of the brain. In the context of perception, this underlying serial picture of how the mind works has been critically described as the 'Cartesian theatre' by the outspoken philosopher of mind Dan Dennett.[400]

Kahneman initially suggested that future progress within neurophysiology may resolve these issues, predicting that 'Continuous records of affective state could possibly be derived from non-invasive measures of localized brain activity', and claiming that 'The movement from science fiction to practical application is likely to be rapid in this domain.'[401] Yet science fiction it remains, and it is difficult to see how the masses of data that might be collected on the neural activity that is the substrate of a pleasurable experience could be adequately characterised by just one number. One cannot so easily measure even a quantity as overtly manifested as physical strength: in strongman contests, different contestants excel at different events, which thus rank-order them differently, and even within a single discipline performance may fluctuate from one competition to the next – in practice, 'time and chance happeneth to them all.'[402] Until an acceptable methodology for constructing nominal measurements in practice is available, then, we will be wise to adapt a suggestion of David Fraser's and speak only of 'assessing' rather than 'measuring' pleasure.[403]

In his more recent writings, Kahneman himself has now come to

[398] Cf. Kahneman, 1999, p. 7.

[399] Kahneman, Wakker and Sarin, 1997, p. 380. I am skeptical about this broader tradition too, though I cannot go into this here – see again the recent JDM literature for contemporary criticism of such basic doctrines as Weber's law, suggesting that because of contextual variation in reports the whole conceptual approach may be misguided – e.g. Stewart, 2009; Luce, 2002.

[400] See Dennett, 1991, p. 107; for defence of the view that 'positive and negative affect are distinct phenomena that have separate neural processes and that serve different adaptive functions', see Fraser, 2013, p. 236.

[401] Kahneman, 1999, p. 10.

[402] Ecclesiastes 9:11.

[403] Fraser, 2013, pp. 234-236; here Fraser compares assessing an animal's welfare to assessing the safety of a lake. See also Hausman, 2015, Chapters 10-11 for some more general doubts about the scientific measurement of well-being based on its contextual variability, and a discussion of policy implications.

acknowledge that these matters are more complex than he had first thought, and for this reason has distanced himself from purely hedonistic views of happiness:

> The experience of a moment or an episode is not easily represented by a single happiness value. There are many variants of positive feelings, including love, joy, engagement, hope, amusement, and many others ...[404]

He therefore now endorses a 'hybrid' view that incorporates elements of 'life satisfaction', and even narrativity.[405] But his remarks here also suggest that the simple notion of pleasure may itself be inadequate as an account of its nature – and if this is true, then as with our initial correlate view of desires, any account wherein pleasure is thus treated as a particular type of sensation might then fail to faithfully reflect agents' Subjective Values as they really are. We therefore return to philosophy and consider some other, more sophisticated conceptions of pleasure that have been developed here – though as we shall see, the issues with quantitative measurement will now become even more acute.

Several generations after Aristippus, a very different kind of hedonist view was espoused by Epicurus (341-270 BCE), who found Aristippus a natural foil in articulating his position. Though Epicurus also held that pleasure is the sole good and pain the sole harm, his notion of pleasure was more akin to rationally-grounded tranquillity than a simple sensation. Moreover, he also argued that the successful pursuit of this more elaborate conception must ultimately coincide with the development of the virtues:

> it is not possible to live joyously without also living wisely and beautifully and rightly, nor to live wisely and beautifully and rightly without living joyously. For the excellences grow up together with the pleasant life, and the pleasant life is inseparable from them.[406]

Pursuing this goal would require pleasure to be sought in a more discerning way: rather than advocating the indiscriminate pursuit of pleasure

[404] Kahneman, 2011, p. 393.

[405] See Kahneman, 2011, p. 402, which also suggests that his underlying epistemology seems to be based on *de facto* preferences: 'We cannot hold a concept of well-being that ignores what people want.' See also Jarden, 2011, Section 2 for an interview wherein he reiterates both views.

[406] Epicurus, 2011.

itself, he attended carefully to the different sources of pleasure, and thus drew a distinction between groundless and natural pleasures, and within the latter between those that are merely *de facto* natural and those that are necessary.[407] Epicurus further argued that the connection between some such sources and the later suffering they tend to cause is robust enough to warrant regarding them as essentially inappropriate as objects of choice, despite this harm being external to them.[408]

Another feature of Epicurus' system is its eye to the long term: whereas Aristippus emphasises opportunism about pleasure in the here-and-now, here the focus is freedom from anxiety, pain and bodily distress over the course of a lifetime. In addition to employing foresight in individual judgements, this is to be achieved partly through character development: Epicurus' followers were encouraged to train, educate, and refine their affective tastes over time, as part of their intelligent pursuit of long-term pleasure – for instance, they were dissuaded from cultivating elaborate desires that were hard to fulfil.[409] Relating this to the present framework, this might plausibly be interpreted as a recognition of the context of Life Planning as well as Action Selection.

The modern hedonist tradition also underwent similar developments, and indeed, although we have seen it is undergoing something of a revival within behavioural economics, the view of pleasure as a simple, homogenous sensation attracted fewer philosophical adherents immediately after Bentham. In particular, his student John Stuart Mill, the second great classical utilitarian, famously did not follow him here, instead acknowledging that pleasures vary considerably in their phenomenal properties.[410] Mill also recognised a need to reply to those critics of hedonism who had proffered instances of pleasure which *pace* the doctrine seemed to be Intrinsically Bad for us independently of their consequences; such possible counterexamples to the theory include 'worthless pleasures' (for instance, those of degrading sexual experiences); 'immoral pleasures' (for instance, the pleasure obtained by a sadist when inflicting pain upon animals); and 'false pleasures' (those essentially based on some kind of

[407]Cooper, 1998; see also Sidgwick, 1981 [1907], p. 125 [footnote] on 'rational' versus 'impulsive' pursuit of pleasure.

[408]Epicurus, 2008, Section 8; see also Plato, *Protagoras*, 351d; and *Gorgias*, 495b; 499b-e.

[409]For a discussion, see Voorhoeve, 2018, Section 1.

[410]On this point, see MacIntyre, 2011 [2007], pp. 76-77, and Parfit, 1984, p. 493: 'Compare the pleasures of satisfying an intense thirst or lust, listening to music, solving an intellectual problem, reading a tragedy, and knowing that one's child is happy. These various experiences do not contain any distinctive common quality.'

illusion: for instance, the apparent affection of a partner that is only pretence). [411]

In response to these two issues, Mill famously introduced the non-quantitative distinction between 'higher' and 'lower' pleasures, understood in terms of what an agent who knew and was able to appreciate both kinds would 'give a decided preference' to.[412] Although these distinctions of quality are consonant with the more sophisticated hedonism of Epicurus, they clearly led to a far more moralised conception of utility than his teacher could accommodate: for Bentham 'push-pin is of equal value with the arts and sciences of music and poetry'.[413] At the same time, in Mill's writings there also emerges a clearer separation between utility as an abstract criterion that determines what is good for a human being, *versus* utility as a concept that serves as an Action-Guiding ideal for ordinary agents. In particular, as we saw in Chapter 1, for Mill it is neither necessary nor even desirable that agents aim solely at pleasure in their ordinary deliberations about what to do.[414] This move marks a key development in philosophical approaches to well-being – though of course, one which the present enquiry aims to reverse.

Both these new aspects of the evolving tradition – a skepticism of pleasure as a determinate kind of feeling, and of the view that individual agents are characteristically motivated only by securing certain mental states – were taken even further by the last of the three great classical utilitarians, Henry Sidgwick. In his monumental text *The Method of Ethics*, which we have met numerous times already, Sidgwick follows Mill in concluding that pleasures are a diverse category as far as their

---

[411]Lengthy discussions of these and other standard philosophical objections to hedonism – both classical and modern – together with associated responses, are readily available in the existing philosophical literature; see e.g. Feldman, 2004, Chapter 3 and 5 respectively; Sumner, 1996, Chapter 4; Moore, 2013, Section 2.3; Crisp, 2021, Section 4.1; and the extensive literature referenced therein.

[412]Mill, 1998 [1863], p. 56. This appeal to something like agents' 'revealed preferences' is similar to Mill's justification of his overall moral theory, since his primary argument for a hedonistic standard of the Human Good turns on the notoriously questionable move from pleasure's being desired to its being desirable in a normative sense. Mill also acknowledges that in practice some individuals do in fact choose sensual pleasures over intellectual ones; he tries to explain these cases away as instances of *akrasia*, but the argument is somewhat question-begging.

[413]Bentham, 1825, p. 206.

[414]Mill, 1998 [1863], pp. 64-65: 'this is to mistake the very meaning of a standard of morals, and confound the rule of action with the motive of it' – though as we have just seen, when he discusses well-being elsewhere, Mill's arguments often appeal to the actual choices of real-world agents. See also Section 1.2.3 above on self-effacing theories, and the discussion of the 'hedonic paradox' in the next section.

phenomenal properties go.[415] Instead of identifying pleasure with a fixed sensation, his approach was instead to demarcate it in a way that depended on its function. On this view, 'what pleasures have in common is not something internal to them – their peculiar feeling tone, or whatever – but something about us – the fact that we like them, enjoy them, value them, find them satisfying, seek them, wish to prolong them, and so on'.[416]

This strategy also improved upon the earlier utilitarian view of the Human Good and the rather crude Benthamite psychology it presupposes, which suggests that our inner lives consist only of two kinds of feeling. For the affective side of human psychology is in fact incredibly rich and complex: we experience not only sensory pleasure but also joy, gratitude, interest, enthusiasm, elation, awe, hope, relief, triumph, pride, amusement, admiration, love; not only felt pain but also discomfort, fear, longing, grief, sadness, disillusionment, dismay, betrayal, anger, rejection, helplessness, anxiety, nausea, fatigue, loneliness, irritation, *ennui*, boredom, outrage, humiliation, panic – each with their own distinctive characteristics, and it is puzzling that in Bentham these other states and their nuanced details are left out of account. However, despite this variety and heterogeneity, the classification of such affective states into those toward which we hold a positive or negative attitude nevertheless seems quite natural and robust; looking at the above list, we find it fairly easy to classify items in this way.[417] Such a classification then becomes available as the basis of an enriched hedonist position, leading Sidgwick to his more comprehensive formulation given in terms of 'agreeable over disagreeable states of consciousness' rather than pleasure and pain specifically. Moreover, some recent philosophical literature simply extends the term 'pleasure' to this broader usage.[418]

In the Twentieth Century, conceptions of pleasure become even more sophisticated, and many contemporary philosophers would perhaps not

[415] Sidgwick, 1981 [1907], pp. 111-115; pp. 125-131.

[416] Sumner, 1996, p. 86; see also *ibid.*, pp. 89-91.

[417] Within positive psychology, see Cohn and Fredrickson, 2009, p. 14 on the differences between positive and negative emotions; and Panksepp, 1998, p. 14; p. 41 for a biological approach. For a warning about oversimplification in appealing to valence here, see Lerner *et al.*, 2015, p. 804-806; consider also surprise as an emotion of ambiguous valence.

[418] Cf. Sidgwick, 1981 [1907], p. 127: "I propose therefore to define Pleasure ... as a feeling which, when experienced by intelligent beings, is at least implicitly apprehended as desirable or – in cases of comparison – preferable'; see also again Moore, 2013, Introduction.

now think of pleasure as a sensation at all, but rather as an emotion. Whereas sensations are raw, felt experiences, emotions have intensional or object-directed elements; many theorists have thus now concurred with Dewey that 'pleasures and pains in reflective individuals are inextricably bound up with ... "ideational" factors — that is, with articulate conceptions of what they are taking pleasure in.'[419] Moreover, in keeping with Sidgwick's approach, emotions are also typically associated with a cognitive 'valence'; an interpretation as good or bad.[420]

In recent years, hedonism has been taken even further in this intellectualist direction by Feldman. Instead of what he terms 'sensory pleasure', Feldman bases his developed Appraisive account of well-being on the distinct notion of 'attitudinal pleasure', which is now something quite different: 'Attitudinal pleasure is a mode of consciousness. It is a way of being aware of a state of affairs. It takes its place among such attitudes as hope and fear, belief and doubt, and recollection and anticipation.'[421]

Feldman also argues that attitudinal pleasure is again measurable. As with sensory pleasure, the argument proceeds by first pointing out that episodes of holding a favourable attitude towards some state of affairs have definite duration, and that these attitudes themselves have a certain 'strength'; that is, that it makes sense to say that someone would be 'more pleased' about one state of affairs than another. Feldman then once more proposes that we 'use the term 'amount of attitudinal pleasure' to indicate the *product* of average intensity and duration.'[422] However, this again appears to be rather simplistic. Attitudes are perhaps even less homogenous than sensations, so it is far from clear that this relation of relative strength is 'complete' in the economists' sense: that is, that any two such episodes can be definitively compared in this manner. But even if we accept that this is possible, Feldman appears to again be fallaciously inferring that ratio scale measurements of strength are possible just because ordinal measurements are possible.[423]

[419]Anderson, 2018, Section 4.1; see also Sumner, 1996, pp. 100-101.

[420]See again Cohn and Fredrickson, 2009, p. 14.

[421]Feldman, 2002, p. 607. For perhaps the first explicit discussion of pleasure as an attitude in the philosophical literature, see Nowell-Smith, 1954, pp. 111-115; though Feldman asserts that the distinction is in fact an ancient one and can be found in Epicurus – see Feldman, 2004, p. 55. See also Sumner, 1996, p. 108, who uses the term 'enjoyment' for an overall positive evaluative response to a situation, object, or proposition, and 'suffering' for the corresponding negative counterpart.

[422]Feldman, 2004, p. 65 [my emphasis].

[423]Cf. the same mistake in Ross, 2000 [1930], pp. 142-143, and Collingwood, 1992

More generally, although these more nuanced accounts of pleasure may more adequately reflect agents' Subjective Values, and – as we shall see in the next section – may also provide reasons to choose one form of pleasure over another, the qualitative dimensions thus introduced render the prospects of precise cardinal measurement of the kind required by EUT even slimmer. The distinctions between higher and lower pleasure drawn by Mill, for instance, are partly characterised in terms of the changing and often elusive judgements of certain preferred individuals over time as they gain more experience; and the later, more explicitly cognitive conceptions of pleasure exacerbate the issue even further because it is difficult to see how the articulate judgements that are partly constitutive of such episodes could be broken down into 'units' in order to be quantitatively weighed. Moreover, beyond laboratory methods that make use of sophisticated measuring apparatus and complicated mathematics, further practical difficulties arise when seeking a viable means of guiding the deliberations of ordinary agents. For even Feldman accepts that ascertaining the relevant numerical magnitudes he envisions would not be a realistic prospect for real-world agents: 'I grant of course that it would be very difficult in practice to locate precise dates for the beginning and end of an episode of attitudinal pleasure.' He continues by pointing out that 'the practical difficulty does not entail that there is no fact of the matter'[424] – yet for our present purposes such alleged free-floating 'facts' are to no avail, since actually implementing a technical hedonistic version of EUT would require agents to arrive at concrete judgements about these magnitudes in practice.

In the remainder of our discussion of hedonism, then, I shall follow Dewey's warning of not 'importing mathematics into morals'.[425] In the next section, I will assume that there is a fairly workable, broadly quantitive conception of pleasure that is cognitively available as a rough guide to practice for real-world agents making decisions, and which we will work with initially; but I will not aspire to expound a precise, technical decision procedure that it is possible to implement algorithmically. But having been thus forced to give up this aspiration once more, we will see that the situated judgement of the acting agent again becomes indispensable.

[1940], pp. 410-411 for a critical discussion.

[424]Feldman, 2004, p. 65.

[425]Dewey, 2002 [1922], pp. 50-51.

2.4 Hedonism as a Guide to Action

Having developed some insights into the different conceptions of pleasure, in this section we now consider the practical application of hedonism as an Action-Guiding theory. Rather than being directly controlled by pleasure or pain in the sense of Bentham's psychological hedonism, here the agent uses a hedonistic standard as a criterion to guide their explicit Practical Reasoning, thus deciding what to do and what kind of life to live – a process which may require them to resist the pull of particular pleasures in the moment.[426]

We begin with the context of Action Selection. Following the conclusions of the last section, here I will only aim for a fairly rough-and-ready account, rather than a precise technical implementation. We also sidestep the need to endorse a particular univocal analysis of the concept of pleasure by allowing for pleasure in any of the senses canvassed so far to contribute – including those that involve nuanced distinctions between higher and lower grades of pleasure, such as we find in Mill. Our hedonist therefore need not be an Aristippus, and may discriminate between particular pleasures in an intelligent way; for instance, by forgoing some sources of pleasure with a view to attaining more pleasure later.[427] But in so doing, they must grant final authority only to a recognisably hedonistic standard.

In the empirical literature on animal and human behaviour, the primary import of hedonic states is taken to be relational: they enable us to evaluate their corresponding objects as good or bad, and hence seek or avoid them accordingly.[428] In contrast, within the contemporary Appraisive study of well-being, philosophers offering a hedonist view are careful to clearly separate the state of pleasure itself from particular

[426] Cf. again Sidgwick, 1981 [1907], p. 125 [footnote]. The need for rational agents to maintain a certain distance from our Subjective Values in general will also be a key theme of the next chapter; cf. MacIntyre, 2009b [1999], p. 69; p. 83; p. 136.

[427] For the reasons identified in Section 1.4, with Action Selection we are not necessarily always aiming directly at maximising *lifetime* pleasure: the distant future may be too uncertain, and the temporal scope of some choices may be fixed exogenously.

[428] In the literature on animal behaviour, see Broom, 2013. p. 256: 'Feelings are part of a mechanism to achieve an end, just as adrenal responses or temperature regulatory behaviour are mechanisms to achieve an end'; and the series of papers by animal welfare scientist Marian Dawkins on the motivational functions of feelings in animals. Within psychology, see Grinde, 2002, pp. 146-147; Nesse, 1991 on the importance of a capacity for bad feelings for life outcomes; and the influential empirical work on 'liking' as a positively-valenced gloss on percepts and its connection to motivational salience, or 'wanting'; e.g. Berridge, 2009.

sources of pleasure, and in general take the position that only the former is of Intrinsic Value, whereas the latter can only be of Instrumental Value, as a means to pleasure.[429] Moreover, for those philosophers for whom pleasure is a simple sensation, these particular objects of pleasure are merely incidental.[430]

Regardless of which of these two theoretical stances is adopted, however, in the present Action-Guiding context, wherein pleasure must fulfil the epistemic role of telling us what the good is, it will remain true that in order for our deliberations to be manageable and meet the demands of real-world decision-making we will generally need to aim at some more-or-less determinate set of specific forms of pleasure – thus orienting ourselves toward particular sorts of experience that we believe that we will find pleasurable, or toward particular actions, items, or states of affairs upon whose achievement we believe that pleasure will supervene.[431] For although if what pleases us changes over time the hedonist must insist that our priorities for which concrete outcomes to pursue ought to change too, nevertheless if an agent is to come to timely decisions about what to do in the here-and-now, then at any one particular time they will require the basic hedonistic principle of action to be supplemented by a set of beliefs about what they will actually find pleasurable or painful. And indeed, in the *Protagoras*, Socrates considers an account of practical wisdom in terms of being knowledgeable about just this subject.[432]

Given this need for agents to have auxiliary beliefs about the sources of pleasure, one major issue with hedonism as an Action-Guiding theory is that in recent years there have been a large number of empirical studies suggesting that humans are in general quite poor at predicting what degree of pleasure and pain will accompany the consequences of their choices: what behavioural economists and psychologists call 'hedonic forecasting'.[433] There is therefore a serious practical concern that human beings cannot effectively perform the assessments that a hedonistic theory of Action Selection would require in order to be successfully implemented:

[429]See e.g. Feldman, 2004, p. 23: 'it is the pleasure we get from pleasant things that is intrinsically good'.

[430]For criticism of this view from an Aristotelian perspective, see again Angier, 2015; especially p. 23.

[431]Cf. Dewey, 2002, pp. 199-201; Mill, 1998 [1863], p. 70.

[432]Plato, *Protagoras*, 357a-b: 'our salvation in life depends on the right choice of pleasures and pains'.

[433]See Kahneman and Thaler, 2006 for an overview.

> to maximize [experienced] utility successfully, one must start by making a forecast about how the various possible outcomes will be experienced. And, if forecasts are systematically biased, then choices may systematically fail to maximize utility.[434]

In much of this literature, pleasure and pain are again thought of as simple sensations, so for now we continue on this assumption too.

One example of such a problem with forecasting is 'focusing bias', wherein agents have a tendency to inflate the hedonic importance of an attribute that is salient in the context of choice.[435] For instance, when evaluating several options presented together, agents may fixate on features that will have little hedonic impact when their chosen option is later enjoyed in isolation, in the absence of this standard of comparison. In one empirical study, subjects were asked to report their willingness to pay for either one or both of a pair of audio systems. In the absence of further information, consumers' responses were heavily influenced by the systems' relative sound quality in the context of a joint evaluation, where direct comparison was possible; but when comparing the reports of those consumers who rated only a single system, the differences were subtle enough to have little effect on the results.[436] And similar predictive failures may occur even if subjects are not explicitly asked to draw a comparison: for instance, subjects' predictions of the extent to which they will enjoy eating a bag of crisps are biased downwards by placing a bar of chocolate in view – as compared to their *ex post* reports – but biased upwards by a tin of sardines.[437]

Another example of a focusing illusion is 'diversification bias'. This arises in experiments where subjects are invited to select several goods simultaneously, but their choices are consumed separately over an extended period of time. Faced with such choices, agents tend to select more variety than is optimal because variety is more salient in the context of choice than in the actual enjoyment of the goods. For example, when a group of students were asked to simultaneously select three snacks from a menu of six options, expecting them to receive them some weeks later across three separate classes, they selected a mixture; but when selecting

[434]Kahneman and Thaler, 2006, p. 231.
[435]Gilbert and Wilson, 2000, 186-188.
[436]Hsee, 2000, p. 552-554.
[437]Morewedge *et al.*, 2010; see also Morewedge, Gilbert and Wilson, 2005 on forecasting biases resulting from salient but atypical past memories.

only at the start of each of the three classes, they choose the same one each time.[438] Summarising this work on focusing bias, some researchers have thus concluded that 'Nothing in life matters [in the hedonic sense] quite as much as you think it does while you are thinking about it.'[439]

A further source of bias in hedonic forecasting is that our predictions about our future affective states are unduly influenced by our present affective states. For example, in one study, agents were more likely to select an unhealthy over a healthy snack if they were hungry rather than sated at the time of choosing (they were asked either in the late afternoon or just after lunch), even though the chosen snack was delivered only a week later.[440] This anchoring of choices that only impact our future selves to temporary affective states in the present has been termed 'projection bias' by psychologists.[441]

Surprisingly, our difficulties with hedonic estimation also arise in retrospective contexts where the incident in question is already over and the agent has only to report their recent affective experiences. In one well-known experiment, patients undergoing a colonoscopy were asked to register the amount of pain they were experiencing using a ten-point scale, with measurements being recorded every second through a hand-held device.[442] Less than an hour after the colonoscopy, patients were then also prompted to assess the *total* amount of pain experienced during the procedure – also on a ten-point scale. It is clear that in general this estimated quantity will depend heavily on the overall duration of the procedure, which varied between 4 and 67 minutes – but in fact, the data showed very little correlation ($r = 0.03$) between how long a procedure took and the patient's assessment of the total amount of pain. This effect is known as 'duration neglect'. However, the estimated total amount of pain *was* correlated with both the maximum pain experienced ($r = 0.64$), and the painfulness experienced at the end of the procedure ($r = 0.43$). Kahneman and others have therefore suggested that 'Retrospective evaluations appear to be determined by a weighted average of "snapshots" of the actual affective experience, as if duration did not matter';[443] and more specifically that agents' estimates of aggregated

[438] Simonson, 1990, p. 153-156.
[439] Kahneman and Thaler, 2006, p. 229.
[440] Read and Van Leeuwen, 1998.
[441] See e.g. Loewenstein, O'Donoghue and Rabin, 2003.
[442] Redelmeier and Kahneman, 1996.
[443] Fredrickson & Kahneman, 1993, p. 45; for further discussion, see Kahneman, Wakker and Sarin, 1997, pp. 383-385; and Kahneman, 1999, pp. 15-17.

affect may be modelled using a 'peak-end' rule.[444] One reason for this may be the relative ease of applying such simple heuristics, compared to attempting to aggregate the entire range of our perceptual experiences directly.[445]

Another experiment provides further empirical evidence for this view that the 'end' of an experiential episode is weighted too heavily in aggregative estimates of total affect.[446] In this setup, each subject was exposed to the same two aversive conditions, and asked about the level of discomfort experienced. In the first condition, subjects immersed a hand in cold water (14.1°C) for one minute. In the other condition, the subjects first also immersed a hand in cold water for one minute; however, this trial was then continued for a further 30 seconds, during which time the water temperature was gradually raised (to 15°C). *A priori* it is clear that the second procedure must involve more total pain than the first, as was indeed registered by the moment-by-moment reports. However, when just seven minutes later the subjects were asked which of the two conditions they would like to repeat for a third and final trial, 22 of the 32 subjects chose to repeat the *second* trial rather than the first one – perhaps because their judgements about total pain were overly responsive to the less severe discomfort experienced in the additional final stage, when the water temperature became less disagreeable.[447]

Such retrospective studies raise the possibility that we may not manage to learn about what will please or pain us even through direct experiences if our memories of these are biased or not coded in the right way, and there is a substantive literature dedicated to exploring this general issue too.[448] Thus, the kind of knowledge envisioned by Socrates in *Protagoras* may never develop.

These obstacles to a hedonistic account of Action Selection seem daunting – but perhaps the problems identified so far might be at least partially overcome. Indeed, the related 'libertarian paternalism' movement acknowledges these biases but nevertheless aims to find interventions at the level of policy that can help people make 'better' choices (that is, according to a Subjectivist standard).[449] And the modern

[444]Kahneman, 1999, p. 19.

[445]Kahneman, 1999, p. 15.

[446]Kahneman *et al.*, 1993; see Kahneman, 1999, pp. 19-20 for further discussion.

[447]See also Schreiber and Kahneman, 2000; and Fredrickson and Kahneman, 1993 for similar studies using aversive sounds, and short, plotless films, respectively.

[448]See Kahneman, Wakker, Sarin, 1997, p. 385-386.

[449]As mentioned in a footnote in the previous section; see again e.g., Thaler and

discipline of happiness studies has now made considerable progress in forming generalisations about what gives us pleasure that agents might usefully employ in their practical deliberations.[450] On an individual level, it is possible that our hedonic forecasting could improve with practice or training; moreover, existing research already suggests that failures of hedonic forecasting may be lessened in highly familiar settings, and that agents may make better decisions when encouraged to think them through carefully rather than making intuitive guesses, or when their attention is directed to appropriate considerations: for instance, in the cold water immersion study, agents did select the shorter procedure when it was described verbally.[451]

Given this ongoing debate, the case against hedonism as presented so far is therefore still a long way from being conclusive. What these studies on the failure of affective forecasting do show, however, is that successfully maximising our affective states is harder than it may first appear – and as we found with desire-based views, thus requires substantial cognitive skills. In particular, determining the efficacy of different potential sources of pleasure requires fairly complex means-end reasoning and sophisticated powers of judgement. This point will become important later; for now, we turn to a second issue that is perhaps even more pressing.

A second potential problem with hedonism is the so-called ‘hedonistic paradox’: the idea that hedonism in the context of Action Selection is self-defeating, since by aiming at pleasure in our choices we tend not to achieve it – even in the absence of the kind of predictive failures just described. Analysis of this point is an active area of discussion within both the social sciences and philosophy today,[452] though it also has a longer history and was raised by earlier philosophers, including Mill:

Sunstein, 2008.

[450] See Layard, 2011 for a popular summary of some key results, arguing for the view that ‘we can train ourselves in the skills of being happy [in a hedonistic sense]’ – p. 189 – and advising that we should focus on family relationships, our financial situation, work, community and friends, health, personal freedom, and personal values (with the first five items here given in order of importance – see pp. 62-63). See also Clark, 2018 for a summary of relevant work in economics, and for some more empirics on factors related to happiness, see Blanchflower and Oswald, 2011. However, most of these studies are merely correlational, meaning that whilst these features may be good *predictors* of our affective states, one perhaps cannot *cause* such states simply by actively pursuing them; see also Annas, 2011, p. 129 for the same point.

[451] Kahneman and Thaler, 2006, pp. 228-229.

[452] Within economics, see e.g. Konow and Earley, 2008; and within philosophy Angier, 2015, p. 16-17; Roger Crisp, 2006; Parfit, 1984, p. 6.

> But I now thought that this end [pleasure] was only to be attained by not making it the direct end. Those only are happy (I thought) who have their minds fixed on some object other than their own happiness ... Aiming thus at something else, they find happiness by the way.[453]

Perhaps the first systematic discussion occurs in Sidgwick, who writes:

> Happiness is likely to be better attained if the extent to which we set ourselves consciously to aim at it be carefully restricted.[454]

As we have seen,[455] Sidgwick's response to the issue was to refrain from advocating hedonism in an Action-Guiding sense, holding it only as an Appraisive account of well-being and an abstract standard for right action. Since this response is not available to us here, how else might we proceed in defending a hedonistic approach in the present context?

On the face of it, the central claim here is somewhat puzzling – hence the term 'paradox'. Consider a particular choice faced by an imaginary young man, Simon: whether to enrol at university and study military engineering, or instead take on a job at the family business, his mother's bakery.[456] Suppose the decision is made on hedonistic grounds: our young man aims to make the choice that will give him the most pleasure. Then the choice advocated by the theory here is simply whichever one will in fact yield him the most pleasure, all things considered. If an unfavourable outcome results, then, we might take this merely as showing that he has made his calculations incorrectly – and it might then come to seem that by definition the theory *cannot* be self-defeating in the manner suggested.

What this reply conceals, however, is that Simon's mode of decision-making here may itself directly contribute to a negative outcome, so that if he makes his choice in *this* way – rather than, say, working out what propriety or familial loyalty requires of him – then whichever option is taken he will later feel unhappy.[457] This interpretation has

[453] Mill, 2018 [1873], p. 82; see also Kant, *Groundwork*, 4:395: 'we find that the more a cultivated reason purposely occupies itself with the enjoyment of life and with happiness, so much the further does one get away from true satisfaction'.

[454] Sidgwick, 1981 [1907], p. 405.

[455] Cf. the previous section and Section 1.2.3 above.

[456] This is a watered-down version of a classic example discussed by Sartre.

[457] See again Angier, 2015, pp. 13-14 for the related Aristotelian view that the causes of happiness are 'formal' (i.e. constitutive) rather than merely 'efficient' (or instrumental).

been supported by recent empirical research in psychotherapy that guards against treating subjective feelings as guides or goals in life. According to this literature, centring our decision-making on achieving certain affective states may lead to mental health disorders. Such a strategy would then be self-defeating because such conditions have a large negative impact on our affective lives. For example, in one study, a causal link was found between a subject's tendency to avoid negative experiences and their susceptibility to later developing a clinically diagnosed mental health problem.[458] Studies such as this one militate against even the considered hedonism of Epicurus, for whom it is a core tenet – intended as an Action-Guiding principle – that one should first and foremost aim to avoid felt anxiety in all its forms.

Though this objection to hedonism at the level of Action Selection is again not conclusive as it stands, the discussion so far does suggest that it may be worthwhile for hedonists to consider an alternative form of the doctrine: one instead posited at the broader level of Life Planning. On this revised view, pleasure and pain thus serve as standards that help us choose between different kinds of lives taken as a whole. Such a version of hedonist must of course counsel that in thus choosing between possible conceptions of what we intend our lives to be like we should choose the most pleasurable. But how pleasurable our lives are might crucially depend upon what kind of person we become, since it has just been suggested that it may turn on the mode of making concrete choices in the here-and-now we adopt. Hence, the Life Plan we thus choose may not be one in which we base all of our local decisions on pleasure at the level of Action Selection, or one wherein we transform our affective commitments so that we 'only care about pleasure'. Indeed, our Life Plan might instead include a commitment to making ourselves into the kind of person that allocates Intrinsic Value to certain other goods at the level of particular concrete choices.[459] So, if choosing based on what is

[458]Spinhoven *et al.*, 2014; for the view that a hedonistic, 'sensation seeking' focus is associated with substance abuse, see Zuckerman, 2007, Chapter 4; though conversely, see Pizzagalli, Jahn and Shea, 2005: these authors find that a lack of responsiveness to reward (as measured by a response bias in a probabilistic signalling task) may be used to diagnose depression, though the effect may be merely correlational.

[459]See Section 1.7 for anticipation of this point. Conversely, Mill shows signs of *not* being a hedonist at the level of Life Planning, introducing the apparently non-hedonic concept of a 'grade of existence' based on intellectual capacity, whose importance derives from the 'sense of dignity' possessed by rational beings, and famously claims that it is 'better to be Socrates dissatisfied than a fool satisfied.' See Mill, 1998 [1863], p. 57.

morally right is what will leave Simon feeling most satisfied, then this is what he should do in the here-and-now; and perhaps over time he should also strive to become the sort of person who has a tendency to choose in this manner in general.

Though this interpretation of hedonism may thus go some way to alleviating the issue just raised at the level of Action Selection, we should also note that the ongoing need to update or revise our life choices later on will still place limitations on the kind of life we have available to choose from if our practical reasoning is to still retain a distinctly hedonistic character at this more abstract level of Life Planning.[460] For whatever determinate plans we might make for our lives, maintaining a purely hedonistic approach to Life Planning requires that these must always be available for revision if circumstances change so that they no longer align with pleasure.[461] Although our local, proximal modes of reasoning in the here-and-now may operate for a time independently of a hedonic standard, as part of our striving to become a particular kind of person or engage in particular projects that we have previously judged will serve hedonic ends, it still remains true that ultimately 'Aiming at pleasure, understood as a way of living my life, will come, on this view, to nothing more than manipulating other circumstances so that I get as many of these episodes as I can.'[462]

The limitations this restriction imposes become apparent when we consider our personal relationships – itself a key topic for us going forward. In particular, although she might contingently allow herself to care for others for their own sake, and thus not treat them in a straightforwardly-Instrumental manner within particular interactions, a pure hedonist cannot engage in relationships that are bound by truly unconditional commitments. For any such association must always be available for later modification as her Life Plan is updated over time in light of changing circumstances; otherwise, pleasure would be relinquishing its authority and no longer serving as the ultimate guiding standard here.[463] Thus although she may, for example, express a commitment of sorts to another through an agreement to marriage, it may always later come about that

[460]See again Section 1.7 above for the same point.

[461]Hedonism might plausibly incorporate 'smoothing' effects, where we avoid possible inefficiencies created by changing our Life Plans too often – thanks are due to my friend Jack Leeming for this suggestion.

[462]Annas, 2011, p. 149; intensity is of course important as well as quantity.

[463]Cf. Section 1.2.1 on the need for ongoing inputs to decision-making as our lives unfold.

she should end the marriage in divorce if she comes to recognise that another option for how to live is superior in hedonic terms. And likewise for action motivated by a concern for the needs of others: other-regarding commitments and immediate priorities at the level of Action Selection can be part of the kind of life we aim towards only insofar as in the larger scheme of things our being such a person continues to serve the Final End of pleasure.[464] If such a view is to remain a Subjectivist one, anything other than pleasure must still be seen in Instrumental terms at the more fundamental level of Life Planning, and so the quest for the good and the quest for pleasure ultimately remain one and the same.[465]

Once it is thus acknowledged that if a Subjectivist conception of practical reason is to be adhered to over time then we must remain open to revising our priorities according to a hedonic standard at the level of Life Planning, it then becomes clear that this expanded account also relies heavily on our capacity for predicting our future mental states – which in this longer-term context has been called 'affective forecasting'.[466] And unsurprisingly, the same issue of predictive failure encountered at the level of Action Selection also arises here. For instance, when thinking about large-scale life choices, we again tend to overemphasise the long-term hedonic impact of a present outcome that is currently in focus – such as the breakup of a romantic relationship, or the failure to achieve tenure.[467] For instance, in one study, U.S. students in both California and the Midwest tended to believe that other students who lived in California would on average be much happier than those in the Midwest, in part due to the perception of an improved climate – yet in fact average self-reported overall life satisfaction was 'virtually identical' in the two locations.[468]

From studies such as this one, Kahneman and Thaler again conclude that 'people are systematically wrong in their expectations about the life circumstances that will increase or decrease their happiness, which in turn implies that life choices that people make in their pursuit of happiness are

[464] Cf. Aristippus's view that friendships are only Instrumentally Good; see here MacIntyre, 2009b [1999], pp. 113-114.

[465] Cf. Plato, *Gorgias*, 500d; see also Taylor, 1992, pp. 58-59 for a discussion of Subjectivism and its connection to taking an Instrumental attitude toward others.

[466] See Wilson and Gilbert, 2003.

[467] Gilbert *et al.*, 1998.

[468] Schkade and Kahneman, 1998; p. 342; the exact estimated difference in overall average was only 0.01 (in favour of CA) on an 11-point Likert scale – well below the threshold of statistical significance.

also likely to be wrong [according to a hedonistic standard]'.[469] Above we found some grounds for hope within narrower choice contexts; yet the added complexity stemming from the broader temporal scope of these predictions means that the difficulties are compounded considerably. Navigating this terrain successfully is therefore again not easy, and as before will require considerable powers of intellect and judgement. Moreover, there is again also a second, possibly even more problematic issue to consider.

As well as failures of affective forecasting, another robust finding in the empirical literature that threatens to cause problems here is that following even a drastic lifestyle change our average affective states tend to revert to a certain fixed, predetermined level – an affective 'set point' or baseline measure – as we become used to our new mode of living. Just as our nervous systems become habituated to a stimulus, in hedonic terms we thus soon adjust to what from an external perspective may seem like a better or worse life, and changes are only registered in the short-term. This effect is known to psychologists as 'hedonic adaptation',[470] and is an instance of a more general phenomenon: 'The human mind is extraordinarily sensitive to changes in conditions, but not so sensitive to absolute levels'.[471] Moreover, though some psychological and medicinal techniques are believed to influence individuals' set point, it has a large genetic component.[472]

Perhaps the best-known example of hedonic adaptation in action is described by a landmark classic study reporting that individuals' happiness levels one year on from either winning the lottery or breaking their spine in an accident are roughly the same as before the event.[473] Many people find these results highly surprising: they might dream of winning the lottery, or believe that they would rather die than become paraplegic, not expecting that as little as one month after their accident, people with paraplegia are on average in a fairly good mood more than half of the

[469] Kahneman and Thaler, 2006.

[470] Frederick and Loewenstein 1999; see also Diener, Lucas, & Scollon, 2006 for some later advances. One proposed explanation for the effect is that the brain attempts to minimise surprise, which is achieved by having the most commonly encountered value be experienced as a neutral default.

[471] Haidt, 2006, p. 85; see again prospect theory, which incorporates a 'baseline' level of consumption: Kahneman and Tversky, 1979. For an illustrative study of adaptation to distortions of faces that 'resets' the baseline, see Young and Burton, 2017.

[472] Haidt, 2006, p. 87; meditation has been argued to help, and long-term use of anti-depressant medication is now common in the United States.

[473] Brickman, Coates, and Janoff-Bullman, 1978.

time.[474]

From the resulting conclusion that 'Most environmental and demographic factors influence happiness very little', and assuming a hedonistic model of well-being, renowned psychologist Jonathan Haidt comes to a surprising conclusion: 'When we combine the adaptation principle with the discovery that people's average level of happiness is highly heritable, we come to a startling possibility: *In the long run, it doesn't much matter what happens to you.*'[475] On the face of it, this seems problematic for assigning pleasure such a central role in moral philosophy: if someone were responsible for the care of a child, for instance, and sincerely expressed this view – that it actually did not much matter what happened to them – we would usually take it to be immediately clear that they were unsuitable for the role. But more importantly for the present enquiry: if felt pleasure and pain alone do not provide us with any means to distinguish different outcomes in the long run, then it would appear that they cannot help us much in guiding Life Planning. And agents would then need to form judgements about what to do that outstrip their estimates of future pleasure and pain, showing hedonism to be seriously incomplete as an Action-Guiding theory.

One possible solution here is to retreat to more sophisticated conceptions of pleasure, such as those bound up with articulate conceptions of what we are taking pleasure in: for instance, Feldman's notion of attitudinal pleasure.[476] In elaborating his preferred version of attitudinal hedonism, Feldman argues for the incorporation of adjustments to the value of episodes of enjoyment based on objective features of the circumstances in question. In particular, following Mill's distinction between higher and lower objects of pleasure, he argues that the contribution of enjoyment to well-being in the Appraisive sense should depend on whether a state of affairs is 'pleasure-worthy' and hence 'deserves' to be enjoyed.[477] If the adaptation principle renders the raw affective experience of pleasure inadequately formative about what kind of life is best, then, perhaps some such distinction can be employed to rescue the theory by directing us toward lives that yield more rather than less valuable sources of pleasure.

[474] Kahneman and Thaler, 2006, p. 230.

[475] Haidt, 2006, p. 87 [my emphasis].

[476] Cf. the previous section.

[477] Feldman, 2004, p. 119; see also p. 73 for a version based on taking pleasure in 'higher' objects, and p. 121 for an agent-relative 'desert-adjusted' version.

Though this is a promising line of thought, it of course leads us only as far as the further question of how agents are to ascertain whether a potential source of pleasure is choice-worthy or not in practice. And as with restricted desire-based views, here again the possibility arises that these evaluations must be conducted according to a standard that is not itself derived from our capacity for pleasure and pain. If this is correct, then enjoyment itself, whether as a feeling or attitude, will again no longer be of sole fundamental importance in deciding what we are to do.[478]

We leave aside the issue of whether the hedonist tradition has the resources to provide adequate guidance here, noting only that even if this is so, then once again certain advanced powers of practical judgement will need to play an essential role in enabling the agent to determine what these higher objects of pleasure are. On Feldman's Appraisive version of the view, however, the issue does not arise, since as we have seen all connection to practice has now been severed: he can therefore take such distinctions of value to be brute, objective facts that lie outside the scope of his theory.[479]

In closing our discussion, it will be helpful to briefly contrast hedonism with an Aristotelian view of practical reason, wherein the role allocated to pleasure is quite different. On Aristotle's view, particular types of pleasure are indeed evaluated according to an independent standard: namely, whether their pursuit contributes to the higher end of *eudaimonia*.[480] Moreover, although acknowledging the motivational salience of pleasure and the consequent intuitive pull of hedonism, Aristotle argues that the attitude of the wise to pleasure should be the same cautious response that the Athenian elders had toward Helen of Troy due to the potency of her physical beauty and its capacity to motivate men to engage in vicious actions.[481] In particular, rather than striving to have as much pleasure as possible, we should instead aim for a condition wherein pleasure discharges its motivational role by moving us towards those genuine goods that we as mature human beings have reason to value, and wherein pain likewise directs us away from genuine harms that we have reason to avoid.[482] On this view, then, as we develop as

[478] See again Rawls, 1971, pp. 554-560.
[479] Feldman, 2004, p. 78.
[480] MacIntyre, 2016, p. 40.
[481] Aristotle, *NE* II.9, 1109b8-11; cf. *ibid.*, II.3, 1105a6-7: 'to feel delight and pain rightly or wrongly has no small effect on our actions.'
[482] Cf. MacIntyre, 2016, p. 90.

practical reasoners we must learn to feel pleasure and pain at the right times, to the right extent, and towards the right objects – a process of education that begins in childhood, when help from others is essential:

> we ought to have been brought up in a particular way from our very youth, as Plato says, so as both to delight in and to be pained by the things that we ought; for this is the right education.[483]

The virtuous person thus finds it pleasurable to pursue the good, so that reason and the passions speak with one voice; yet pleasure itself is not the overall goal of her activity, and must sometimes be guarded against.[484] Most importantly, for Aristotle, the practically wise person will sometimes knowingly forgo pleasure – even long-term, all-things-considered pleasure – in order to cultivate the virtues.

When all is going well with our character development, so that we are indeed pleased by what we have good reason to be pleased by, pleasure might then still function as a useful proxy for our practical deliberations to aim at on this Aristotelian view.[485] Yet we must nevertheless also remain ever open to the possibility that what we currently take pleasure in has diverged from what is truly valuable for us, so that our affective drives now require realignment rather than satiation. And as we shall discuss in detail in the next chapter, this openness to finding reasons to educate our Subjective Values that are not themselves based on maximising their satisfaction is incompatible with Subjectivism in any of its forms.[486]

2.5 Conclusion

In this chapter, we have drawn on different theoretical perspectives to explore the practical application of our two Subjectivist conceptions of the Human Good, and discussed a number of issues that arise here. We have seen that both our key theories may be framed at either the level of Action Selection or that of Life Planning, and hence must accommodate

[483]Aristotle, *NE* II.3, 1104b11-13; see also; *NE* III.12, 119b10-17; *NE* X.4, 1174b14-23.

[484]Though see Aristotle, *NE* X.4, 1175a18-10 for a less committal view.

[485]Cf. MacIntyre, 2011 [2007], p. 187.

[486]Cf. MacIntyre, 1993, p. 15 '[on this view] there can be no good reason for intentionally bringing about such a transformation which is not itself grounded upon an existing preference.'

each of these possibilities in the critical discussion that follows. Yet regardless of which of these alternatives the Subjectivist follows, we have further argued that no purely technical, algorithmic framework based on consulting our raw Subjective Values has been found that is adequate to the particular problem we have set ourselves, and that a capacity for sound practical judgement has remained essential – though we have not yet ruled out the possibility of a more developed Subjectivist theory itself providing adequate standards for this deliberation to follow. In the final chapter, we shall attend to what prerequisites must be in place to enable the acquisition and maintenance of this capacity, as well as certain other cognitive and practical powers that are also required for implementing any Subjectivist account in practice.

Chapter 3

The Social Context of Practical Reason

> *Individualism conceives a man as if he were a God, a self-contained and self-sufficient creative power whose only task is to be himself and to exhibit his nature in whatever works are appropriate to it. But a man, in his art as in everything else, is a finite being. Everything that he does is done in relation to others like himself.*
>
> – R. G. Collingwood[487]

In this final chapter, I offer a negative argument to the effect that the individualist assumptions shared by our two Subjectivist approaches necessarily render them inadequate as comprehensive Action-Guiding theories of the Human Good. Here I focus on two dual capacities at the heart of practical reasoning: forming our own conception of what is good for us, and being able to follow this in practice. Since these are preconditions for us benefitting from rational reflection on the good life at all, they are powers that *any* Action-Guiding account must presuppose from the outset. Yet I will claim that in order to first develop these capacities and to then maintain them throughout their lifetimes, human beings must participate in certain formative social relationships. Moreover, full participation in these relationships requires that we do *not* understand what is good for us in purely Subjectivist terms. If we are to become and remain rational agents at all, then, we cannot be guided only by consulting our untutored inclinations and feelings.

[487] Collingwood, 1958 [1938], p. 316.

3.1 Practical Reason and its Requirements

In both contemporary analytic philosophy and economics, work on well-being and practical rationality today typically considers only the situation of an independent and competent adult whose affective tastes, desires, preferences, and life ambitions are already fully formed – somewhat as if human beings 'sprung out of the earth, and suddenly, like mushrooms, come to full maturity without any kind of engagement to each other.'[488] In cataloguing the limitations of Subjectivism in this chapter, we will instead look at human lives as a whole, beginning in infancy – a time when we are not yet capable of action proper, and of only limited behaviour. In the first part, we investigate the social prerequisites for our being able to develop our conceptions of the Human Good and to incarnate these ideals in our lives. Then, in the second part, we examine the support required from others if these dual capacities are to be sustained later in adult life.[489]

The argument will chart a course between Subjectivist and objectivist approaches, thus aiming to avoid the issues endemic to each.[490] Rather than following the objectivist strategy of defending a specific list of concrete items that agents *should* value, or indeed attempting to systematically defend a positive ethical view at all, I shall instead aim to proceed with the critical argument by illuminating the actual empirically-observable *processes* by which successful rational agents come to identify goods in practice – and in particular, which resources they must acknowledge and make use of in order to gain and maintain their status *as* rational agents who are able to do this.[491] I shall argue that

[488]Hobbes, 1949 [1651], p. 100; cf. Bowles, 1998 on mainstream economics. For the same point in orthodox psychology, see Bowlby, 1997 [1969], p. 4, and Dewey, 1998, Vol. 2, p. 47; Dewey, 2002 [1992], pp. 93-94; in philosophical sociology, see Mead, 1934, Chapter 29; in political theory, see Taylor, 1992; MacIntyre, 2009b [1999], pp. 81-82; and MacIntyre, 2011 [2007], p. 290-291 on Rawls and Nozick's view of 'entry into social life as – at least ideally – the voluntary act of at least potentially rational individuals with prior interests'. For notable exceptions in philosophy, see Kauppinen, 2008; and Rosati, 2006.

[489]See MacIntyre, 2009b [1999] for an extended discussion of this dependence and vulnerability and its implications for moral philosophy, especially Chapters 8-10; see also the important work of Eva Kittay; e.g., Kittay, 1999; especially chapter 2.

[490]For an 'objective list' theory of the good life, see e.g. Nussbaum, 2011, p. 33 in the context of the 'capabilities approach'.

[491]This is somewhat similar to what Dewey calls a 'theory of criticism'; a notion introduced alongside his attack on the assumption that 'it is the province of moral theory to reveal moral goods; to bring them to consciousness and to enforce their character in perception'; see Dewey, 1998, Vol. 1, p. 99. See also MacIntyre, 2011 [2007], p. 317 on some ends being discovered through ongoing activity rather than

some of these goods only appear in a social context, being made available to us only through those from whom we must learn what we need to learn to become practical reasoners, and on whose continuing support we must depend if we are to retain this competence throughout our lifetimes.[492] What we need cannot always be supplied by our extant ideals and values, or come from within ourselves at all, but is only accessible through having an attitude of openness towards these supportive others. Hence, neither Action Selection nor Life Planning can ultimately consist in merely drawing out what we *qua* individuals already desire or find pleasurable – and indeed, sometimes we must critique, modify, and educate our Subjective Values rather than merely satisfy them. We begin by drawing attention to four commonplace characteristics of human nature and development, as background to the central argument.

One striking fact about human development is that a newborn baby is amongst the most vulnerable and dependent offspring in the animal world.[493] This forms a marked contrast with the progeny of some other mammals, such as horses, for instance: a foal can walk minutes after birth, a feat which takes humans some 9-15 months.[494] In many insect and amphibian species, meanwhile, the parents are not even present when their young hatch from their eggs. And for some spider species, life begins by eating the carcass of the mother.[495] It is not quite true that human babies are born with no instincts: they intuitively know how to suckle, for instance, and to cry for food when hungry.[496] But in terms of practical reason they could scarcely be less proficient, not even being able to lift their own heads. And though we feel pleasure and pain as infants, and may perhaps be said to have desires in a limited sense, these nascent psychological features are not yet coherent or developed enough to serve to guide our day-to-day lives. At this stage, then, neither our instinctive behaviours nor our inchoate Subjective Values add up to a viable strategy

specified in advance by theory.

[492]Cf. again MacIntyre, 2009b [1999], p. 71.

[493]Slater, Johnson, and Muir, 2017, p. 5: 'The newborn infant is a helpless creature'. In biological terminology, we are an altricial rather than precocial species – though with some aspects of precocial development; see Zeveloff and Boyce, 1982 for a classic discussion; and see also Kellman and Arterberry, 2000, p. 27-28. Unsurprisingly, altriciality is highly correlated with degree of parental care – the subject of the next section.

[494]Though other types of mammals are highly altricial, including primates, marsupials, and rodents.

[495]Salomon *et al.*, 2015.

[496]Dewey, 2002 [1992], pp. 65-69; see also Section 3.2 below.

for living.[497]

As we shall soon see, this meagre behavioural range of baby humans at birth entails that they are at first radically dependent on their caregivers; but it is also linked to another notable feature that is key to our competence later in life and success as a species. This is our second core characteristic of human development: a profound capacity for learning, which we understand very broadly as behavioural change in light of experience.[498]

One advantage of learning over instinctive behaviour is that it can generate more apt responses to the specific environment an organism finds itself in, which may be different to those encountered in its phylogenetic history. Rather than just making organisms innately disposed to behave in specific ways, evolution has instead interposed a mechanism that facilitates rapid adaptation to new situations, thus enabling behaviour that is much more flexible.[499] This is particularly useful in an environment that is especially changing and unpredictable.[500] Moreover, partly because the specific experience acquired during an organism's lifetime is typically at least somewhat idiosyncratic, a more pronounced capacity for learning tends to result in greater individual differences in behaviour across members of a species, as well as greater changes within the same individual over time; this includes those differences in our Subjective Values that form the basis of the Subjectivist theories under discussion.[501] Importantly for us, human beings also have an especially enhanced capacity for learning from other members of their species, rather than directly from the physical environment.[502]

At around three years old, human beings move from infancy into the phase of childhood; here our reliance on others is somewhat lessened, but still profound.[503] Beyond this, dependence continues into adolescence and in recent decades has been extended even further in the West: psy-

[497]Cf. Rosati, 2006, p, 35.

[498]On this trade-off, see Sagan, 1986 [1977], p. 3.

[499]Skinner, 1965 [1953], pp. 54-55. Often learning takes the form not of acquiring an entirely new behaviour, but coming through experience to deploy a hard-wired behaviour in an appropriate way and under appropriate conditions; cf. Panksepp, 1998, p. 25.

[500]Dennett, 2017, pp. 88-89.

[501]See O'Dea, Noble, and Nakagawa, 2022 for an approach to modelling these individual differences.

[502]There is a rich literature on 'social learning theory' which I cannot adequately summarise here; see especially the now-classic work of Albert Bandura.

[503]See Bogin, 1997 for some insights into the biological origins of this distinct developmental stage.

chologists have noted that on average young adults may now delay their development by up to ten years.[504] Many citizens of advanced Western nations today are still in full-time education and at least financially dependent at 21 – an age which early humans would be lucky to live to, and well beyond the entire life cycle of many other mammals such as cats and dogs. Our impressive capacity for learning continues to be important even later in life, too; and more generally, the retention of juvenile features throughout the lifetime will be a recurring theme of the latter part of the chapter.[505]

A third core characteristic of our species also lies at the heart of the subject matter under investigation: our capacity for rational, cognitively-guided behaviour. Relative to our overall size, the human brain dwarfs that of any other animal – and indeed is so large that human babies must be born with their skull in two pieces to allow further cortical expansion before later fusing into a single structure.[506] Yet despite having this impressive hardware from birth, our rational powers reach their full height only after the frontal cortex – known as the 'executive reasoning module' – comes online, which does not happen until around our mid-twenties.[507]

As with our condition of initial dependence, it has been suggested that our long route to maturity here is not merely a mistake or defect, but plays an important role in our later competence. Because of its delayed maturation, the frontal cortex has an opportunity to be shaped to a greater extent by learning rather than genetics:

> probably the most important fact about genetics and culture is the delayed maturation of the frontal cortex – the genetic programming for the young frontal cortex to be freer from genes than other brain regions, to be sculpted instead by environment, to sop up cultural norms.[508]

Although the behavioural sophistication resulting from our learning and cognitive abilities is a great strength, these also entail that the environmental prerequisites for adequate development are much more demanding

[504] See Sheehy, 1996 for a popular discussion.
[505] See especially Section 3.5 below.
[506] See Piantadosi and Kidd, 2016, p. 6874 for a discussion of the mutually-reinforcing nature of brain size and intelligence.
[507] Sapolsky, 2017, pp. 154-155; cf. *ibid.*, pp. 165-169 on adolescents, and Aristotle, *NE* I.3, 1095a2: 'Hence a young man is not a proper hearer of lectures on political science'.
[508] Sapolsky, 2017, pp. 326-327; cf. *ibid.*, pp. 171-173.

compared to simpler species.[509] Consequently, there is a wider range of ways in which things can go wrong for human beings during maturation.[510]

This point relates to a fourth central feature of our species: our inherently social nature. Though this is a large topic in general, the key claim focused on in this chapter is that some of the prerequisites for our proper development are social: that following on from our behavioural and cognitive plasticity during infancy, the emergence of our powers of practical reasoning is heavily mediated by our surroundings and culture, and is conditional on certain sorts of formative relationships with others.[511] Moreover, because these core features of our species – our initial radical dependence, our capacity for learning, our rationality, and our intensely social nature – continue later into life, a particular sort of social environment is necessary for maintaining these powers over time too. We are now finally ready to present our central line of reasoning.

The argument against Subjectivism will have the structure of a transcendental argument, in the style pioneered by Kant.[512] As is well-known, Kant deploys such arguments in the first *Critique* in attempting to determine which categories of the understanding are necessary for us to experience the *phenomenal* world as we do, and what features the *noumenal* world must logically have in order to causally give rise to these same experiences.[513] Yet Kant also uses the same strategy in the context of uncovering what must be presupposed in order to understand ourselves as agents capable of rational action.[514] The core of the present argument will have a similar focus – though the goods I will identify as prerequisites

[509] The meaning of the term 'adequate' here will be made clear below; for a classic criticism of 'normal development' as a normative ideal, see Hull, 1986, Section 2.

[510] Bowlby, 1997 [1969], pp. 129-130; MacIntyre, 2009b [1999], p. 68; p. 72; MacIntyre, 2016, p. 39.

[511] Cf. MacIntyre, 2012: 'To be a human individual is to have the potentialities of a rational animal ... those potentialities are for the most part actualized in and through that individual's relationships with others'.

[512] Thus avoiding appeal to shared 'intuitions', which I have argued is problematic in this normative context – see Section 1.3 above.

[513] Kant, *Critique of Pure Reason*, B106; B116-169; B274-279; see also Stern, 2019 on transcendental arguments in general, Pereboom, 2018 on Kant in particular, and Korsgaard, 1996 for a transcendental argument inspired by Kant, concluding that as rational agents we must value our own humanity.

[514] Kant, *Groundwork*, 4:446-463; see Williams, 2006 [1985], pp. 56-70 for a critical perspective suggesting that 'freedom' is the only good that is essential to rational agency, whereas Kant draws up a rather more expansive list; and see also Rawls, who suggests the 'primary goods' of 'liberty, opportunity, wealth and self-respect' to be 'in general necessary for the framing and the execution of a rational plan of life'; Rawls, 1971, p. 433.

to practical agency are not intended to be logically necessary, formal requirements of the sort that Kant took himself to have found. Rather, I shall instead claim that they meet certain specifically human needs that stem from a correlate empirical account – an account once general and partial and incomplete – of how human beings in particular develop into rational agents.[515] This identification will consist of an argumentative pattern that will be instantiated several times throughout the chapter, within different social contexts of practice. It has four parts, which I will now outline.

The first part of the transcendental argument pattern consists in elaborating a complex capability that I shall denote 'Practical Reason', which has cognitive and practical aspects. As a rough first approximation, I will use this phrase to mean both the acquisition of an articulate, reason-responsive conception of what is good for us that is suitable for use in the contexts of Action Selection and Life Planning, and the ability to successfully implement this ideal in our lives in practice.

As we shall see, the development of both of these abilities is gradual, and at first they come into being only in partial form. And because foresight is slower to develop, our emerging conception of what is good for us will at first relate only to the here-and-now perspective that informs Action Selection, before later developing into a partial specification for how to move forward with our lives as a whole in the context of Life Planning.[516] These points are important to note because over the following three sections we will be interested in what attitude towards their own good an agent whose powers are still inchoate must adopt if these are to be brought to full fruition. Moreover, later in life our rational powers may go into decline or be lost entirely; managing this ongoing vulnerability will be the central theme of the final two sections the chapter.[517]

This minimal conception of Practical Reason will be further expanded presently; in particular, we shall focus on the specific form this capacity

---

[515]On this teleological conception of 'needs', see Griffin, 1986, p. 42: 'Statements of need are of the form: x needs a in order to *ø'*. This entails that the final argument stands or falls with the empirical claims presented in this section; I have therefore tried to restrict myself to claims for which there is robust evidence issuing from a number of empirical perspectives.

[516]MacIntyre, 2009b [1999], p. 74-75; on the development of planning, see also Bowlby, 1997 [1969], pp. 350-351.

[517]For discussion of this vulnerability, see MacIntyre, 2009b [1999], Chapter 1; p. 73; and Snead, 2020:, p. 88: 'because all human beings exist as corruptible bodies, periods of serious illness, injury, and senescence create cycles of often-profound dependency throughout the life span for everyone'.

must take within the context of our two Subjectivist approaches, bearing in mind the discussion in Chapter 2. But since the argument will focus on what is needed to acquire these powers, first we must ask: Why should we value Practical Reason in this sense?[518]

Some human beings indeed never attain these capacities, and thus do not have rational control over their actions or the direction their lives take. However, here we must consider the *telos* of the project we are engaged in.[519] For organisms that are not rational agents in this sense, the deliberative contexts of Action Selection and Life Planning do not even arise; if an agent cannot choose what to do with their lives, then no Action-Guiding theory will have anything to offer them. Yet the same of course goes for the Subjectivist accounts discussed in the previous chapter. Practical Reason is a necessary prerequisite for an agent to make use of the results of the enquiry attempted here; otherwise, not only the central questions I have raised but also the skeptic's question of why we should value Practical Reason itself cannot concern them.[520]

The assumption that our powers of Practical Reason are an essential good that we must strive to achieve and maintain, then, is something *any* Action-Guiding theory must be committed to from the outset. For if Practical Reason is lacking, the goals of these theories cannot be achieved, since they are intended to apply in contexts of choice and action that presuppose them. And the force of this commitment remains in place regardless of whether or not Practical Reason turns out to be an 'objective good' in the superlative, metaphysical sense. In terms of practice, for the present, critical argument we may therefore likewise assume that the agents (or would-be agents) we consider must indeed strive towards gaining and maintaining these powers of Practical Reason.[521] But as this practical issue is our sole concern here, the theoretical issue of objectivity is moot – and we thereby also avoid the need to take a stand on a notoriously thorny philosophical issue.

On this minimal characterisation, the value of Practical Reason is thus a shared commitment of any Action-Guiding account of the Human Good. However, a more detailed description of exactly what is required

[518]For criticism of constructivism in ethics which presses this line of thought, see Enoch, 2006; Tiffany, 2012; and Leffler, 2019.

[519]Cf. Section 1.3 above.

[520]See Velleman, 2009, Chapter 5 for a related line of thought; see also Aristotle, *NE* I.3, 1095a10-11: 'to those [and only those] who desire and act in accordance with a rational principle knowledge about such matters will be of great benefit.'

[521]Cf. Section 1.3 above on the *reductio ad absurdum* nature of the project.

here will depend on the specific theory we are considering and its particular demands. For instance, any Aristotelian conception of the Human Good requires for its full implementation an agent who has developed the virtue of *phronésis*: a demanding capacity for practical wisdom that in practice may be achieved by only a few exceptional moral agents.[522] This in turn requires for its operation both extensive practical knowledge and other character traits as support, and has complex constituent components such as *nous* (intuitive perceptiveness of particulars), *gnōmē* (sympathetic judgement) and *sunesis* (understanding).[523]

For the Subjectivist views we will consider, the detailed specification of Practical Reason needed is a rather less demanding one; in essence, only the development of adequate Subjective Values and the powers needed to mobilise these for expression in our actions. I now give a list of five particular constituent powers, on which the official core argument will focus. Again, some of these are broadly cognitive, and others broadly practical.

1. *Conception of Self*
   To put our Subjectivist views into practice, an agent first requires an understanding of who they are. Within Action Selection, this is the self whose Subjective Values our choices are to express; for Life Planning, it supplies a sense of the boundaries of their particular life as a whole.

2. *Coherent Preferences*
   Our initially plastic and indeterminate Subjective Values must develop into a stable, coherent, and realistic specification of what is best for us and what our lives should be like, which can then be rationally implemented. We must also have access to their contents, yielding explicit premisses for our practical reasoning.

3. *Language Skills*
   In this project, we are considering *articulate* conceptions of the Human Good that are put to use in practical deliberation about action that is propositional in nature, as agents make explicit rational judgements about what to do in particular situations and concerning their lives as a whole. To do this, agents must first

[522]MacIntyre, 2011 [2007], p. 180.
[523]See Dunne, 1993, pp. 277-279; cf. *ibid.*, Section 9.6 on *nous*.

acquire the underlying concepts and linguistic capacities necessary for reasons to be framed and adequately reflected on.

4. *Means-Ends Reasoning*
   In order to successfully incarnate our Subjective Values in the world, we must also first acquire the causal knowledge of our physical and social environment necessary to connect particular proximate actions to the achievement of our ultimate goals – which in the context of Life Planning may occur only at a great temporal distance.

5. *Practical Judgement*
   In the previous chapter, we found that no mechanical, algorithmic implementation of either of our Subjectivist Action-Guiding theories was forthcoming. Practical Reasoning in the Subjectivist mode therefore requires quite extensive powers of practical judgement: minimally, the ability to adjudicate the rival claims of competing desires or of different forms of pleasure, thus rationally choosing between them.

The second part of the argument pattern consists in detailing what is required for us to either develop these five components of Practical Reason, or to sustain them later on in life. As I have said, what I shall focus on in the former context, drawing on developmental and social psychology, is how the achievement of these abilities depends heavily upon our making good use of the contributions of supportive others.[524] This empirical literature thus reveals the inadequacy of any simple 'maturation' theory here, wherein Practical Reason develops automatically and inevitably, requiring merely the passage of time.[525] Rather, for our rational powers to emerge at all, our initially flexible and indeterminate behavioural predispositions must be shaped through a process of socialisation that requires us to actively participate in certain formative relationships at different stages in our development:

[524]Cf. Snead, 2020, p. 91: 'Charles Taylor noted that even the traits required for thriving under the ambit of expressive individualism depend on social structures and conditions that nurture the development of such capacities.' An important type of source here is studies of 'feral' children or those raised in conditions of extreme social deprivation; see here Tartter, 1998, pp. 104-111, and see Davis, 1940 for a classic discussion of the consequences of isolation.

[525]Indeed, a simple maturation view has even been shown to be too simple even in the more prosaic case of motor development; see Thelen, 1999, p. 103; Johnson, Slater, and Hocking, 2017, p. 50.

> The inchoate and scattered impulses of an infant do not co-ordinate into serviceable powers except through social dependencies and companionships. His impulses are merely starting points for assimilation of the knowledge and skill of the more matured beings upon whom he depends. They are tentacles sent out to gather that nutrition from customs which will in time render the infant capable of independent action. They are agencies for transfer of existing social power into personal ability; they are means of reconstructive growth.[526]

In particular, I shall argue that the successful development of Practical Reason requires the emerging rational agent to recognise what I shall call 'Relationship Goods': that is, goods only made available through participation in certain relationships with others, such as our parents, teachers, and friends.[527] The Relationship Goods I will discuss will be of three broad types. Firstly, there are those goods which are strictly external to our relationships but which we can nevertheless only come to recognise through being open to the influence and guidance of these others. Secondly, there are those goods which are constituted by the intellectual and practical development that is achieved through following the direction of these others. Finally, the third class of Relationship Goods I will appeal to is simply these ongoing relationships themselves.[528]

In the final two sections of this chapter, I will complete the case against Subjectivism by further arguing that beyond this initial development there are also certain social prerequisites for maintaining our rational abilities over time, from when they are achieved in late adolescence onwards.[529] This will enable us to extend our conclusions about the importance of Relationship Goods beyond the context of infancy, childhood, and adolescence to human life as such, and thereby provide a reply to those Subjectivist theorists who might respond to the initial

[526]Dewey, 2002 [1922], pp. 93-94. The present account is also supported by the classic research of Vygotsky, whose theory treats 'higher cognitive structures as coming from the social world, becoming internalised as a result of interactions with knowledgeable others' – Johnson, Slater, and Hocking, 2017.

[527]For some ideas from Rawlsian political philosophy on the importance of interpersonal relationships for our capacity to develop conceptions of our own good, see Cordelli, 2015; and Brake, 2017.

[528]For an overview of the philosophical literature on the goods of relationships and their significance, see Gheaus, 2018.

[529]See Sections 3.5-3.6 below; this theme is explored within the conditions of isolation depicted in Defoe's classic novel, *Robinson Crusoe* – whose title character economists often use to illustrate the theory of choice and production.

argument by being careful to only advance their claims in the context of already-developed adult reasoning.[530]

At this stage, we must also anticipate another line of objection that could be raised in defence of Subjectivism here. This retort would accept that we must indeed acknowledge such Relationship Goods, on pain of stifling our rational powers – but also maintain that this acceptance is nevertheless compatible with pure Subjectivism because these can in fact be seen as *Instrumental Goods* that merely afford us a means to satisfy our Subjective Values – whether the theory is applied at the level of Action Selection, or else retreats to the broader perspective of Life Planning.[531] And if the later option is taken, it may also be denied that such goods involve an Internal Means to the best kind of life: for instance, it may be argued that the concept of a pleasurable life can be adequately specified without them.[532]

This objection will occupy us a great deal more, and brings us to the third part of the argument. What I shall argue in response is that due to certain contingent facts about the nature of human development and vulnerability, fully engaging in these relationships in the requisite manner will, in fact, require us to value certain Relationship Goods Intrinsically rather than merely Instrumentally. For instance, we will see that at a sufficiently young age, approaching our most formative relationships with a purely Instrumental attitude is not even a coherent possibility. For the empirical evidence shows that such relationships are necessary even for a Conception of Self to develop at all, and further that infants also come to recognise commitments to their caregivers *prior* to being capable of the sophisticated Means-Ends Reasoning that this Subjectivist response entails.[533] Moreover, later in life, the state of holding a purely Instrumental attitude toward others is in various ways treated as pathological by psychotherapists: the concept of 'Ego-syntonic process', for instance, denotes a state wherein 'a person is so afraid of social relationships that he or she consistently perceives all issues in terms of their payoff for self'.[534] And we shall also see that such mental health conditions may undermine our rationality too.[535]

[530]Cf. Section 3.5 for extended discussion of this response to the argument of the first part of the chapter.
[531]A view expressed by Aristippus; see Section 2.3.
[532]See Section 1.7 for discussion of 'Internal Means.'
[533]See the following two sections, respectively.
[534]McLeod, 2009, p. 202.
[535]On mental health problems caused by relationship deprivation undermining the

Lastly, the fourth and final component of the argumentative pattern concludes that our Subjectivist accounts are inadequate to the practical tasks of Action Guidance we have set for them.[536] The observations made within the second part of the argument reveal the importance of formative and supportive relationships for development and beyond, and the third part emphasises that to discharge this role they must be valuable Intrinsically. Yet we have previously seen that a purely Subjectivist response to our core problem requires the agent to instead hold a merely Instrumental attitude towards their social relationships – if not directly at the level of Action Selection, then at least within their reasoning at the broader level of Life Planning.[537] Our central need to learn from and Intrinsically Value the input of others is therefore something that our Subjectivist views must systematically fail to accommodate, and its individualistic commitments thereby make the theory unworkable in practice.

To summarise: we have seen that our powers of Practical Reason are necessary for benefitting from any Action-Guiding account of the Human Good, and for even speaking of 'reasons for action' at all. In particular, the five constitutive elements laid out above are necessary requirements for implementing our Subjectivist views in practice, and are therefore achievements that any Subjectivist theory must recognise as at least Instrumentally Good. However, attending to the development of Practical Reason from infancy onwards – rather than only its finished expression in adulthood – will show that proper participation in the types of formative relationships we will discuss is prerequisite to the achievement of these same powers. Moreover, consideration of the social context of Practical Reason in adulthood, which supplies resources that must be present if these powers are not to atrophy and decline, will establish a similar conclusion across the lifetime. Yet to engage in these key relationships in the required way, the Relationship Goods they generate must also be Intrinsically Valued, rather than being merely Instrumentally Valuable as a means to desire-fulfilment or pleasure.

The chapter will proceed by focusing on several key periods of human life, and each corresponding discussion will instantiate this four-part

conditions needed for autonomy, see Brownlee, 2013; on individualistic approaches to relationships with others being ultimately self-defeating, see Taylor, 1992, p. 35, and p. 52: 'it would seem that having merely instrumental relationships is to act in a self-stultifying way.'

[536] As per the Normative-Constructivist framework developed in Chapter 1; see especially Section 1.3 above.

[537] See again Section 1.7 above.

argumentative structure within these different social contexts.[538] We begin our sketch of our route out of our initial position of extreme dependence in the next section, concerning infancy – a time when we remain largely unable to fend for ourselves, and our reliance on others is obvious and overwhelming.[539] But first, one last *caveat* should be made, in order to avoid a potential misunderstanding.

Before embarking on our investigation, it should be noted that my position is not that we should *never* pay heed to our Subjective Values, always listening only to others rather than consulting our inner inclinations and feelings.[540] Mill was not entirely wrong: oftentimes, it is indeed us who are the most suitable judge of what is best for us, given our intimate – if sometimes biased – knowledge of the details of our lives and our capacities and limitations, and our deeper understanding of what coherent long-term goals we have formed.[541] In reining in the excesses of individualism, then, we must be careful not to under-emphasise our personal contribution either. There are times when we do best by going our own way, as well as particular occasions where advice offered to us is not to be trusted – indeed, there are perhaps some people we should never listen to at all. As adults, if not before, we should therefore accept outside influence only in a considered, discriminating way.[542]

What agents must do, then, is to find the right balance between the two poles of being guided by outside influence from others and following their internal Subjective Values – both in specific choices at the level of Action Selection, and in cultivating the implicit tendencies of their character at the broader level of Life Planning. This is similar to Aristotle's doctrine of the mean, where virtue is seen as an intermediate disposition lying between the vices of excess and deficiency; for instance, a brave person is one who experiences fear to an appropriate degree, and

[538]Though the chapter is not intended as a stage-based theory of development, in the sense of Piaget, Kohlberg, or Erikson, wherein stages cannot be 'skipped out' and a child must progress through each stage to engage in the processes of later stages. For instance, they will not in general occur in a strict linear temporal order, and due to space constraints the list leaves out other important considerations, such as romantic relationships.

[539]Though we might have begun even further back, since the social environment does not begin at birth: see Sapolsky, 2017, p. 210-211 for an overview of prenatal influence on development.

[540]On the need for independence of mind, see MacIntyre, 2009b [1999], p. 96.

[541]Mill, 1989 [1859], p. 67, quoted in Section 1.3 above; having our wants consistently frustrated or being exposed to chronic pain may also have debilitating effects on our powers of practical agency.

[542]Cf. MacIntyre, 2016, p. 75, and the potentially critical attitude to received values proffered in Section 1.3 above.

as such is intermediate between the rash person who fears too little and the coward for whom fear is overwhelming.[543] Moreover, to echo another Aristotelian position, when others are listened to this must be done in the right way, towards the right person, at the right time, and to the right extent.[544]

Although any positive Action-Guiding account would be incomplete without addressing these questions in much greater detail, I shall not pursue these matters further in the present work, since here I am aiming only at a negative critique of our two Subjectivist views. Finally, in closing, it is worth also pointing out that even our personal values themselves ultimately stem in large part from what we have picked up from our social context.[545]

3.2 The Long Childhood

In this section, we begin tracing out the path by which human beings acquire our capacity for Practical Reason by considering infancy – a time when our dependence on caregivers is virtually complete; and infant-parent relationships – which I shall argue are critical for our future development.[546] Our primary focus will be the development of our first two background prerequisites for Practical Reasoning in the Subjectivist mode: a Conception of Self, and Coherent Preferences. We will ask what the infant's attitude towards their parents and mode of interaction with them must be like if these are to emerge.[547]

What guidance can our Subjectivist views offer here? Earlier, we saw

[543]Aristotle, *NE*, II.6, 1106a26–b28; see also 1106a36–b7 on the relativity of the mean to the individual's particular context.

[544]Cf. Aristotle, *NE*, III.7, 1115b15-19.

[545]Cf. Watts, 1966: 'We seldom realize ... that our most private thoughts and emotions are not actually our own. For we think in terms of languages and images which we did not invent, but which were given to us by our society.'

[546]For simplicity, I consider the familiar case where an infant is raised by their biological parents, focusing primarily on the mother – though I expect the remarks to apply to surrogates. Indeed, I was not raised by my biological parents, and in some indigenous tribes, an infant is raised by many caregivers; see Bowlby, 1997 [1969], pp. 303-306; McHale, 2007 on co-parenting and the impact of broader family 'relationship systems'; and Sagi *et al.*, 1994 on children living in a Kibbutz. See also MacIntyre, 2009b [1999], p. 101 on the normative importance of the typical case despite exceptions.

[547]I thus set aside the interesting topic of what values guide good parents as they oversee a child's development: whether their actions are motivated by the baby's or their own Subjective Values, or some other source – though on this, see Scanlon, 1993, pp. 194-195.

that if the satisfaction of our Subjective Values is all that is Intrinsically Good for us, then we must ultimately view all of our relationships in Instrumental terms.[548] And even if this holds only at the level of Life Planning, rather than at the direct level of Action Selection, it nevertheless entails that any commitments to others that we accept as binding must be revisable as these inclinations dictate.[549] Moreover, this Instrumental attitude to others must apply here, within a family setting, as much as it does elsewhere; and the resultant conception of the family is thus merely an arena for self-interested strategic interactions. For although we may form contingent alliances whereby the good of others is respected within our here-and-now reasoning, our valuing of these must remain conditional on the advancement of our own Subjectively-determined ends, so that they must always remain provisional and at least potentially antagonistic.

Although infants are not yet capable of rational action and of only limited behaviour, if our Subjectivist accounts are to have anything to offer here, they can only insist that their behaviour should be guided in the same manner. However, I shall argue that this conception of infants' engagement in family life is inadequate if their development is to proceed satisfactorily. If infants were to behave as though their existing Subjective Values are constitutive of their own good, then this would prevent them from being sufficiently receptive to the transforming influence of socialisation, and they would be unable to internalise what they need to learn in order to develop their powers. Rather, *contra* the Subjectivist position, infants must treat the Relationship Goods available through engagement with their parents as having Intrinsic rather than merely Instrumental Value.[550]

We begin with some empirical insights into infant-parent relationships. Any adequate study of these matters must now reckon with Bowlby's work on attachment.[551] One of the most enduring contributions to the developmental psychology literature, the core of Bowlby's

[548] Cf. Sections 1.7 and 2.4 above.

[549] See again Section 2.4; see also Taylor, 1992, p. 43.

[550] See Brighouse and Swift, 2014 for more on the Relationship Goods that arise in supportive family contexts, suggestions for how these can be accommodated within a broadly liberal point of view, and especially Chapter 3 on their importance for child development. In the case of very poor parenting, development may be thwarted until an acceptable substitute is found.

[551] For the original papers, see Bowlby, 1958; Bowlby, 1959; and Bowlby, 1960; for a more recent overview of attachment and its developmental consequences, see Prior and Glaser, 2006.

attachment theory is still considered valid today, many decades after its proposal.[552] Important research themes here have included the stages of attachment,[553] multiple attachments and the role of the father,[554] the critical period for attachment, attachment in other animals,[555] the affective side of attachment and its basis in the nervous and endocrine systems,[556] cultural variation in attachment,[557] and types of attachment, maternal deprivation, and the influence of both of these on later development.[558] The latter will be especially important for our purposes, but first we must elaborate the concept of attachment itself.

At the most basic level, attachment behaviour is merely the tendency of the young to maintain physical proximity to a particular caregiver.[559]This overt behaviour is broadly continuous with what we find in other species – such as imprinting in the greylag goose, as famously studied by the ethologist Konrad Lorenz.[560] However, in his inaugural paper of 1958, Bowlby had distinguished five further patterns of behaviour that infants are innately predisposed to acquire: sucking, clinging, following, crying, and smiling.[561] Hence, attachment encompasses not only the disposition to maintain proximity itself, but also to produce these other behaviours that promote it.[562] Moreover, as the infant becomes more advanced, attachment behaviour comes to be 'goal-corrected' and hierarchically organised, rather than merely instinctive – for instance, following behaviour may employ a variety of novel means in pursuit of the fixed goal of maintaining proximity to the caregiver.[563]

Noting that attached infants are especially prone to proximity seeking when they perceive a threat or feel discomfort, Bowlby also drew on the ethology literature to argue that the key evolutionary function of attachment behaviour is to protect the infant from predators.[564] By the early 80s, Bowlby was therefore careful to separate off attachment

[552]Schaffer, 2003, p. 100; Waters and Cummings, 2000, p. 11.
[553]Following the pioneering work of Schaffer and Emerson, 1964.
[554]See e.g. Lamb, 1997.
[555]Following Lorenz, 1935 on geese; and Harlow and Zimmermann, 1959 on macaques.
[556]Feldman, 2012; Sapolsky, 1997, p. 135.
[557]E.g. Sagi *et al.*, 1994.
[558]Ainsworth *et al.*, 1978; Main and Solomon, 1986; van IJzendoorn, Schuengel, Bakermans-Kranenburg, 1999 on 'disorganised attachment' and its sequelae.
[559]Prior and Glaser, 2006, pp. 16-17; Schaffer and Emerson, 1964, p. 6.
[560]Lorenz, 1935; Bowlby, 1997 [1969], p. 223.
[561]Bowlby, 1958; cf. Bowlby, 1997 [1969], p. 180 for discussion.
[562]Ainsworth and Bell, 1970, p. 49-51.
[563]Bowlby, 1997 [1969], p. 180; p. 248.
[564]Ainsworth and Bell, 1970, p. 51; p. 64.

behaviour as motivated by reducing risk of harm and relieving anxiety from other forms of behaviour that also happen to involve maintaining proximity to a special figure – for instance, as a playmate, or as a bearer of shared interests.[565] This anti-predation hypothesis has been corroborated by studies of factors that tend to intensify attachment behaviour – such as when a child is alarmed, or unwell, or placed in an unfamiliar environment.[566]

Finally, another important aspect of attachment relationships is that they function as a 'safety net' that enables increasingly confident and independent exploration of the world in the knowledge that there is a secure base to return to if problems arise.[567] We will return to this idea later; next, we give a rough overview of the timing of the onset of attachment.

Attachment behaviour in any sense naturally requires the baby to recognise the mother via perceptual discrimination, which is possible from the second or third month onwards.[568] However, fully developed, goal-corrected attachment occurs rather later: though there is great individual variation, and some degree of cultural variation too, the purposive maintenance of proximity to a specific attachment figure is usually established only at some 6-7 months of age – much later than we find in other mammal species.[569]

As attachment relationships develop further *en route* to childhood, from the middle of the third year onwards they lead to the formation of what Bowlby has called a 'goal-corrected partnership', wherein 'By sharing a common set-goal and participating in a joint plan to achieve it, partners have a rewarding sense of common purpose'. This occasions a child's first experience of acting according to an articulate plan – albeit one which is shared with the mother.[570]

Attachment behaviour then reduces in intensity from the end of the third year of life onwards, which often coincides with the onset of formal schooling – though it may appear much later in a more muted form throughout our subsequent practical and cognitive development.[571]

---

[565] Bowlby, 1997 [1969], pp. 374-375 [introduced in the second edition of 1983]; cf. Bretherton, 1980, to whom Bowlby is responding here.

[566] Bowlby, 1997 [1983], p. 226; p. 256; p. 258-259.

[567] Waters and Cummings, 2000; cf. Bowlby, 1997 [1969], p 208; Ainsworth and Bell, 1970, p. 52.

[568] Schaffer and Emerson, 1964, p. 7.

[569] Bowlby, 1997 [1969], pp. 200-209; pp. 267-268; p. 183.

[570] Bowlby, 1997 [1969], p. 355.

[571] Bowlby, 1997 [1969], pp. 204-206.

This brings us to the importance of attachment for later developmental outcomes. Many empirical studies have supported Bowlby's theory of maternal deprivation, which claims that having at least one successful attachment relationship is necessary for normal emotional, cognitive, and social development. This literature is so richly suggestive that we briefly mention some of these findings before moving on to our main theme.[572]

Concerning emotional development, the onset of attachment behaviour has been associated with important changes in the limbic system that occur around 7-15 months; this is an evolutionarily-ancient brain system that is centrally involved in the emotions and their expression.[573] Moreover, successful attachment relationships have also been found to be necessary for the proper development of the orbitofrontal cortex – an area associated with the regulation of emotion – which again occurs over roughly the same time span.[574] Emotional regulation occurs first interactively – during exchanges with other humans, and later autonomously – in contexts when there is no one else present. Yet both kinds of regulation depend on the prior existence of a successful attachment relationship.[575] This observation led to further development in the concept of attachment, which now comes to be understood as partly defined by the shared regulation of emotion.[576]

There is also an extensive literature on the importance of attachment relationships for normal cognitive development, with many studies showing that the growing infant's intellectual powers are severely stunted in their absence.[577] Moreover, we just noted that the maturation of the orbital prefrontal cortex depends upon the formation of a 'goal-corrected partnership'; yet this brain region facilitates goal-directed behaviour in

[572]For a recent overview and evaluation of later developments following Bowlby's classic account here, see the meta-analysis of Groh *et al.*, 2017; and see also the now-classic Minnesota parent-child study; Sroufe, Coffino, and Carlson, 2010. The literature suggests that attachment is important even for physical development – see the influential Romanian orphan study: Rutter *et al.*, 1998.

[573]Schore, 2000, pp. 29-30; Sapolsky, 2017, p. 135; see Walker *et al.*, 2022 for a meta-study on attachment styles and emotional intelligence.

[574]Schore, 2000, p. 29-30; see also Chugani *et al.*, 2001 for a study on the effects of early deprivation in Romanian orphans.

[575]See Bowlby, 1997 [1969], pp. 363-364 on the importance of attachment for 'ego-control' and 'ego resilience'; Reindl *et al.*, 2018 on the importance of 'interactional syncrony' (discussed below) for infants' emotional self-regulation; and Sapolsky, 2017, p. 191 on the mother as teaching the regulation and proper deployment of pre-existing instinctive behaviour in the animal world.

[576]Schore, 2000; 1996; Sroufe, 1996, pp. 172-174.

[577]See e.g. Ding *et al.*, 2014 for a longitudinal study in China on attachment type and cognitive outcomes; see also Bowlby, 1997 [1969], pp. 223-228 on infants learning cognitive skills from the mother.

general, and hence is essential for the emergence of Life Planning.[578]

Finally, in lieu of systematic human experiments, due primarily to the presence of obvious ethical issues, early studies on baby monkeys have found that exposure to conditions of isolation early on in life leads to severe and lasting maladjustment.[579] This shows markedly the importance of attachment relationships for social development too – which according to the present argument is itself crucial in the development of rationality in human beings.[580]

The importance of attachment relationships for these broad developmental outcomes suggests that they are also necessary for a capacity for Practical Reasoning to appear. I will now argue that this is the case for our first two prerequisites for Practical Reason in the Subjectivist sense; hence, the role of others in our early lives is far more profound than merely facilitating the satisfaction of our pre-existing Subjective Values.[581]

Although it may sometimes be claimed that babies are egoists, at birth this can only be true in a limited sense at most, since a requisite component of the Subjectivist choice architecture is missing at this early stage in life. For initially, the infant lacks a Conception of Self as separated off from the rest of the world: 'Up to the third quarter of the first year, Piaget has shown, a state of "adualism" exists, in which there is no distinction between the self and the environment.'[582] In theorising about how this initially-absent Conception of Self develops, theorists from a wide variety of perspectives and disciplines have concurred on the overriding importance of social interaction – albeit in rather different ways.

Within developmental psychology, it has been argued that an infant's Conception of Self first appears through an awareness of the mother and baby as a combined unit during attachment. Thus the paediatrician

[578]See Tremblay and Schultz, 1999 for discussion and a neurological study of reward discrimination in macaque monkeys.

[579]Griffin and Harlow, 1966; one monkey in the study suffered such severe depression that he ceased to eat and died several days later. On the importance of attachment for later human relationships, via specification of a paradigmatic 'internal working model', see Hazan and Shaver, 1987.

[580]See Bowlby, 1997 [1969], pp. 221-222 for discussion; see also Sroufe, 1996, p. 596 for some data on the relationship between social competence and attachment style.

[581]Cf. again MacIntyre, 2009b [1999] for extended discussion.

[582]Schaffer and Emerson, 1964, p. 8; for a later discussion of the issue, see Kagan, 1998, who dates the emergence of infants' Conception of Self to some 18-24 months of age, based on behaviour such as the use of self-referential vocabulary, the conscious imitation of adult models, and acknowledgement of failure to meet expectations.

and psychoanalyst Donald Winnicott writes that 'there is no such thing as a baby ... a baby cannot exist alone, but is essentially part of a relationship.'[583] Moreover, this process of fully individuating ourselves from the mother may take a surprisingly long time – and even once established, inconsistencies or relapses may be present later on: 'A lack of ego boundaries is shown when a toddler isn't all that solid on where he ends and Mommy starts – she's cut her finger, and he claims his finger hurts.'[584]

Within psychoanalysis, Melanie Klein's 'object-relations theory' similarly holds that a Conception of Self is born in a distressing emergence from an original symbiotic relationship with the mother, as the infant moves through a 'depressive position' wherein they first become able to recognise themselves and others as separate entities.[585]

In the behaviourist tradition in psychology, a key research theme is understanding how operant conditioning enables us to acquire the power to discriminate between different stimuli.[586] Importantly, this includes the process of coming to recognise our own bodies and behaviours as different from the rest of the world – and in this case, the conditioning in question has been argued to be primarily social. Thus, on this view, we come to an awareness of ourselves through recognising the consequences of our own sensory-motor initiatives, and these contingencies are arranged by the background culture in which we develop as agents:

> self-observation is also the product of discriminative contingencies, and if a discrimination cannot be forced by the community, it may never arise. Strangely enough, it is the community which teaches the individual to "know himself."[587]

In sociology, G. H. Mead has propounded the influential view that our Conception of Self is the consequence of organised, co-operative interac-

[583] Winnicott, 1964, p. 88; cf. McLeod, 2009, p. 92; see also the classic work by Vygotsky noted in the previous section, Sroufe, 1996, pp. 587-588; and Bowlby, 1997 [1969], p. 349: 'the behavioural organisation of a child of [twelve months] is much less stable than is that of the couple of which he is a partner'.

[584] Sapolsky, 2017, p. 177.

[585] Klein, 1946; see also Kohut, 1972.

[586] Skinner, 1965 [1953], p. 66; pp. 107-109; see also Sapolsky, 2017, p. 42.

[587] Skinner, 1965 [1953], pp. 260-261. Compare: 'Humans, "possessing language" and culture, *do* arrange (verbal) discriminative contingencies that produce discriminations of one's own body and one's own behavior ... Consciousness, in the sense of "self-aware" is, therefore, of social origin and depends on the emergence of verbal behavior and culture. Only a full-blown culture (i.e., not one that consists of units involving only imitation) can arrange the verbal contingencies that produce a self-observational repertoire.' – Glen Sizemore [personal correspondence].

tions with others, and especially of taking on social roles that allow us to view ourselves from an external perspective:

> The self is something which has a development; it is not initially there, at birth, but arises in the process of social experience and activity, that is, develops in the given individual as a result of his relations to that process as a whole and to other individuals within that process.[588]

In continental philosophy, meanwhile, Habermas follows Hegel by expressing a similar view with his 'theory of communicative action':

> individuation can be comprehended only as a process of socialization. The moral subject, the subject of praxis, is inconceivable in abstraction from communicative relations with others.[589]

Finally, MacIntyre's Thomist approach also argues that our social context is always the starting point for forming our individual identities, and what initially 'gives my life its own moral particularity'; moreover, on this view, the self also continues to be in part constituted through our ongoing relationships with others and the social roles that we thereby come to inhabit.[590]

Our Conceptions of Self are thus formed socially: we become aware of who we are in large part through our relationships with others. So, if they are to offer viable guidance for the behaviour of pre-rational infants *en route* to achieving Practical Reason, our Subjectivist theories must acknowledge the critical need for infants to orient themselves toward and engage in these formative relationships.[591] They must, therefore, implicitly recognise them as at least Instrumentally Valuable.

---

[588]Mead, 1934, p. 135; p. 138; see also *ibid.*, Chapters 18-19; and Taylor, 1992, p. 33: 'The genesis of the human mind is in this sense not "monological," not something each accomplishes on his or her own, but dialogical.'

[589]Habermas; quoted in Dunne, 1993, p. 176.

[590]MacIntyre, 2011 [2007], pp. 254-256; see also MacIntyre, 2009b [1999], pp. 89-90; pp. 94-95; and MacIntyre, 2006a, p. 83: 'our understanding of ourselves is open to correction by what we learn about ourselves from others'. For further concurrence within contemporary philosophy, see also Held, 1993, p. 70-71: 'It is mothers and mothering persons and children who turn biological entities into human social entities through their interactions ... New persons are created within families'; and see also Mead, 1934, p. 164: 'No hard-and-fast line can be drawn between our own selves and the selves of others, since our own selves exist and enter as such into our experience only in so far as the selves of others exist and enter as such into our experience also.'

[591]See Schaffer and Emerson, 1964, p. 63 on the active role of the infant.

Another aspect of attachment theory pertinent to our enquiry is the aforementioned idea that the attached infant uses the mother as a base for exploration that is returned to when a significant stressor arises. Bowlby reports on a classic study of indigenous peoples:

> In her study of Ganda infants Ainsworth (1967) notes how, soon after an infant is able to crawl, he does not always remain close to his mother. On the contrary, he makes little excursions away from her, exploring other objects and people and, if allowed to do so, he may even go out of her sight. From time to time, however, he returns to her, as though to assure himself she is still there. Such confident exploration comes to an abrupt end if either of two conditions occurs: (a) if the child is frightened or hurt; (b) if the mother moves away. Then he returns to her as quickly as possible with greater or less signs of distress, or else cries helplessly.[592]

This again suggests that the infant alone does not constitute a complete organic unity that exists separately from its parents – but I shall instead put this observation to use in the context of our second component of Practical Reason: developing Coherent Preferences.

The tendency to return to a parent when a challenging situation arises – at first instinctive, but later increasingly purposive – also automatically effects the parent's taking over and handling any overly difficult, complex, or unfamiliar situation. Hence, this mechanism can function as a remedy to the general fact that infants are not yet ready to deal with the world by themselves. In some instances, this incapacity is due to the infant at first lacking an adequate understanding of the causal structure of the social and physical world; we return to this idea in the next section. But in other cases, the absence of a coherent, independent response here is instead a consequence of another phenomenon that remains common throughout youth and may arise even in adulthood: that the infant *does not understand the situation well enough to have a preference*.[593] Where Coherent Preferences over the outcome of an interaction are thus not forthcoming, the infant must thereby also lack the ability to formulate an intelligent response of their own; in retreating to the more experienced parent, then, the infant implicitly places their trust in their judgement

[592] Bowlby, 1997 [1969], p. 209; see also *ibid.*, p. 208; Ainsworth, 1967.

[593] Cf. Wittgenstein, 2003 [1953], p. 42^e on 'not knowing our way around'.

and guidance instead.[594]

In moving from this initial condition – where at the level of Action Selection our Subjective Values may not yield any conception of how to act in situations we do not yet understand, and do not add up to anything like a coherent overall Life Plan – the infant's social environment is again key. In particular, as well as functioning as both our protectors and guides in situations we cannot yet comprehend, it is our parents (or surrogate caregivers) that must 'serve as the primary formers of our preferences'.[595]

The power of parents in shaping the emerging Subjective Values of the growing infant is immense.[596] In some cases, there are obvious reasons why this receptivity to parental values is necessary for our practical success in the world. For instance, infants must learn from their parents to avoid – and hence to implicitly treat as bad – potential harms that they cannot yet recognise as harms: perhaps because engaging with them does not appear to be connected to pain in any obvious way – such as touching a plug socket, or playing with matches; things which the infant may instead feel a strong desire to do.[597] Conversely, in other cases, infants must come to acknowledge that certain items parents deem to be good are in fact so, though they cannot yet understand why: hence the painfulness of the injection does not count against giving vaccinating an infant, even if the infant's own Subjective Values are not yet developed enough to identify this as a good. And in still other cases, parents may shape our Subjective Values in more subtle ways, such as by opening us up to possibilities we would not have considered on our own, and hence for which we have hitherto lacked any determinate preference either.[598]

In claiming that such parental influence is necessary if the growing infant is to gain Coherent Preferences, I am not advocating that the

---

[594]For some philosophical thoughts on secure attachment relationships as necessary for trust in general, see Kirton, 2020.

[595]Rosati, 2006, p. 35; cf. *ibid.*, p. 48.

[596]Sapolsky, 2017, p. 202; see also e.g. Benton, 2004 on parental influence on food preferences; Bauer, Chytilová, and Pertold-Gebická, 2012 for an experimental study on parental influence upon whether children develop other-regarding social preferences, using the dictator game; and Aristotle, *NE* II.1, 1103b24-25 for a classic statement of the importance of parents for moral development. Conversely, see Harris, 1998 for an influential statement of the view that peer influence can outweigh that of parents.

[597]See here Winnicott's discussion of the experience of saying 'no' as a response to the common experience of infants wanting to do something that is bad for them in the home: Winnicott, 2002, pp. 112-113; see also Brighouse and Swift, 2014, p. 63; and MacIntyre, 2009b [1999], p. 70.

[598]Cf. MacIntyre, 2009b [1999], pp. 75-76; MacIntyre, 2016, p. 36.

growing infant treat their parents as absolutely correct all of the time. Indeed, beyond early infancy and into childhood, it may be that our parents have failed to bring us up adequately if we always only follow their direction and do not seek our own reasons to support our choices; a process which may itself require us to sometimes put in question what we have already learnt.[599] And later on, a certain distancing from the parental relationship may be necessary if we are to thus progress from our initial position of radical dependence and gain some measure of self-sufficiency – which is itself important for Practical Reason.[600] Moreover, even from a young age, we must also be given some room to express ourselves rather than remaining merely passive if Coherent Preferences are to ever emerge.[601] But nevertheless, this development cannot continue unaided, and our movement beyond our initial starting point is always dialogical. In particular, for Coherent Preferences to arise at all, the growing infant must first recognise the parents' values as a central source of knowledge about what is good and best for them, and have an openness to absorbing and internalising these values.[602] Hence, in addition to seeking and engaging in attachment relationships themselves, recognition of this further type of Relationship Good is also necessary for Practical Reason in the Subjectivist's sense to be possible.

We now consider an important Subjectivist response to the arguments of this section so far; one anticipated in the previous section. So far, we have seen that any Action-guiding theory that applies in this context must acknowledge infants' need to recognise certain Relationship Goods if they are to develop a Conception of Self and Coherent Values. However, in reply to this, the Subjectivist may claim that ultimately infants should only participate in these relationships *in order to* satisfy their Subjective Values, and hence maintain that the Relationship Goods we

[599]MacIntyre, 2009b [1999], p. 84; cf. MacIntyre, 2011 [2007], p. 256 on transcending the limitations of our given moral identity. See also Bowlby, 1997 [1969], p. 355; p. 369 on authoritarian parenting *versus* a functioning parent-child relationship wherein the child is recruited as a partner; and Kohlberg and Piaget on moving beyond merely doing something because the parent commands it – see here Duska and Whelan, 1977, p. 39; p. 53; p. 69.

[600]This process is called 'individuation' in psychoanalysis; cf. McLeod, 2009, p. 116. This may fail to occur in more subtle ways than the most obvious cases, such as profound disability or impairment; e.g., individuals remaining at home with overbearing parents and hence not really in control of their lives. Mature agents may also suffer relapses here; e.g., long-term hostages who must get used to making simple decisions after years of stagnation.

[601]On the influence of parenting styles here, see Sapolsky, 2017, pp. 202-203.

[602]On self-sufficiency as consisting in part in internalising the standards of parents, see again Rosati, 2006, pp. 35-36; 43-48; and especially p. 51.

have identified have only the status of being Instrumentally Valuable.

This Subjectivist objection fails because the cognitive architecture required to vindicate it is not yet in place.[603] In the present context, for the objection to be convincing, it would have to actually be incorporated into a viable Action-Guiding account of how infant behaviour should be governed in practice, rather than a mere *post hoc* philosophical interpretation intended only to save the theory.[604] But for this to be possible, even at the implicit, behavioural level we are currently considering, infants would still need to be able to first discriminate between situations in which engagement in attachment relationships with their parents is indeed a means to the satisfaction of their private Subjective Values and situations where this is not so, and then either engage or refrain from engaging with them accordingly. Yet making this distinction requires both a Conception of Self as separate from the mother and Coherent Preferences that the resulting outcomes are measured against – and as we have just seen, these two features of their psychology arise only later on. Indeed, in early infancy, when this Conception of Self as distinct from others is not yet in place, the goods that our behaviour is implicitly directed towards cannot even be considered as 'private properties' at all: rather, at this age at least, my implicit conception of my good must be 'one and the same as the good of those others with whom I am bound up in human community'.[605]

If they are to develop as rational agents at all, then, the only remaining option is that infants must Intrinsically Value attachment relationships and the shared Relationship Goods they generate after all. This counter-response will now be supplemented by digging deeper into what infants' relationships with their parents are actually like during their initial radical dependency.

Two key concepts that have been developed in the empirical literature here are interactional synchrony and reciprocity. 'Interactional synchrony' (or 'Parent–Infant Synchrony') has been defined as the 'tem-

[603] See Bowlby, 1997 [1969], p. 242 for the view that real human infants do value attachment relationships and the contact they afford for their own sake rather than as a means to some further end: 'For no other behavioural consequences, perhaps, are standards of appraisal in man more clear-cut from the start, or more environmentally stable. So stable indeed are they as a rule that for babies to love mothers and mothers to love babies is taken for granted as intrinsic to human nature.'

[604] Cf. Section 1.7 above.

[605] MacIntyre, 2011 [2007], p. 266; cf. Section 1.6 above on 'well-being as commodity'.

poral coordination of micro-level social behaviour'.[606] In interactional synchrony, then, we observe a high degree of similarity in the facial or bodily movements of the baby and mother.[607] Such behaviour has been demonstrated right from birth, and with previously-unseen behaviours that could not have been learned through their post-natal experiences; yet the variety of imitative responses that are possible also points away from an explanation in terms of hard-wired 'fixed action patterns'.[608]

With behaviour known as 'reciprocity', meanwhile, the infant and the adult respond to and elicit further replies from each other non-verbally in real time, taking turns to lead the interaction.[609] Here regular, patterned communication flows in both directions: the adult responds in a sensitive way to the infant's responses and needs, and the infant expresses these needs whilst mimicking the more advanced behaviour of the adult. Studies also suggest that infants are again not merely exhibiting conditioned responses, but engaging in a proactive way that exhibits for the first time a limited form of purposiveness; intentional behaviour is thus inherently socially constructed from the outset.[610]

In contrast to the Subjectivist approach, what we see in the twin concepts of interactional synchrony and reciprocity is a certain kind of closeness in the relationship between baby and caregiver wherein there is no psychological distance between them: emotions are shared, influence

[606] Feldman, 2007, p. 340; see *ibid.*, pp. 343-344 for discussion of its importance for a range of developmental outcomes, including self-regulation. For a recent discussion of current research on human behavioural synchrony in general, see Schirmer, Fairhurst, and Hoehl, 2021; for a summary of work on imitation in babies and its relation to communicative development, see Nadel and Butterworth, 1999.

[607] For a groundbreaking early study correlating facial and manual gestures between babies and adults, see Meltzoff and Moore, 1977. Matching of affective states may also occur; see again Feldman, 2007.

[608] See Meltzoff, and Moore, 1983, who demonstrate the effect with mouth opening and tongue protrusion in neonatal babies less than 3 days old, with babies responding differentially in kind when these gestures are produced by an experimenter, yet also write that 'There is little in the nature and organization of the response that tempts us to describe it as a classic fixed-action pattern that is released by the adult's display', favouring instead a hypothesis of 'active intermodal matching' involving an implicit awareness of the equivalence of seen and performed actions – see pp. 707-708.

[609] For a classic early study on reciprocity, see Jaffe, Stern, and Perry, 1973, who model the temporal pattern of infant-adult gaze behaviour using Markov chains; for a more recent study on the importance of both mother-child and father-child reciprocity for later developmental outcomes, see Feldman, Bamberger and Kanat-Maymon, 2013.

[610] See Murray and Trevarthen, 1986 for a classic study wherein 2-month-old infants react very differently to live and recorded maternal responses, suggesting intentionality on the part of the baby; see also MacIntyre, 2009b [1999], pp. 15-17 on interpersonal recognition and understanding beginning in mutually responsive pre-linguistic activity.

is accepted uncritically, and behaviour is determined jointly.[611] The game-theoretic conception of each partner engaging strategically with the other and seeking only to pursue their own individual interest is therefore entirely out of place here, and it is indeed only infants suffering from pathological conditions that do not engage in this immediate, engaged, and trusting manner at this early juncture.[612]

Crucially, any attempt to construct an equivalent of the 'internalist' view of reasons in the domain of pre-rational behaviour must therefore fail as a normative account of the infant's engagement here. For this must require that their emerging intentional behaviour is directed by 'motivational resources that are already to hand'.[613] But in thus insisting that these guiding principles lie within the infant themselves, merely waiting to be discovered by introspection, such a view would impose an epistemological barrier that becomes implausible once we note that in reality this behaviour can arise only through engagement with the parent:

> The pattern of interaction that gradually develops between an infant and his mother can be understood only as a resultant of the contributions of each, and especially of the way in which each in turn influences the behaviour of the other.[614]

In particular, rather than being led only by their internal Subjective Values, what is required for the infant's development into a rational agent is that they instead have a certain openness to the influence of these caregivers and receptivity towards their values. This openness and receptivity will also be a central theme going forward into the rest of the chapter.[615]

In this section, we have remained at the level of mere behaviour, and have not yet reached the stage of explicit rational deliberation to which our Subjectivist Action-Guiding theories were originally intended

[611] See Reindl *et al.*, 2018 for a recent discussion and fNIRS study of the underlying neural mechanisms, and further discussion of the importance of synchrony for the infants' growing to regulate their own emotional states.

[612] Treating others exclusively in a purely Instrumental mode is indeed recognised as a key trait of psychopaths; see Bjork, Chen, and Hommer, 2012 on the relation between psychopathy and differences in the neural circuity mediating instrumental reward.

[613] Wallace, 2020, Section 3; cf. Williams, 1979, and MacIntyre, 2016, p. 155.

[614] Bowlby, 1997 [1969], p. 204.

[615] For a recent philosophical discussion of the concept of openness, distinguishing openness as a mental state from openness as a deliberative practice, see Davis and Finlayson, 2021.

to apply. The case against them is therefore still far from complete. But it has been made plain that at this implicit level the infant's required mode of engagement with their parents cannot be made to fit a Subjectivist model. Instead, we have seen that attachment relationships at first constitute the whole world of the young infant, and are part of their very being from the outset; in being thus identity-forming, they cannot be of merely Instrumental Value.[616] Moreover, an Instrumental understanding of these relationships also fails to do justice to the degree of intimacy, openness, and genuine partnership required for the infant to progress towards independent rational agency.

Though the infant's Subjective Values may be important, then – particularly their social and affective inclinations towards the mother figure – we may conclude that they cannot be the ultimate guiding influence on their behaviour if they are to acquire the abilities necessary for Practical Reason. In particular, in order to facilitate their rational development, sometimes infants should simply listen to what their parents have to say – regardless of whether they perceive that doing so will lead to the satisfaction of an existing desire or bring them pleasure, either in the here-and-now or within their lives as a whole. In the next section, we extend these remarks about infancy to childhood by considering the further development of our powers through formal schooling.

3.3 Formal Education

In this section, we move beyond infancy to childhood and discuss the formal education that most maturing agents in our contemporary society will receive through attending primary school.[617]

We shall now assume that the child has a Conception of Self and Coherent Preferences in place, at least to some extent, and will focus on the acquisition of our next two components of Practical Reason in the Subjectivist sense. These are: the advanced Language Skills needed to frame and critically reflect on propositions concerning our own good, and the understanding of the social and physical world necessary for implementing these ideals through Means-Ends Reasoning.[618]

We have seen previously that a purely Subjectivist view of the Human

[616] Cf. Taylor, 1992, p. 52.
[617] Again, focusing on the typical, paradigmatic case – though this is not universal.
[618] See Section 3.1 above.

Good requires that we ultimately view our social relationships in a merely Instrumental light – including the infant's attitude toward their parents, if an attempt is made to apply them here.[619] And likewise, insofar as our Subjectivist theories have anything to offer by way of guidance in the present context, they must also recommend that children treat their teachers and the knowledge and intellectual skills they have to impart as merely a potential means to the satisfaction of their own Subjective Values – again, if not directly within Action Selection, at least at the level of Life Planning. Hence, they should follow the direction of their teachers only in cases where this is a successful strategy for achieving this end. However, I shall argue that this model of how their behaviour is to be guided is again inadequate if their powers are to develop. Children are hindered in their studies by the view that they invariably know best, and to make intellectual progress they must instead recognise that sometimes it is wiser to listen to more experienced others rather than only consulting their own inner feelings and inclinations.[620]

By the time their formal education commences, children are of course generally able to speak and to converse with each other and their teachers in a fairly extensive way. However, for developed Practical Reason in the Subjective sense, rather more than this is needed. For this involves fully-fledged rational agents drawing upon their Subjective Values to formulate explicit propositions about what is good for them, and then utilise these premisses in their deliberations about what to do – including about how to engage with others in their social sphere. And the Language Skills required for framing and reflecting on reasons for action in this way are rather advanced.[621] Moreover, we shall see that the development of our prowess here is again something that can only occur within a supportive social context, of which our formal education is an important part.[622]

In addition to these linguistic limitations, young children also lack the solid understanding of the causal structure of the physical and social world needed for the developed Means-Ends reasoning that would enable them to implement their preferences successfully. Consider here a moving

[619] Cf. again Taylor, 1992, p. 59.

[620] For some empirics on the relation between students' receptivity to services provided and drop-out rates within higher education, see Smith, 2004.

[621] Cf. Taylor, 1992, p. 33: 'We become full human agents, capable of understanding ourselves, and hence of defining an identify, through our acquisition of rich human languages of expression'. See MacIntyre, 2009b [1999], pp. 53-56 for an extended discussion; and see also MacIntyre, 2016, p. 26; and Dunne, 1993, pp. 139-140.

[622] Cf. Taylor, 1992, p. 33: 'No one acquires the languages needed for self-definition on their own.'

example of their shortcomings, collected by Bowlby:

> Inevitably, the earliest attempts that a child makes to change his partner's behaviour are primitive. Examples might be pulling and pushing, and such simple requests or commands as 'Come here' or 'Go away'. As he grows older, however, and it dawns upon him that his mother may have her own set-goals and, moreover, that his mother's set-goals may perhaps be changed, his behaviour becomes more sophisticated. Even so, the plans he makes may be sadly misconceived, owing to the inadequate working model he yet has of his mother. An example is a little boy, just short of two years, who, having been deprived of a knife by his mother, attempted to get her to return it by offering her his teddy bear.[623]

Again, it is a core function of education itself to foster the growth of these cognitive skills, enabling developed Practical Reasoning to emerge.[624] How, then, does this education proceed?

According to a view defended by Socrates in the *Meno*, all education is merely recollection. Anything we might learn at school is therefore just the retrieval of knowledge held in a previous life, and the role of the teacher can only be to draw out what is already present in some form.[625] This mythical view would attract few followers amongst education theorists today: disregarding its obvious circularity and questionable metaphysics, it is clearly inconsistent with what we now know about the brain and its development.[626] However, recent pedagogical theory has also cast doubt on the converse view of education as the simple transmission of knowledge from one mind to another. According to 'constructivism', a currently-influential paradigm with roots in the work of Piaget, students are not passive receivers of information but rather *construct* their emerging knowledge through the educative process, building on existing prior knowledge that they already possess.[627]

Due to its emphasis on the student's contributions, constructivism has also inspired practical pedagogic approaches whereby teachers play

[623] Bowlby, 1997 [1969], p. 281.

[624] For justification of the fostering of rationality as a key goal of education, see Siegel, 1995.

[625] Plato, *Meno*, 80e-82b.

[626] Cf. Sober, 1998.

[627] For a recent discussion, see Amineh and Asl, 2015; for an earlier exploration of some key themes, see Steffe and Gale, 1995.

only a facilitating role. Therefore, it may initially seem to fit well with Subjectivism as construed here. Yet although these reformed teaching strategies do stress the active role played by students, they nevertheless also recognise that their competencies can only develop in the specially-structured social context of a supportive educational environment.[628] And theorists have therefore again pointed out the poverty of any simple maturationist view, wherein intellectual development requires only the passage of time and needs no outside influence.[629] In particular, the modern conception of learning as an active, student-centred process still requires that children engage with their educators if their Language Skills and Means-Ends Reasoning abilities are to emerge.[630] For example, with Language Skills: if our linguistic powers are to reach their full height, we must inherit specific cognitive tools and modes of thought which are neither ingrained within us from the outset nor solely of our own making *ex nihilo*.[631] Hence, any adequate Action-Guiding Subjectivist theory must also recognise that these important linguistic and cognitive powers are in fact Relationship Goods that can only be gained through active participation in the educative process.[632]

In a move that is by now familiar, the Subjectivist may respond to these remarks by again appealing to a long-term strategy here, arguing that the child should value these Relationship Goods only *because* doing so is means to satisfying their Subjective Values; or, put in more explicitly Action-Guiding terms, only *if* engagement is a means to their satisfaction – either on a piece-meal basis at the level of Action Selection, or else in guiding their own implicit attitudes within Life Planning. For instance, if a child has a deeper, long-term desire for material success in life, then

[628] For criticism of these 'hands-off' approaches to teaching and support for substantive guidance early on in the educative process, drawing on studies of human working and long-term memory, but concurring with the constructivism view of how learning occurs, see Kirschner, Sweller, and Clark, 2006.

[629] On maturationist views in general, see Section 3.1 above; in an education context, see DeVries *et al.*, 2002: 'The romantic maturationist stream is based on the idea that the student's naturally occurring development should be allowed to flower without adult interventions in a permissive environment.'

[630] For a classic statement of this view, see again the work of Vygotsky, and Cole, 1978; for further empirical criticism of 'pure discovery' teaching strategies as an adequate translation of the constructivist theory of learning into pedagogy, and support of either explicit instruction or the alternative method of 'guided discovery', wherein key concepts are presented to aid independent problem-solving, see Mayer, 2004; and see again Kirschner, Sweller, and Clark, 2006 on the importance of guidance for novices in a subject.

[631] Cf. Dewey, 2002 [1922], pp. 58-59.

[632] I set aside here exceptional individuals who are primarily self-taught, focusing again on the typical case.

perhaps making an effort to develop an open, receptive attitude to their teachers will be what most effectively serves their interests as the desire-based theorist construes them, even if this conflicts with some of their more local desires in the here-and-now.[633] But again, I shall argue that this response again fails.

As before, for the Subjectivist's response to be valid, children would have to be capable of actively distinguishing between cases where participation in such relationships would lead to the satisfaction of their Subjective Values and those where this is not so, and then acting accordingly; otherwise, the response is as otiose as Wittgenstein's wheel and adds nothing to an Action-Guiding account.[634] However, despite the enormous advances made since infancy, when a child first enters formal education they are still not yet theorist enough to form this kind of judgement by themselves. For in order for the suggestion that the child's behaviour should be guided by such strategic, Instrumentalist considerations to actually be implemented in practice, they would first have to enjoy a solid grasp of the Means-Ends reasoning required to determine the consequences of their engagement in lessons. But this is not yet possible – precisely because of the lack of knowledge that renders them able to benefit from formal education in the first place.[635]

Similarly, Language Skills are also an essential part of reasoning itself, and acquiring linguistic tools through education is thus a necessary part of learning how to think clearly at all. Hence, learning these same skills cannot be solely the outcome of a process of advance rational deliberation aimed at implementing our Subjective Values.[636] And as we can learn these Language Skills only in a social context, this undermines any conception of interpersonal engagement that is purely Instrumental right from the outset:

> if thought depends on language, and language is a social institution, how can rational agency have a pre-social foun-

[633]For some empirics on the actual relation between educational attainment and Subjective well-being, see Witter *et al.*, 1984, who conduct a meta-analysis to argue that 'education is a small but positive contributor to adult subjective well-being' – p. 165.

[634]Cf. Section 1.7 above.

[635]See also Lepper and Green, 1978 for a classic discussion of 'overjustification' and the limitations of merely extrinsic motivation.

[636]Cf. again MacIntyre on learning the rules of a community prior to engaging in mature deliberation: 'Hence, allegiance to these particular rules has to precede any set of arguments, any theorizing.' – MacIntyre, 1992, pp. 9-10.

dation?[637]

We now consider two further reasons why the Relationship Good of acquiring advanced Language Skills cannot be seen as merely Instrumentally Valuable by the developing child.

In *After Virtue*, MacIntyre presents an observation credited to Sir Karl Popper: that regardless of the truth of causal determinism, it is impossible for situated agents to form rationally justified predictions about future radical conceptual innovations.[638] This is because in order even to frame such propositions it is necessary to employ the concept whose invention is being predicted; yet to do so simply *is* to invent it. Thus, it would be impossible to predict the invention of the concept of the wheel, since to even express this prediction one would have to already possess the concept of the wheel itself.

On an individual level, a similar line of reasoning shows that there are limits to how far a child can predict the consequences of their own future acquisition of concepts, since to frame such a prediction they would again need to already have these same concepts within their possession. Of course, it is admittedly true that in a general sort of way I can predict some effects of my gaining new knowledge of a fairly specialised nature. In mathematics, for instance, I might predict that one consequence of my learning about modular forms will be an advance in my enquiries in number theory. Yet this would require understanding rather a lot of other things about mathematics too, and in general such a hypothesis can be issued only from a mature adult standpoint – rather than the limited conceptual world of early-years education, where a child may indeed not yet have grasped *what* mathematics is.

A second further reason that the learning of Language Skills cannot be a merely Instrumental Good is the radical nature of the transformation that it effects in the child.[639] 'To know a second language is to possess a second soul', pronounced Charlemagne, and indeed the languages we acquire are imbued with normative commitments that may direct our thought into quite determinate channels – for instance, by allowing easier and more fluent expression of some thoughts rather than others. Consider, for instance, the Ancient Greek ethical terms *eudaimonia*, *arete*, *phronesis*, *nomos*: a human being who learns to speak using

[637] Larry Siedentop, 2014, p. 65.
[638] MacIntyre, 2011 [2007], p. 109; cf. Section 1.2.1 above.
[639] Cf. Kraut, 2007, p. 107: 'The learning of a language will transform what a child will want, and the plans she will make, in ways that are good for the child'.

this vocabulary is likely to think quite differently to our contemporaries from the outset.[640] And even in developed adult human beings whose capacities are already in place, new modes of expression we pick up tend to carry us forward with their own momentum.[641]

As we acquire new concepts, then, our Subjective Values are inevitably transformed, and new preferences given in terms of concepts with which we were previously unacquainted may also propagate. But since they so reshape our existing desires and affective responses, the acquisition of Language Skills cannot be understood as a mere means to satisfying these same Subjective Values.

Having now seen that a purely Subjectivist view of children's engagement in education is not viable, we again supplement this negative argument with some descriptive observations about what the educative process is actually like. Here I abstract from the two components of Practical Reason we have focused on so far, and consider what is required for learning in general to be successful – in part, drawing on my own experience of teaching mathematics as a private tutor.[642]

We begin with the teacher's view of the process. What conception of a child's good must guide the practice of teachers if the child's abilities are to develop? And what resources must they make use of in achieving these educational goals? Here the Subjectivist has a second potential response to what has been said so far: that ultimately, the primary role of teachers is in fact to contrive the satisfaction of their students' desires, or to bring them pleasure, so that education should still be guided by the students' Subjective Values after all, albeit indirectly.[643]

Here it may be conceded at once that a child's Subjective happiness is indeed important to bear in mind if their education is to be successful. For instance, if lessons are experienced as tedious or distressing, they are unlikely to be productive.[644] However, if a teacher merely brings the child pleasure or entertainment but contributes nothing of greater intellectual substance, then the educational outcomes we are presently interested in

[640] Cf. Dewey, 2002 [1922], pp. 58-59: 'An individual usually acquires the morality as he inherits the speech of his social group'; see also Dunne, 1993, p. 82.

[641] See Dunne, 1993, p. 136, who quotes Heidegger's adage: 'language speaks'.

[642] Having taught some 10,000 hours of one-to-one sessions in this subject.

[643] As with parents, we set aside the issue of whether this can be made consistent with the teacher's actions being guided by *their* Subjective Values.

[644] One prominent London-based private tuition agency, with whom I have worked for many years, is indeed called 'Enjoy Education'; see again Witter *et al.*, 1984 on subjective well-being and educational outcomes.

may also not be achieved.[645] For this purpose, what is required is not to satisfy students' wishes, or imbue them with positive feelings – but rather, to foster within them certain skills and capacities.[646] Hence, although students' Subjective Values are an important consideration for teachers, they cannot have the overriding centrality that a purely Subjectivist account of their good – as conceived from this third-person perspective – must accord to them.[647]

This notion is illustrated by a type of case that is a common experience in teaching: when a student falsely believes they understand an idea.[648] Here what is often required for the student to make progress is to move them to a condition wherein they are less satisfied but where their perceptions of their own understanding are more accurate and realistic, as inappropriate confidence is replaced with doubt or temporary confusion. The student is unlikely to enjoy this process: having one's existing beliefs contradicted is not usually a pleasant experience, and the resulting lack of certainty may be aversive. Yet it is necessary nonetheless, since 'a person is not going to undertake to learn anything that they think they already know.'[649]

In a more extreme type of case, teachers may encounter a student with no desire to succeed in a subject at all, or a felt aversion toward the entire teaching goal. In order to move forward with such a student, the teacher may first need to effect a transformation of their Subjective Values to become better aligned with their intellectual development – and more generally, the best route to success may often be to artfully instil a love for the subject. However, if this positive feeling is at first entirely absent, then this requires changing the student's preferences rather than merely

[645] Cf. O'Dea, 1993, p. 24: 'Teachers should see themselves as professionals whose central purpose is not to amuse their students but rather to engage their interest and curiosity, whose central task is not to make complex content easily palatable but rather to make it come alive in all its complexity'.

[646] See Weinert, 2001 on teaching as aiming at developing 'competencies'.

[647] This has been acknowledged even by those who advocate the adaptation of teaching methods to individual differences in students' personal interests: 'While we do not recommend that instruction be guided solely by learning style preferences, we believe teachers should make informed decisions about the areas or units within which style differences can be incorporated.' – Smith and Renzulli, 1984, p. 45; see also *ibid.*, p. 47; and pp. 45-46 on deliberately 'mismatching' teachers and students to foster positive change.

[648] This is of course also a key theme in Plato's early 'Socratic' dialogues, such as *Euthyphro* and *Gorgias*, wherein Socrates relieves the antagonists of false beliefs about their state of understanding; see also MacIntyre, 2016, p. 197 for discussion of dealing with the 'happy wastrel' *versus* the 'unhappy perfectionist'.

[649] Epictetus, 2008, p. 117 (II.17).

contriving the satisfaction of their existing preferences without passing judgement on their content.[650] But by assumption, this formative process cannot be entirely guided by resources available within the child's existing Subjective motivational set. And this idea applies more generally – as we will now illustrate through some further examples of educational best practices.

In discharging their role effectively, teachers often find it useful to set homework to be completed outside of lessons. Yet completing homework may not be especially conducive to pleasure, or something for which students have a pre-existing desire; indeed, it may be considered good for a child only in ways that the child themselves do not yet have the resources to appreciate.[651] But although imposing requirements that the student is not inclined to follow may risk making them irritable, teachers must not thereby conclude that continued encouragement in this direction is not good for them.[652]

Another common issue is that a student may express a desire to take a lesson down a path that diverges from what their teacher has planned: for instance, to skip a particular type of mathematical problem because they find it taxing, even though learning to solve such problems is important in light of the outcomes that the lessons are directed towards achieving. Again, it may be wise to acknowledge these preferences, since a student whose wishes are always ignored is likely to become frustrated and hence make no progress. But nevertheless, if a student does not enjoy a particular kind of problem, this is typically not adequate grounds for avoiding tackling them, and is perhaps even an indication that the reverse is true. Teachers cannot allow lessons to be derailed by arbitrary whims; hence, students' demands must sometimes be resisted.

If the teacher is unimaginative or impatient, they may simply invoke their authority at this point: 'Do as I command!' Yet a better teacher may instead try to offer an explanation of the *rationale* for knocking

[650] Compare with the traditional concept of *Bildung* in Germany, as envisioned by Hegel and Humboldt, which has moral as well as theoretical aspects, and which for Hegel must involve conflicts wherein the student overcomes their existing standpoint and values to achieve harmony with universal reason; see e.g. Wood, 1998. For an article on education that distinguishes what students need from what they happen to want, see also MacIntyre, 2006b.

[651] Cf. Craig, 2007, p. 194: 'The benefits of a good education ... may be visible only to those who have already had one or are well on the way to it.'

[652] For some empirical work on students' attitudes to homework, and the suggestion that teachers may need to work on altering their perceptions of its value, see Wilson and Rhodes, 2010.

back a student's requests, explaining why it is important to continue the lesson in the way they have recommended.[653] The Subjectivist theorist may argue that if this justification is to be rationally appealing to the student, it must point towards a greater amount of pleasure in the long run – perhaps bought with the large salary that good grades in certain subjects tend to invite, or invoke a stronger or more deeply help desire – such as their wish to pass their exams.[654] But nevertheless, if the teacher's explanation is successful, the student's Subjective Values will become reordered so that now they wish to practice the difficult type of problem; moreover, once some progress has been made they may come to enjoy tackling them too, with this bringing a satisfaction they had not anticipated.

This last point brings us to considering things from the student's point of view. What attitude from the student is required for such transformations of their preferences to occur? And more generally, what attitude is needed on their part for the knowledge and skills offered by an accomplished teacher to be adequately absorbed?

As a private tutor, one occasionally encounters students who hold something like a purely commercial model of education, in keeping with the Subjectivist spirit. They or their parents are paying for the sessions, so they argue, and hence the time should be used however they happen to see fit: 'Let's do things *this* way, because this is what *I want* to do.' But generally speaking, such an attitude tends to hold students back and get in the way of progress. The primary reason that students can benefit from private tuition at all is that due to their being at a more advanced stage of education a good tutor has a superior understanding and grasp of the intellectual principles of the subject in question. Yet if the student is concerned only with imposing their personal preferences, they cannot fully benefit from having access to this experience. Rather, doing so requires them to instead adopt an attitude of openness and receptivity to learning, thus permitting their transformation into more competent practitioners of the subject:

> Here one is involved and belongs with the other ... one listens to it and not only acknowledges its otherness but

[653] On the range of benefits of justification-based teaching as opposed to rote-learning in the context of school-level mathematics, including 'promoting conceptual understanding', see Staples, Bartlo, and Thanheiser, 2012.

[654] See again Lepper and Green, 1978, which suggests that the proposed strategy may be ineffective.

> remains *open* to the possibility of being affected by it oneself. This openness or receptivity to the other, a kind of *docta ignorantia* in which one is unsuspicious without being gullible, is the most essential requirement for fruitful interaction with a person or tradition.[655]

In particular, a key requirement for a successful teaching relationship is that the student is willing to acknowledge the teacher's superior understanding, and on this basis recognises their authority – thus deferring to their knowledge and insight when required.[656]

As a concrete example of this idea in action, consider a mathematics tutorial where a particular type of problem is being discussed: dividing polynomials by a linear factor, say. Suppose the student presently prefers to do this with long division, which they have practised many times, and are averse to approaching problems through the alternative method of comparing coefficients.[657] The tutor, in contrast, knows that the latter method is in fact better by any reasonable criterion once properly understood: it requires far fewer calculations whilst having the same range of application.[658] Under these circumstances, the tutor can clearly see that – unless there is the immediate pressure of an exam – in the long run, it is best for the student to abandon long division and focus on learning the comparison method instead. Yet these reasons may not yet be visible from the student's point of view, due to their being insufficiently experienced – both with the proposed new method and with mathematics in general.

If the student is to ever reach the point of mastery, and thus become able to find independent reasons for deciding between such techniques themselves, then at innumerably many junctures like this they must trust in the teacher's experience and allow them some measure of authority that may override what they have come to believe or to value so far. In particular, in developing themselves as mathematicians, students must be open to the idea that there are principles at work beyond their Subjective whims and existing 'powers of response', and adopt an attitude of trust towards the intellectual tradition of mathematics and the established

[655] Dunne, 1993, p. 133. On the concept of a 'growth mindset', wherein students see their own intellectual abilities as dynamic and pliable, see Dweck, 2006; for implications for teaching practice, see Boaler, 2013.

[656] Cf. Dunne, 1993, pp. 112-113.

[657] I.e., considering the expansion of part of the quotient to find its next term.

[658] Its complexity is of order n rather than n^2.

standards of excellence it embodies.[659]

This latter way of putting the matter also makes clear that although the educative process is essentially dialogic, it is not merely a case of the student passively deferring to the *teacher's* personal values and absorbing these; rather, education at its best is a shared quest for understanding at which the teacher is currently at a more advanced stage. And in addition to academics, this applies also to another facet of the teacher's superior knowledge: their understanding of *how* education should take place – including the nature of a good teaching relationship itself.[660] For such ideals come not only from the teacher's private, individual preferences, through the light of *their* inner direction, but are gleaned from many experiences with students, parents, and colleagues throughout their training and working life.[661]

Returning now to the teacher's perspective; very rarely, one may encounter an especially difficult student who completely rejects the resulting natural power dynamic of the relationship and does not accept their authority at all, despite the teacher's experience and expertise. For instance, asking them to complete a task, even if this is done politely, may be seen as an affront to their freedom; they may have a grievously inflated view of their abilities; they may insist that they understand concepts which they manifestly do not; they may behave in an impertinent or condescending manner. Although attempts to upset this rhythm in order to build a functional relationship may again make the student irate, humouring or placating the student and allowing this dynamic to become entrenched is an error that more experienced instructors would try to avoid. Instead, the teacher must find a way to reassert their authority – perhaps through a quiet but incontestable demonstration of their superior ability in the subject being taught.

[659]Cf. Williams, 2006 [1985], p. 125, quoted in Section 1.3 above; on these standards of excellence, see my Berry, 2019b; see also a quote usually attributed to Richard Feynman: 'Be humble. Be teachable. The universe is bigger than your view of the universe. There's always room for a new idea. Humility is necessary for growth.'

[660]On the importance of good student-teacher relationships for student engagement and associated developmental outcomes, see Hamre and Pianta, 2006; see also Carr, 2005 for a virtue-ethics perspective on how the personal nature of these relationships may conflict with the professionalism required from teachers.

[661]For some recent work on teachers' professional development that focuses on teachers learning through their own experiences and inquiries, but which also expresses related concerns about the impact of neoliberalism in this context, see Boyd and Szplit, 2016; on complex relationships having their own logic and dynamics that are partially independent of the goals and intentions of the two parties, see Barrett-Lennard, 2013.

As well as this respect for the teacher, there are also certain other basic requirements that students may have to follow in order for the educational goals we are considering to be achieved: for instance, having good attendance and being punctual, paying attention in sessions, replying to communication, doing set reading or homework, and actively engaging by thinking through and responding to questions.[662] But again, unfortunately one may occasionally encounter students who are exceedingly loath to meet these basic expectations: for example, they have no desire to complete their homework, or have an aversive feeling when doing so.[663] Hence, these conditions for positive student participation must be characterised independently of students' personal Subjective Values – and indeed, might make little sense if construed as mere tactics for satisfying them.

To summarise: though in the ideal case the educative process does appeal to the student's Subjective Values – the lessons are enjoyable, and they have a desire to progress in their studies – by considering these less exemplary cases, we see that a good teaching relationship cannot be *constituted* by the satisfaction of the student's existing preferences – even if this turns out to be a fairly reliable mark of one in the short term. In particular, regardless of students' private attitudes, without a basic attitude of compliance the necessary transfer of intellectual abilities cannot take place. Moreover, we have also seen several reasons why this attitude could not be contingently adopted as a Subjectivist stratagem at the level of Life Planning – for instance, at the point of development we are presently considering, students are not yet sufficiently intellectually advanced to formulate and justify this complex conditional thought.

This section has focused on our behaviours and attitudes as our linguistic and cognitive powers develop during childhood, and the deference to and openness to the transforming influence of our teachers that this education requires. We have seen that the inward-looking ethos of Subjectivism again imposes an implausible epistemological barrier here, since the principles that must guide the child's educational achievement are not yet present anywhere within themselves. Instead, they are only accessible through a certain kind of productive engagement with others, wherein certain Relationship Goods are recognised as Intrinsically Valu-

[662]See Meyer, 2003 for an influential perspective on good teaching practice; on student engagement and its importance for educational attainment, see Abulela and Bart, 2016; and again, Hamre and Pianta, 2006.

[663]See again Wilson and Rhodes, 2010.

able. But since these goods may extend beyond what the child presently desires or takes pleasure from, if children view their existing Subjective Values as constitutive of their good then this will systematically stunt their education and thus frustrate their progress in becoming Practical Reasoners.

In the next section, we move from childhood to adolescence, and consider an important arena in which our powers of Practical Judgement are tentatively expressed in an independent manner, as we begin to form and act on our own explicit judgements about goods.

3.4 Practices

Our third formative social context is engagement in 'Practices', in the sense of this term developed by Alasdair MacIntyre – whose work I will rely on heavily in this section. The focus will mainly be on adolescence here, since the importance of Practices for the present argument is most evident at this level of development – though initiation into some Practices may begin earlier or later.[664] We will also assume that the first four components of Practical Reason in the Subjectivist sense are now intact, and go on to consider the maturing agent's developing ability to make independent Practical Judgements about goods; an ability which we saw in the previous chapter is pivotal for implementing our Subjectivist views. But following the argumentative pattern of previous Sections, I will argue that the development of this ability requires a certain kind of participation in Practices which Subjectivism systematically precludes – whether applied within Action Selection or Life Planning.

A 'Practice' here is a 'coherent and complex form of socially established cooperative human activity' that is partially characterised by standards of excellence governing the performance of the activity, in such a way that through the historical progression of the Practice these standards are systematically extended.[665] Hence architecture, playing the violin, chess, and research in pure mathematics are Practices; as are the 'activities of members of string-quartets, of the crews of fishing-fleets, of architects and construction workers jointly engaged in developing good housing, of members of families making and sustaining the familial com-

[664]For some insights into the neurobiology of adolescence, see Sapolsky, 2017, pp. 154-159.

[665]MacIntyre, 2011 [2007], p. 218.

munity, of farmers and of physicists'.[666]

Also key to the definition of a Practice is that in striving to attain the standards of excellence by which it is partly constituted, certain 'Internal Goods' are thereby achieved. 'Internal Goods' in this sense are those goods which can only be specified through elaboration of the Practice itself, which are only accessible through participation in the Practice, and which can only be accurately identified and appreciated by those with sufficient experience of the Practice. Hence for chess, the Internal Goods include 'the achievement of a certain highly particular kind of analytical skill, strategic imagination and competitive intensity'.[667] These contrast with items such as money, power, or status that may be contingently attached to the mastery of Practices, which I shall call 'External Goods'.[668]

It follows from the above definition of 'Internal Goods' that 'Those who lack the relevant experience are incompetent thereby as judges of internal goods'.[669] And likewise, a grasp of the associated standards of excellence that characterise the Practice must be acquired alongside learning to participate in it. We therefore ..

> .. first encounter goodness as evaluated by standards in the light of which we fall short and the understanding of which is a task rather than an achievement, a theoretical task inseparable from the practical task of becoming good at farming or chess or physics.[670]

It is thus through practical engagement in a Practice that we learn to make rationally-grounded judgements about both goods and good performance within it, at the same time as learning how to overcome obstacles to the achievement of those goods and to how to move forward when progress in our performance seems to have stalled.

[666]MacIntyre, 1992, p. 7; see MacIntyre, 2011 [2007], pp. 218-219 for further examples. There is some debate in the secondary literature on MacIntyre concerning what exactly constitutes a Practice; I cannot enter into this here.

[667]MacIntyre, 2011 [2007], p. 219.

[668]MacIntyre, 2011 [2007], p. 219. This distinction may have been drawn slightly differently in the later secondary literature, with more of a focus on the role of External Goods as the 'currency of institutions'; given our current focus, I again set such considerations to one side. Another important distinction MacIntyre draws is between Practices and the institutions in which they are housed; we will not need to address this point either.

[669]MacIntyre, 2011 [2007], p. 220.

[670]MacIntyre, 1993, p. 4; see also MacIntyre, 1998, p. 31: 'in all these areas there is not only progress in achievement but also progress in our conception and recognition of what the highest perfection is.'

From the perspective of the present enquiry, the import of this essential connection between the practical and theoretical aspects of making genuine progress with Practices is that the latter both requires and is necessary for the development of our Practical Judgement;[671] a capacity which we have seen that the Subjectivist must recognise as central to the rational pursuit of our good.[672] It therefore follows that to make progress with Practices and to develop the requisite powers of Practical Judgement in general are in fact one and the same thing. How, then, is this progress achieved?

In keeping with the conclusions of the chapter so far, such abilities do not simply appear from nowhere, and the social context is again of paramount importance:

> [We] become poets or painters or musicians not by some process of development from within ... but by living in a society where these languages are current.[673]

In particular, the primary way our Practical Judgements about goods and good performance are improved is through correction and critical feedback from others who are presently at a more advanced stage than ourselves.[674]

In order to gain adequate access to and to make apt use of these Relationship Goods, serious students seeking to attain true mastery of a Practice must undergo a period of initiation by assuming the role of an apprentice.[675] And as we found with education in general, this in turn involves a certain attitude of openness and responsiveness to the constructive criticism offered by these more advanced others, as well as a degree of humility about our current standing:

> Such education requires a recognition that our judgments can be in error ... and will be apt to be in error until we have been adequately initiated into and educated in accordance with the standards of the relevant practice.[676]

If they are to learn what they need to learn, then, the attitude of the apprentice must be of one who is teachable because he or she acknowledges 'his or her own incapacity for judgment in respect to the goods

[671] See below for an illustration of this point in the context of a particular Practice.
[672] As argued at length in Chapter 2 above.
[673] Collingwood, 1958, p. 317.
[674] Cf. Greene, 2014, p. 86-87.
[675] On becoming an apprentice in a craft, see MacIntyre, 1990, p. 61-62.
[676] MacIntyre, 1993, p. 15.

internal to that practice'.[677] But in recognising this, apprentices must at least implicitly draw a more general conceptual distinction: one between 'what merely seems good to them and what really is good (a good way to plough a field, for example, or to write an elegy)' according to the standards of the Practice, as represented by the judgement of more experienced practitioners.[678] Moreover, as with all rational public discourse concerning excellence within a Practice, the formative critical feedback on our judgements and performances that we receive from these others cannot take as its guiding standard the satisfaction of *our* private Subjective Values.[679] Rather, it must be directed toward the socially-established *shared* standards that partially define this particular Practice; standards which orient practitioners towards the achievement of its associated Internal Goods, and which emerge only through the history of the Practice as it has developed hitherto:

> the question of what such goods are and what such excellence is – excellence, for example, as of a fisherman or a pitcher or a chess-player or a violinist or a physicist – is answered by the standards of each particular practice and not by what any particular person happens to prefer to choose to be pleased by.[680]

For instance, when our golf coach comments upon whether we have selected the right club for a particular shot, her response will not be based on the private criterion of whether we have picked the club that we *most wanted*, or even the one that will give us the *most pleasure* to use; but rather, the club that is most conducive to the publicly-verifiable outcome of positioning our ball on the green. The type of Practical Judgement initially developed and exercised through engagement in a Practice is therefore broadly Aristotelian in character, rather than the Subjectivist mode specified by decision theory and its kin.[681] And these

---

[677]MacIntyre, 1993, p. 4.

[678]MacIntyre, 1987; see also here MacIntyre, 1990, p. 61-62; MacIntyre, 2009b [1999], pp. 68-71; and see MacIntyre, 2016, p. 49; p. 134 on coming to distinguish what we want from what is good for us.

[679]My argument here loosely draws on Wittgenstein's 'private language argument'; see Wittgenstein, 2003 [1953], p. 75^e.

[680]MacIntyre, 1993, p. 6.

[681]See MacIntyre, 1982 for extended discussion; see also MacIntyre, 1987, pp. 4-5: 'Hence the question of whether or not something is a good reason for action is and comes to be understood as independent of the will or preferences of any particular individual.'

same observations also cause problems for Subjectivism in a more direct way.

We may concede to the Subjectivist that what *seems* good to us at the present moment is indeed largely a reflection of our present desires and affective preferences. But it then follows that the apprentice's initial incapacity for Practical Judgement about goods and good performances is at least partly 'an incapacity deriving from the inadequacy of his or her attitudes, tastes, feelings, and thought in the uneducated condition'.[682] Hence, in coming to recognise the value of critical feedback and what they can learn from it, apprentices must also draw a conceptual distinction between ..

> .. what I would do, if I did what would please me most here and now, and what I would do if, in the light of the best instruction available to me, I were to do what would make me excellent in the pursuit of the goods internal to the particular practice or practices in which I am engaged.[683]

As well as specifying what it is to perform well within a Practice, then, the standards of excellence in play here also provide a means of evaluation of our Subjective Values themselves:

> To enter into a practice ... is to subject my own attitudes, choices, preferences and tastes to the standards which currently and partially define the practice.[684]

This mode of thought is, of course, one which the Subjectivist cannot endorse; yet it is essential for attaining proficiency with a Practice, and hence also necessary for developing just those powers of Practical Judgement that Subjectivism itself requires for its own practical implementation.

Finally, it is not merely that the pull of our Subjective Values may on occasion need to be resisted at the level of Action Selection, in those particular cases where they clash with the pursuit of the Internal Goods of the Practice. Ultimately, because of their central motivational importance and their inevitable impact on our judgements, to achieve mastery

[682]MacIntyre, 1993, p. 4.
[683]MacIntyre, 1992, p. 7.
[684]MacIntyre, 2011 [2007], p. 221; see also MacIntyre, 2016, pp. 49-50, and MacIntyre, 1993, p. 4: 'from the standpoint afforded by a practice, the standards of good cannot be understood as an expression of anyone's tastes, feelings and thoughts, for taste, feelings and thoughts are themselves to be characterized and evaluated in terms of those standards.'

we must also strive to develop, train, and educate our Subjective Values in accordance with the standards specified by the Practice. That is, over the longer time scale that is the province of Life Planning, the aspiring apprentice must ..

> .. confront the tasks of bringing his or her tastes, feelings, and thoughts into alignment with the judgments of those others who are the masters of the particular craft and who have the authority of education and experience. [685]

Yet rationally engaging in and successfully completing this transformation is again something that holding a Subjectivist view about our own good will systematically preclude. Subjectivism thus further impedes us from making progress with Practices here, and we must instead cultivate a certain distance and critical detachment from our existing Subjective Values if our powers of Practical Judgement are to develop.[686]

We turn now to exploring a concrete example of a Practice. Here I aim to illustrate the crucial connection between our engagement in Practices and the growth of our ability to exercise independent Practical Judgement as such, as well as to corroborate the observations made so far about how this engagement must proceed in order for us to make solid progress.

In previous work, I have discussed pure mathematics as a Practice, outlining normative standards of excellence that characterise what it is to do mathematics well – the most important of which for my enquiries has been that the underlying concepts employed in a mathematical argument must have clear defining conditions.[687] Here I will consider quite a different Practice: cooking and appreciating good food.

I choose this particular Practice for a number of reasons. Firstly, human biological needs being what they are, the preparation and consumption of food is an activity that we must all have some particular relation to: even asceticism may be taken as a definite, substantive attitude here.

Secondly, here we will also find an illustration of the qualities of mind and character needed for successful deliberation at both of our core levels

[685] MacIntyre, 1993, p. 4; cf. *ibid.*, p. 9: 'with the assistance of others further advanced than ourselves and therefore standing in authority relative to us, [we must] reshape ourselves, reorder our feelings and desires, develop powers of discrimination and distinction-making'.

[686] On this point, see MacIntyre, 2009b [1999], p. 69; p. 83; p. 136.

[687] See again Berry, 2019b; cf. Berry, 2018.

of enquiry. In terms of Action Selection, we have here a clear arena where Practical Judgement about what to do is constantly required – whether to add more salt to a sauce; when to cease cooking pasta or risotto; what combination of courses to serve at a dinner party, and so on. Agents must learn to make such Practical Judgements about goods and the good in this context in the same way they must learn to make them in general. Meanwhile, at the level of Life Planning, as our relationship with food evolves, we can guide the development of our tastes and dietary habits, and our culinary prowess as it grows or fails to grow over time. In both of these aspects, then, we find our main problem in microcosm: deliberating about what is it make a good meal and what it is to pursue our good *qua* aspiring chef are simpler versions of the larger tasks of working out how to act in general and how to direct our lives as a whole.

A third reason for choosing this Practice is that the culinary arts – as opposed to, say, open heart surgery – are a type of Practice for which engagement in the spirit of Subjectivism seems most plausible, *prima facie*. Although the effectiveness of specific techniques *qua* means to a fixed culinary end may be debated rationally – more or less efficient ways of chopping onions, say – in terms of the end product itself, it may be plausibly contended that *de gustibus non est disputandum*; that individual preference is sovereign in precisely the way that Subjectivist theory would require here.[688] This choice of Practice will therefore constitute a good test case for the validity of the injunctions introduced above.

On a sociological level, food is indeed often evaluated according to a Subjectivist standard in practice. For instance, we often eat to satisfy the innate drive of hunger, and it seems unlikely that someone would deem food to be excellent unless it tasted agreeable to them. For this latter reason, Plato's Socrates was critical of cookery, since he regarded it as aiming only at pleasure – unlike medicine, which was aimed at the good of the body – and thus would likely have denied it to be a Practice in the present sense.[689] And indeed, if the phenomenology of eating the same dish actually varies in systematic ways amongst its consumers – perhaps partly due to relatively fixed genetic factors – then in preparing food, it would seem foolish not to take these individual differences in

[688] See again Stigler and Becker, 1977 for a statement and criticism of the traditional view.

[689] Plato, *Gorgias*, 462d-463.

their Subjective Values into account.[690]

On a larger scale, it may also be argued that if we consider received standards of excellent cooking across different societies, we find a radical cultural relativity here; consider the contrasting roles of spices in West Asian *versus* European cooking, for instance. And to the response that cookery is then best thought of as fragmenting into a number of distinct Practices, such as Italian, French, or Japanese cuisine, each with its own objective standards of excellence, the Subjectivist may also counter that it can indeed be individuated even more narrowly than this – for instance, the latter into Bolognese, Venetian, and Roman cooking, amongst many others; yet as this fragmentation continues, these standards of excellence come to look increasingly like merely local proclivities.[691]

A Subjectivist view of cookery may also seem to find some support when observing discussions amongst cooking enthusiasts today, where utterances with individualistic overtones – such as 'I liked it, and that's all that matters' – are often heard even from those who are fairly committed to improving their skills, and where a strong negative reaction to those purists or traditionalists who believe otherwise is also commonplace. However, things are not so simple as these preliminary observations about some views represented within our culture would make them appear. For at the same time as the presence of a radical Subjectivism, a view quite inconsistent with it is expressed through widely respected institutions such as the Michelin guides, whose famous star system is awarded based on apparently objective standards of excellence. This moves us beyond the sovereignty of individual preference: whilst enjoying our own food may be nice, to incur the approbation of our friends too is better; and its reception is rais'd to its utmost perfection by the universal applause of the cultivated world.[692] Yet others are skeptical of official rating systems too, and loudly declaring 'Each to their own!' consider anything beyond a pragmatic hedonistic standard as mere prissiness or snobbery.[693]

[690]Cf. Baltzly, 2020. For example, some individuals seem to have a partly-genetic aversion to coriander, which causes it to taste like soap: see Eriksson *et al.*, 2012.

[691]See Hazan, 2011, p. 1: 'Ask an Italian about Italian cooking and, depending on whom you approach, you will be told about Bolognese, Venetian, Roman, Milanese cooking or Tuscan, Piedmontese, Sicilian, Neapolitan. But *Italian* cooking? It would seem no single cuisine answers to that name.'

[692]Paraphrasing Hume, 2000 [1739-1740], p. 121 (*T* 1.4.1).

[693]Again, Williams' claim that – *contra* the ancients – we do not always desire things under the 'guise of the good' would complicate our relations to these official standards here even if their validity is accepted. For then there would then be no inconsistency with my announcing that I prefer the delights of my local kebab shop to *haute cuisine* whilst also recognising the objective superiority of the latter. However, as noted

The existence of conflicting attitudes here entails that neither attending to the received views of our culture nor appealing to shared 'intuitions' which ultimately derive from these same sources will by themselves permit us to settle the philosophical question of the objectivity of the standards of excellence associated with cooking. Moreover, from our Action-Guiding perspective, these resources will likewise not suffice for resolving disagreement about the type of reasons that should guide us in our culinary preparations: whether these are to be objective, 'external' reasons that are independent of our individual tastes, inclinations, and preferences, or else only 'internal' reasons that are nothing but reflections of this same '*subjective motivational set*'.[694] For it would appear that there is no logical inconsistency in either defining or deliberating in terms of either type of reason, since both are coherent concepts which are indeed in active use in this sphere today.

Fortunately, within the present methodological framework, we have recourse to another strategy for making progress here: to withdraw from these skirmishes at the level of Action Selection to the broader standpoint of Life Planning, thus considering the development of our relationship to food and cookery over time. More specifically, we can move past this *impasse* by asking what attitudes towards goods and what conceptions of reasons are most appropriate *if* we are to make progress with the Practice of cookery. For our concern is not with abstract metaphysical questions, such as 'What is a reason?' and 'Are goods objective?', but rather with normative-practical questions, such as 'What conception of 'reason' is most appropriate to adopt here?' – and 'How should a successful practitioner think about these standards of excellence?'[695]

In their initial encounters with a Practice such as Italian cooking, an individual may at first act as a pure consumer who is indeed guided by their Subjective Values alone. From the menu at restaurants, they pick whichever combination of items they have a felt wish for, those dishes that are perceived to be most delicious to them at that particular moment, and make such additions or substitutions as they happen to see fit. In attempting to prepare their own Italian-style food, they may introduce foreign elements such as tomato ketchup, which in their experience so

in Section 1.3 above, this is not a concern for the present enquiry, where we are specifically interested in Action-Guiding standards *of the good*.

[694]Williams, 1979, p. 18, quoted in Section 1.5 above; cf. MacIntyre, 2009b [1999], p. 86 for discussion of the general issue.

[695]Cf. the discussion of our 'normative constructivism' methodology in Section 1.3.

far seem dependable and familiar. In making a sauce for pasta, they are again guided by their existing inclinations: it is a matter of picking a few ingredients they happen to like and combining them in a way that strikes them as most appropriate or convenient at the time. Wine is selected without knowledge or regard for conventional advice about parings; perhaps they have a favourite vintage they take with any meal.[696]

Such an individual may, however, decide to become more serious about Italian cooking; to participate more fully in or even try to make new contributions to the Practice. Let us assume for the time being that this is indeed the case; that the individual in question is serious about learning to excel here.[697] In keeping with the claims made at the beginning of the section, I shall argue that the serious novice must then take a different attitude if their culinary powers are to reach their full potential.

In this context, one often hears talk of an innate preparedness – of 'having it in your blood', or that 'cooking is one of the natural instincts'.[698] Yet in endeavouring to further their understanding and improve their skills and abilities, an aspiring chef will soon encounter situations where the tradition of the cuisine imposes *its* standards on *them.*[699] And in order to adequately absorb what the tradition has to teach them, I will argue that such apprentices must be open to recognising and being moved by external reasons that extend beyond their own Subjective Values; improvement thus indeed requires us to have a certain detachment from our existing preferences, rather than regarding them as the sole determinant of the good. Let us then look in more detail at how this process of development unfolds.

As argued for Practices in general above, in learning to cook Italian food well, one generally needs to begin from and submit oneself to the guidance of others. This social learning may take place in a number of different ways, at different levels of proficiency.

Since the standards of excellence that partially define the Practice

[696]On engaging in the activities associated with Practices but without commitment, see MacIntyre, 1993, p. 16. It may in fact be the case that this casual attitude is not optimal even by the standards of the consumer's existing untutored affective preferences, since as well as not yet knowing how best to implement these on a technical level, they may not yet even fully understand them either; see here Section 2.1 on decision *versus* experienced utility.

[697]See below for a Subjectivist response that questions this assumption.

[698]As put by Uncle Monty in Bruce Robinson's superb film *Withnail and I.*

[699]Paraphrasing MacIntyre's remarks on great art – MacIntyre, 2016, p. 144; see also MacIntyre, 1993, pp. 13-14.

of Italian high cuisine themselves derive in part from the achievements of the great culinary masters, one way to acquire and internalise these standards is to look directly to them for inspiration:

> Those achievements are assigned a canonical status within the practice of each type of activity. Learning what they have to teach is central to apprenticeship in each particular form of activity.[700]

In particular, in order to make rapid progress it is often helpful to find a specific practitioner to act as a mentor or role model; one whom we can acknowledge as a reliable authority whilst our abilities develop.[701] A wise apprentice will choose someone skilled and knowledgeable here, such as the late Marcella Hazan – who was not only a foremost expert on the history of Italian cuisine, but also held a PhD in biochemistry.[702] For a career chef, the resulting mentorship may be constituted by a professional relationship whilst working together at a restaurant; for serious amateurs, it may take a more explicitly educational form, as it did for those attending Hazan's famous cookery classes. But although such direct, interactive relationships may be the ideal way to learn, many of us must instead make do with learning from the masters from a much greater distance – most often in the form of reading what they have bequeathed to us in writing.[703]

For many aspiring young cooks, then, their first steps in advancing from the initial condition of a complete beginner are taken by following written recipes. However, although these instructions may be read with great diligence, such novices cannot yet act with a full understanding of what they are doing. For although at this point our rational powers in general may already be developed, we will at first lack the conceptual resources required to fully understand the reasons behind certain instructions these recipes issue – such as using a large pan of salted water for cooking fresh pasta, or peeling and removing seeds from tomatoes, or adding fresh basil only at the end of cooking a sauce. Yet although these steps may seem unnecessary or overly complex to our untutored inclinations, if we are to absorb what we need to learn here then the

[700]MacIntyre, 1988, p. 31.

[701]Cf. Greene, 2014, Chapters 2-3 for insight into finding a mentor and the value of accepting the corresponding relationship dynamic.

[702]Hazan, 2011 [front matter].

[703]Of course, today many aspiring young chefs watch online videos rather than reading books.

process requires us to place our trust in the experience of the author of the recipe, thus following the recommended procedures dutifully even where we currently lack this understanding of why they should be performed. Before we can learn to make our own judgements with confidence, we must first act partially under the direction of more advanced others; and if asked for justification at this initial stage, we must sometimes simply say: 'That's how I've learnt to do it.'

Although written recipes are thus a good starting point, however, the quality of our preparations can only progress so far through reliance on these alone. For instance, even if we follow prescribed weights, timings, and temperatures with scrupulous accuracy, doing so cannot always exactly reproduce the intended results because of the great differences in innumerable variables that may occur from one kitchen to another, such as quality and freshness of ingredients, thermic and other material properties of equipment, background temperature and humidity; differences which no writer can adequately account for in advance.[704] And more generally, there are always limits to what we can learn from any explicit, propositional account of the techniques of a Practice:

> What can never be done is to reduce what has had to be learned in order to excel at such a type of activity to the application of rules. There will of course at any particular stage in the historical development of such a form of activity be a stock of maxims which are used to characterize what is taken at that stage to be the best practice so far. But knowing how to apply these maxims is itself a capacity which cannot be specified by further rules.[705]

In gaining true mastery as a chef, then, one must develop beyond merely following rigid written specifications for what to do by rote. Rather, what is required here is to develop something approaching the kind of flexible practical wisdom that Aristotle captured with his notion of *phronesis*.[706] An experienced chef may intuitively see, for instance, that on a particular occasion she must add some water to an egg wash to prevent her breading from sticking, without this having been suggested to them explicitly by any recipe book. Hence, as claimed above, making progress with

[704] The same issue with laboratory conditions also causes problems for replication in psychology – see Anderson and Maxwell, 2016, p. 3; on the variability of moral agency and its concomitant conditions, see Aristotle, *NE* I.3, 1094b14-18.

[705] MacIntyre, 1988, p. 31.

[706] Cf. Dunne, 1993, p. 81.

Practices goes hand in hand with developing our capacity for Practical Judgement, as we become more responsive to these contextual factors. And at the same time, we will also slowly begin to understand the reasons underlying the procedures we first learn by rote, thus becoming increasingly rational and self-directing in this area of our lives. But how does this capacity for good judgement come about here?

Although in learning to cook, the personal experience developed through exposure to many possible eventualities is invaluable, the resources supplied by our social context are again essential too.[707] Access to an experienced teacher who can give us direct and immediate guidance and critical feedback on our performance is again the ideal case; yet even if this is lacking, we can nevertheless learn to improve our judgement through others by presenting our food in social contexts – the paradigmatic occasion being catering for our friends at a dinner party.[708]

As a beginner, it is often apt to listen carefully to and try to learn from the comments and reactions of these others, thus taking note of a consensus that a dish we have prepared is too sweet, bitter, watery, dry, raw, or burnt. Through such feedback, we might also come to see that our tastes are somewhat peculiar in ways that it is in our interest as maturing chefs to correct – for instance, now realising that we have become especially unresponsive to salt, leading us to add what others consider to be too much seasoning to our preparations. However, in order to acknowledge the value of such insights we must again exhibit a certain openness, receptivity, and willingness to reconsider what have until now been our personal preferences.

A Subjectivist might respond to the argument so far by claiming that any new knowledge that we might glean from others here can be rationally recognised as worthwhile only as a means to attaining certain External Goods that we already value – such as in enhancing our ability to impress these others with our cooking, and thus securing their allegiance or love or respect – rather than as a discovery of the objective Internal Goods available through the culinary arts. Yet although such considerations may indeed be part of our private motivations for attempting to contrive the approval of others, what this objection misses is how our development as chefs is nevertheless constituted in part by the education of our palates

[707] See again Bandura and social learning theory on the relative efficiency of learning from the experiences of others rather than directly from the physical environment.

[708] Though again, today many enthusiasts now also photograph their dishes and post them in online discussion forums for comment.

and taste for food through the types of social learning detailed above, rather than being merely our finding more efficient ways to engineer the satisfaction of these Subjective Values.

What is required to make progress with cooking, then, is to educate and develop our preferences by making available to ourselves the best principles, techniques, and standards that have emerged from respected practitioners of the culinary arts during the course of their history so far. Yet this process cannot be guided solely by our existing Subjective values. For at first, our appreciation of good food may be uncultivated and our aptitude for recognising particular tastes undeveloped – and if such a novice, would-be chef looks only inwards within herself, then she will find no resources to direct her and no coherent plan to move forward. As we attempt to move beyond our initial uneducated position, a purely Subjectivist conception of the good is therefore critically limiting; and if an agent recognises only internal reasons for action in this context, this will render them blind to the distinctions required to progress from the initial condition of a beginner. For such an approach to practical deliberation ..

> .. obscures from view the way in which agents have to learn at various stages how to transcend what have been up till this or that point the limitations of their motivational set and will fail badly in their moral [and culinary] development, if they remain within those limitations.[709]

Instead, in pursuing the education of their inclinations and palates *en route* to mastery, the novice chef must be open to the 'transformation of her or his motivational set, so that what were originally – in Williams' terminology, although now differently understood – external reasons also become internal'.[710] Moreover, this is a process that cannot be entirely foreseen in advance, since our knowledge of the relevant Internal Goods and standards of excellence appropriate to cookery is, as of yet, still lacking. It therefore cannot be the outcome of even indirect Means-Ends Reasoning that aims at satisfying our prior Subjective Values at the broader level of Life Planning.

These remarks about the early stages of our culinary education may also be answered with a contrary claim at the other end of the scale of achievement: that one who only slavishly follows the techniques of

[709] MacIntyre, 2009b [1999], p. 87 [my addition].
[710] MacIntyre, 2009b [1999], p. 87.

others cannot become a true master, and that 'the greatest achievements in each area at each stage always exhibit a freedom to violate the present established maxims'.[711] Even accepting the above claims about the beginner at cooking, then, this observation may be thought to occasion a return to a more individualistic attitude later in a practitioner's career. Thus a great chef who has earned the right to go beyond the received principles of the tradition and whose relationship with other authoritative practitioners is now that of equals may originate their own distinctive style of cooking that is grounded in and expresses their unique identity, so that their personality and values are imprinted on the dishes they create. And so, as our powers of judgement and execution in this area reach fruition, it is possible that our Subjective Values then come to supply us with reasons for altering the received directions of the Practice as it has existed hitherto.

It is indeed part of the definition of a Practice that its associated standards of excellence are systematically extended over time, as the Practice advances and human powers to achieve excellence in this area thereby increase.[712] However, although there is always room for innovation, such an attitude is appropriate only for those with advanced knowledge of the tradition from which they deliberately and consciously choose to depart. The expert chef may indeed base her recipe choices on her private inclinations; but she follows these not merely because they happen to be her *de facto* preferences, but rather, because they are the developed preferences of an experienced and authoritative practitioner, and as such have already been corrected, educated, and transformed in light of the existing standards of excellence that characterise the tradition in its present form:

> The rules of practices are made to be broken, but only by those who have become so proficient in practical reason that they know better than those who previously framed the rules.[713]

Thus it is perhaps no accident that Cézanne and Picasso were trained to draw and paint in the classical style before moving so far beyond it; and at first, when such budding innovators are still beginners and even intermediates, the best strategy remains to be open to guidance from the

[711] MacIntyre, 1988, p. 31.

[712] See again MacIntyre, 2011 [2007], p. 218.

[713] Knight, 1996, p. 13; cf. MacIntyre, 2009b [1999], pp. 86-87 on *de facto* versus good-oriented desires.

established masters.

A final Subjectivist response would be to acknowledge the above account of what is required for success within a Practice to which one is already committed, but add to this the further claim that the only reasons we could have for thus seeking mastery within any particular Practice in the first place must necessarily be based on deeper preferences already existing within ourselves.[714] Where such fundamental inclinations towards success in this area do happen to be present, the response would thereby characterise the apprentice's progress in terms of the increasingly adequate satisfaction of these, with the above directives being seen as hypothetical imperatives that are merely means to this end; where they are absent, this account of our engagement may be rejected entirely. The answer to which conception of 'reasons' an agent should adopt when acting in this sphere would then turn on the prior question of whether they want to become (or would be pleased by becoming) a chef; and indeed, although earlier we assumed that our novice is serious about developing their skills, this is an assumption that is often false in practice.

To respond to this objection, we must again take a step back: this time from the particular Practice at hand to the place of Practices as such within our lives as a whole.[715] It is indeed possible to think critically about the role of Practices such as cookery within this larger context: that is, to consider whether or not to become serious about our participation in a Practice, and hence whether to follow the account of Practice-Based reasoning given above when we engage in this area. However, deliberation of this sort is an instance of Life Planning, and so to be conducted rationally requires our powers of Practical Reason to already be in place. In particular, reasoning about this matter in a Subjectivist mode requires us to make immensely complex choices between the committed pursuit of very different sorts of goods, based on predictions of the impact of doing so on our Subjective Values.[716]

[714] Cf. Anscombe, 2000 [1957], p. 66, discussing Aristotle: 'Similarly 'Dry food ... suits anyone etc., so I'll have some of this' is a piece of reasoning which *will go on only in someone who wants to eat suitable food.* That is to say, it will at any rate terminate in the conclusion only for someone who wants to eat suitable food'; see also MacIntyre, 1982, pp. 301-305 for a critical discussion of this attitude in the context of practical reasoning within Practices.

[715] On the place of particular Practices within our lives as a whole, see MacIntyre, 1992, pp. 7-8; MacIntyre, 2009b [1999], pp. 66-67; and MacIntyre, 2011 [2007], pp. 233-234.

[716] Cf. Section 3.4 above on the difficulty of making such large-scale choices according

Yet I have also argued that engagement in Practices is essential to the development of precisely the same capacity for Practical Judgement that would be required to make an independent choice here. So, because this process of developing our rational powers must be well underway before we are even able to answer these difficult questions about particular Practices, our earliest engagements with Practices as developing rational agents cannot themselves be guided solely by a Subjectivist standard. For although as a mature agent I may choose to reject the claims of a particular Practice such as cookery for reasons the Subjectivist can endorse as rational, I will first need to fully engage in at least *some* Practices before I can be a fully-fledged rational agent at all.[717] The Subjectivist therefore cannot successfully evade the claims of a Practice-based understanding of goods altogether.

How, then, does the selection of those particular Practices that will claim our allegiance occur? At a young age, we again often proceed partly under the guidance of others here; those parents, teachers, friends, or role models who not only contribute to shaping our emerging system of values, but who may also provide us with new opportunities or help make us aware of opportunities we already possess. Critical reflection on Practices may often instead consist in opposing the standards of excellence associated with one Practice to those of a different Practice we are already committed to: perhaps we might conclude that food which is excellent by the standards of Italian fine dining is not wholly consonant with the art of maintaining our health, for instance – Marcella does suggest that we use rather a lot of butter![718] And in other instances, we may follow the received views of our culture about what excellent conduct consists in and which types of accomplishment are to be commended, and about goods and the good life as such.

To summarise: in this section, I have defended four key claims. Firstly; making progress with at least some Practices is necessary for the development of our powers of Practical Judgement. Secondly; this progress must be guided by detailed critical feedback that is only available through engaging in the Practice within the relevant social environment. Thirdly; accessing these Relationship Goods further requires both a recognition of the associated Internal Goods and authoritative

to a hedonistic standard.

[717]Cf. again MacIntyre, 1992, pp. 9-10 on commitments that must precede any practical deliberation.

[718]See Hazan, 2011, p. 165; cf. again Plato, *Gorgias,* 462d-463e.

standards of excellence of the Practice, and a certain deferential openness to absorbing the techniques, knowledge, and values of more advanced practitioners. And finally; this is a process which adherence to our Subjectivist theories tends to systematically preclude. Hence, because the powers of Practical Judgement that must be acquired in this arena are themselves required for implementing our Subjectivist conceptions of Practical Reason, these Subjectivist views are thereby rendered inadequate as comprehensive Action-Guiding accounts of the Human Good.

3.5 The Sapling and the Caterpillar

So far in this chapter, I have argued that for not-yet fully-mature agents, such as infants, children, and adolescents, any purely Subjectivist Action-Guiding conception of the Human Good must fail, since adopting such a view would hold them back from developing just those powers of Practical Reason whose discharge is at issue in our enquiry. In these final two sections, I extend this conclusion to human life in general. To this end, I will first argue that the status of being an independent Practical Reasoner which we achieve in early adulthood is always liable to go into decline without certain kinds of help from supportive others, so that the rational capacities peculiar to human beings can only be sustained within a certain kind of community.[719] Then, I will further claim that participation in the life of this community in the required way is again not possible for one who holds a purely Subjectivist conception of their own good.

We begin with an important line of response aiming to block this extension of the claims made thus far. As we have said, the focus of the earlier parts of this chapter – infancy, childhood, and adolescence – is not a typical one, and within analytic philosophy and economics work in this area instead tends to centre on adult human beings whose powers of Practical Reason are already fully developed. Hence, it is natural for those who hold an Action-Guiding Subjectivist view to try to sidestep the above arguments by explicitly restricting the application of their accounts to these fully mature agents. Whilst acknowledging that neither desire nor pleasure can furnish us with a successful Action-Guiding theory of what is good for those who are still young, then, such

[719] Cf. MacIntyre, 2009a, p. 130 on the 'illusion of self-sufficiency'.

theorists would nevertheless maintain that later in life our Subjective Values can now function as the basis of an adequate conception of the Human Good. And such a view would also fit well what has been called called 'Aristotle-Friendly Liberalism' within political philosophy, wherein the official liberal neutrality about the good life only applies once a human being has come of age.

Whether some such bifurcating view can be theoretically motivated depends upon the answer to a question raised by philosopher Patrick Tomlin: whether children are more like saplings or caterpillars.[720] A sapling is essentially a smaller version of a fully grown tree, whereas a caterpillar is radically different to the butterfly it transmorphs into later in the organism's life cycle. We would thus expect roughly the same substantive conditions to govern what is good for saplings and trees, but might find that quite different accounts are needed for butterflies and caterpillars. Likewise, if children are in some relevant respects like caterpillars, then it may be plausible to hold that we also need two quite different conceptions of the Human Good too: one for children, and a quite different one for adults.[721]

In the remainder of this thesis, we shall concentrate on responding to a modified Subjectivist view wherein the crucial turning point occurs when independent rational action becomes possible; that is, that the achievement of Practical Reason itself is the counterpart to the emergence of the butterfly from the cocoon. The importance of this developmental milestone has already been acknowledged, but this version of Subjectivism goes further and claims that it makes all the difference here.

In the first section of this chapter, I drew attention to our great behavioural flexibility compared to other animals. Generally speaking, species that evolve more in this direction also tend to exhibit higher degrees of 'neoteny': that is, the preservation of child-like traits into adulthood.[722] Hence, as well as being arguably the least specialised and most behaviourally labile of all extant species, human beings are also highly neotenous; the palaeontologist Stephen J. Gould thus writes that the 'evolutionary story' of our lineage is that of 'retaining to adulthood the originally juvenile features of our ancestors'.[723] In reply to this Sub-

[720] Tomlin, 2016; Tomlin instead begins from intuitions about the differences between child and adult well-being, and then argues from this to the caterpillar view.
[721] See also Skelton, 2016 from another statement of this view.
[722] Steiner, 2017, pp. 70-79; Bogin, 1997, pp. 66-67.
[723] Gould, 2008.

jectivist counter-strategy, then, we will focus on two particular neotenous features of human beings – each giving rise to a way in which we are like saplings rather than caterpillars.

Our first key neotenous trait of human beings is the profound capacity for learning discussed above. We have seen already that our great behavioural flexibility is inseparable from a heightened sensitivity to environing conditions as we mature, and considered the complex background prerequisites for humans to successfully develop into rational animals; we thereby also noted that the development of Practical Reason is thereby vulnerable to various setbacks if these are not in place.[724] Yet this potential for learning is also retained into our later years – even in those who appear to be stubbornly set in their ways, as expressed in the increasingly popular pedagogical concept of a 'lifelong learner'.[725]

Crucially for our present purposes, we shall see that this ongoing capacity for cognitive and behavioural change entails that there is also always a real ongoing danger that we human beings might lose our capacity to organise and rationally direct our own lives once this has been achieved. We therefore also exhibit another kind of vulnerability in addition to those we have considered so far, and consequently there is a further range of ways that things can go wrong in our lives from the standpoint of Practical Reason.[726] In particular, what we shall explore in the present section is how we are all to varying degrees susceptible to our Subjective Values themselves becoming distorted in ways that undermine our overall capacity for rational action.

In seeing how such changes in our preferences might occur later in life, we return to the 'adaptation principle' discussed earlier.[727] Recall that the core idea was that our affective responses tend to 'settle down' to a certain set level as we become used to new circumstances in our lives, with key early examples from the literature being winning the lottery and incurring a serious disability.[728] Consider now a more everyday case: suppose that circumstances arise wherein for an extended period of time we cannot avoid subsisting entirely on food that is bad for us because it is overly salty or high in fat – perhaps whilst travelling through a region where nothing else is available. Then even if we have been brought up

[724]Cf. MacIntyre, 2009b [1999], p. 68; see also Tate, 2020 for an analysis of pediatric suffering in terms of absense of the conditions needed for proper development.
[725]Field, 2000.
[726]MacIntyre, 2016, p. 27.
[727]Cf. Section 3.4 above.
[728]Brickman, Coates, and Janoff-Bullman, 1978.

with good habits whilst we are young, and initially have a dislike for eating such food in excess, we may become used to this new diet and have a strong wish to continue it once the initial circumstances are removed.

We turn now to a more dramatic case of wayward preferences; one whose impact is of special relevance to the present enquiry. Human beings are always to varying extents vulnerable to falling victim to various forms of addiction – most commonly, to substances such as alcohol or hard drugs, our desire for which may become overwhelming and all-consuming. Moreover, all too often the end result is a severe reduction of our powers of Practical Reason. For serious addiction may either prevent us from rationally reflecting on our conception of what is good for us, as our inability to adequately distance ourselves from these desires undermines our capacity for Practical Judgement; or else may impede our progress in moving towards such an ideal in practice, as our willpower fails us and we find ourselves now unable to retain cognitive control over our actions or lives as a whole.[729]

This decline of our rational powers is a condition which any Action-Guiding theory of the Human Good must enjoin us to avoid, on pain of being self-defeating.[730] Yet at the same time, we will not find the resources we need to overcome such challenges by looking inwards to our Subjective Values, since the problems stem from these same preferences themselves. For instance, our desire for an addictive substance may far outstrip all others, and its satisfaction typically brings a great deal of pleasure, too – at least in the short term, before the adaptation principle comes into effect or serious health problems develop.[731] Moreover, even if a hardened addict recognises their state as one of sickness and does not identify with their felt cravings, their Subjective Values may not point to any positive way out of their current predicament. Rather, if our powers are to be restored or remain intact, I shall argue that what is again needed is a certain kind of supportive social network on which we can rely.

In coming to see this, we turn to a second neotenous trait of human beings: the susceptibility to social influence that was an important theme

[729]See Clarke-Doane and Tabb, 2022 on addiction and rational agency; see also again MacIntyre, 2009b [1999], p. 69; p. 83; and p. 136 on the capacity for detachment from our desires as necessary for the proper discharge of our rational powers.

[730]Cf. Section 3.1

[731]On pleasure (or 'liking') coming apart from desire (or 'wanting') in jaded addicts, and the distinct neural circuitry underlying the two responses, see Berridge and Robinson, 2016.

of the first part of the chapter. Though perhaps less pronounced in mature adults, this second juvenile feature of our species is nevertheless also retained later in life; indeed, even here its impact on our actions is hard to overstate.[732] For instance, there is a huge empirical literature on conformity in our judgements, much of which is inspired by a classic series of experiments by Solomon Asch wherein individual subjects were made to give clearly-wrong answers to simple visual perception tasks by having the other participants – who were confederates of the experimenter – form a majority consensus.[733] Outside of the laboratory, our personal friends often influence us implicitly through our engaging in shared activity, and when we are getting to grips with a difficult decision – at the level of Life Planning, for instance – rehearsing the problem to one or more of our friends and gauging their reaction is often an important first response.[734] Moreover, concerning addictive substances in particular, we remain subject to various forms of 'peer pressure' that actively re-orient our behaviour – whether for good or bad.[735]

In the present social order, one often encounters a radical individualist attitude wherein to accept this social influence is seen as a sign of indecisiveness or even weakness: 'Surely you don't care what your friends think?!', it is thus sometimes exclaimed. And of course, the Subjectivist cannot but agree; ultimately, it must be our internal Subjective Values that move us to action rather than any external source, and friends can at most offer Instrumentally Valuable advice about how to satisfy these most effectively.[736] Yet social influence is so pervasive in our lives that any claim to follow this guidance without exception must be mere posturing – although we can critically reflect on its sources, we cannot reject it entirely:

> It is not an ethical "ought" that conduct should be social. It is social, whether bad or good.[737]

[732] See Sapolsky, 2017, pp. 165-166 for a comparison between adolescence and adulthood here.

[733] Asch, 1956; participants were asked to make apparently easy comparative judgements about the lengths of different line segments.

[734] Cf. Aristotle, *NE* III.3, 1112b10-11: 'We call in others to aid us in deliberation on important questions, distrusting ourselves as not being equal to deciding'; see also MacIntyre, 2009b [1999], p. 148 on rational justification proceeding through dialogue with friends.

[735] See Morris *et al.*, 2020 on alcohol; and Stacy, Newcomb, and Bentler, 1991 on sensation seeking.

[736] For a classic discussion of 'normative' versus 'informational' social influence, see Deutsch and Gerard, 1955.

[737] Dewey, 2002 [1922], pp. 16-17.

From a normative perspective, in the remainder of this chapter, I shall argue that being receptive to supportive social influence is necessary for us to remain competent rational agents in the face of the threats to our rational powers introduced above.[738] Despite their status as independent Practical Reasoners now being established, even grown adults still at times require from others the caring attitude of a good parent, teacher, or mentor here too – though now from supportive family, friends, lovers, colleagues, and others:

> Within this framework, one's gaze is not fixed, limited to her inner self and its depths. One's attention instead turns outward, understanding that flourishing is becoming a participant and steward of the network of giving and receiving that sustains life as humanly lived.[739]

The Aristotelian conception of the invulnerable *megalopsychos* is therefore a myth.[740] The need to rely on supportive others is not a mere temporary condition ..

> .. not just a fact about *genesis,* which can be ignored later on ... the contribution of significant others, even when it occurs at the beginning of our lives, continues throughout.[741]

In particular, maintaining the status of independent Practical Reasoners requires us to be open to learning from and absorbing the advice of others – even if on many occasions we should not listen to other people and do better to follow our own Practical Judgement instead, and indeed even if such others may sometimes be the *cause* of our problems.[742]

In the next section, we will consider three types of circumstances in adult life where we are especially vulnerable, and thereby must rely

---

[738]Other empirical hypotheses for the function of social norms are that they guide the navigation of complex social situations and thus reduce uncertainty about the appropriate way to behave; that they allow the actions of others to become predictable to us; and that they facilitate the coordination of individuals towards complex shared goals – in part by negating the need for a coincidence of private incentives.

[739]Snead, 2020, p. 98.

[740]Aristotle, *NE* IX.9, 1169b3-8; see MacIntyre, 2009b [1999], p. 127 for further criticism of this aspect of Aristotle's view.

[741]Taylor, 1992, p. 33; see also Bowlby, 1997 [1969], pp. 207-208 on attachment behaviour recurring in a new form later in life; and Waters and Cummings, 2000 on the import of attachment relationships beyond infancy.

[742]Cf. again MacIntyre, 2009b [1999], p. 71: 'My physician, or my trainer, if I am an athlete, or my teacher, if I am a student, may well be better placed to make judgements about my good than I am. And so on occasion may my friends.'

on the help provided by various medical professionals including doctors, clinical psychologists, and carers. In the remainder of this one, we will explore the above claims further by focusing on the place of friendship within adult life in general.[743] I will argue that with human nature being what it is, everyone has the ongoing potential to find themselves in need of friends in order to maintain their powers of Practical Reason – even though the extent to which this need is actually realised may vary, and some people may therefore function well for a time whilst leading quite isolated lives. I will then further argue that committing to our friendships in the manner required for our friends to adequately support us here – if and when they must be called upon – entails that we recognise certain Relationship Goods that extend beyond our Subjective Values.

We have pointed out that changes in our Subjective Values like those wrought by addiction may undermine our powers of Practical Reason, and I have further claimed that what is required to guard against this is a supportive social network. In particular, what we need in order to avoid falling victim to our own desires in such instances is the support of understanding and caring but sufficiently detached others on whom we can rely. And good friends are especially well-suited to this role, since they are close enough to us to understand our problems whilst not being directly under the sway of the distorting desires in question.

One sort of case where the benefits of friendship may be felt is when we have acknowledged that some aspects of our Subjective Values are not serving us well, but are nevertheless struggling to overcome the power of their motivational pull. Not feeling the same desires as we do, our friends can support us in behaving in a way consistent with our more considered views about what we want our lives to be like, rather than what we happen to feel compelled to do in the present moment – inclinations which we may not cognitively identify with even when the level of feeling they convey is overwhelming. For instance, friends may help us commit to such reasoned-out choices as sticking to a new diet, stopping smoking, or leaving a toxic relationship – all of which may be unpleasant or painful; especially in the short-run, whilst our Subjective Values are still adjusting to our new mode of living. And indeed, we often announce such planned changes to our friends as a means to ensuring that we persevere with

[743] See Helm, 2021 for an overview of some recent philosophical work on friendship; Aristotle famously distinguishes three kinds of friendship based on pleasure, utility, and virtue: see Aristotle, *NE*, VIII.

them.[744] Somewhat paradoxically, then, it is through being open to the influence of such supportive others that we can gain greater control over our impulses and instinctive behaviour than would be possible with our individual cognitive apparatus alone.

In instances such as these, the Subjectivist might plausibly claim that the assistance others may provide here can be of value to us only when it enables us to better satisfy our existing second-order preferences. However, in other cases, we may not be able to recognise the predicament we are in, and consequently these corresponding second-order preferences may never be formed. Indeed, we are particularly susceptible to this lack of clarity in our self-perceptions, due to the biases inherent to an introspective point of view. This systematic tendency toward error has been amply documented in psychology and behavioural economics over the past few decades: for instance, in the extensive literature on social desirability bias, wherein agents' perception of their own traits are more in line with what they would *like* to be true rather than what is really the case.[745] As such, we are always liable to be stuck in a 'closed loop' that our own inner resources cannot supply a way out of.[746]

Here friends can help us in a different way: by diagnosing threats to our rationality that we cannot recognise for ourselves. For the judgement of our friends is typically less clouded through direct personal investment and the emotional blinkering this can bring:

> it is often from others that we are able to learn how to correct the one-sidedness and partiality of our own particular point of view and to see things as they are, rather than as our desire-driven phantasies represent them.[747]

In order to maintain our status as independent Practical Reasoners, then, we human beings have need of friendships wherein our first-person judgements about what is good and best for us are open to being corrected from the standpoint of others. But if we are to thus benefit

[744] Rubin, Shmilovitz, and Weiss, 1993.

[745] See Krumpal, 2013 for a review; see also MacIntyre, 2016, p. 75.

[746] Cf. Skinner, 1965 [1953], pp. 260-261: 'There appears to be no way in which the individual may sharpen the reference of his own verbal repertoire in this respect. This is particularly unfortunate because he probably has many reasons for distorting his own report to himself.'

[747] MacIntyre, 2008, p. 290; cf. see also Dunne, 1993, p. 118 on Gadamer, and Schleiermacher's dictum that the historian understands events better than the contemporary actor whose deeds he describes in a detached way; and again MacIntyre, 2006a, p. 83 on others correcting our misperceptions of ourselves.

from these supportive friendships when our powers of Practical Reason are jeopardised by pathological or eccentric Subjective Values that need revising, we must listen and pay heed to our friends when they have valuable advice to offer us. And this requires that we must – in spite of some contemporary protestation – be open to respecting, taking seriously, and inevitably on occasion conforming to the values of these friends – even if they conflict with our own present judgements and standards:

> The basic requirement here is openness: a readiness to allow the questionableness of one's own contribution or to be persuaded that the other's contribution may enrich it or even have to prevail over it.[748]

This again runs contrary to Mill's individualist claim that when minimal conditions of competence are in place it is invariably the acting agent themselves who can best see what is good for her.[749] Moreover, in instead maintaining an inward-looking focus and attitude of Instrumentality toward others, our Subjectivist conceptions of the Human Good preclude us from rationally accepting the support we need in such cases of distorted preferences.

A Subjectivist may respond here that this is a fairly minor criticism that applies only in those special circumstances where we are especially liable to degeneration, or where the influence of distorted preferences is beginning to take effect – such as when we are well on our way to becoming addicted to a harmful substance. Moreover, even in such instances, the required openness and receptivity at the here-and-now level of Action Selection might be ratified by a prior endorsement at the level of Life Planning, if taking on this attitude promotes the aims of Subjectivism. However, as before, this response again fails. For an Action-Guiding view cannot coherently restrict endorsement of the requisite attitude to such cases only, because – as we have seen – we are often unable even to distinguish between these perilous situations where our Practical Reasoning is under threat from others where this is not so.[750]

The attitude of openness and receptivity required to guard against our rational powers going into decline must therefore be one that is always ongoing, rather than being adopted only when we judge that it happens

[748] Dunne, 1993, p. 118 (discussing Gadamer).
[749] See again Mill, 1989 [1859], p. 67.
[750] Recall here our discussion of the 'situated perspective' in Section 1.2.1 above.

to coincide with the satisfaction of our Subjective Values. We must stand ever ready to receive advice and constructive criticism from those friends whom we rightly trust, thus allowing them to provide the help we need on those occasions when – sometimes unbeknownst to us – we do need it.[751] Yet to hold this standing attitude is also to at least implicitly acknowledge that it is possible that what we desire or find pleasure in is, in fact, not what is good for us. And so there is again a divergence between any Action-Guiding conception of what is good for us that enables us to maintain our status as Practical Reasoners on the one hand, and what we as individuals happen to desire or enjoy on the other.

To summarise: the type of bifurcating Subjectivist view introduced at the start of the section neglects important facts about our continuing dependence on others later in life. Rather than a hard-and-fast divide paralleling the caterpillar's emergence from the cocoon as a butterfly, at which point our Subjective Values can take over the reins, what we find instead here is an ongoing vulnerability to losing our rational abilities, and consequently, an ongoing potential to need support from others at any time.[752] Moreover, this ongoing vulnerability entails a need to be ever open to re-educating our preferences through learning from supportive others – especially our friends. Yet our Subjectivist views preclude the possibility that it can be rational to accept such help, since in recognising only internal reasons for action, such views can neither make coherent sense of nor guide us toward sources of Intrinsic Goods other than those identified by our Subjective Values – including those Relationship Goods we need to rely on when our powers of Practical Reason are imperilled. Participation in such relationships in the required way is indeed only possible if we instead remain open to the possibility that the satisfaction of our existing Subjective Values – however deep-seated, well-established, or even higher-order – may in fact be bad for us; but to acknowledge this is of course also to draw a conceptual distinction that is incompatible with seeing our Subjective Values as constitutive of our own good.

In making these claims, I should again point out that I am not advocating a radical *anti*-Subjectivism either. It may be that much of the time we should be guided by our internal Subjective Values: for

[751]See again Davis and Finlayson, 2021 on the general issue of openness in an epistemic context. Such ongoing openness and respect may also be either causally or constitutively necessary for friendships of the required kind to develop, though I cannot enter into this matter here.

[752]MacIntyre, 2009b [1999], p. 128.

instance, sometimes our friends may hold us back by being resistant to a positive change in our character or lifestyle – perhaps because of a preference for continuing to see us in a way that is familiar to them.[753] And it is undeniable that we do heavily rely on these inner sources of motivation for guidance in practice. But nevertheless, we must be wary of suffering from the attitude of holding on too tightly to our existing vision of how we should live and whom we would like to be if this ideal is no longer working for us, and instead be open to revising some of the values, principles, and preferences we have adopted so far.

Finally, although rationality might require us to only accept outside influence in a critical way, it is worth pointing out that its preservation may also require us to select our friends in a manner that is counter to our existing Subjective Values rather than in line with them.[754] In particular, in choosing our friends wisely, we might need to place our trust in those who put in check rather than encourage our worst impulses or proclivities, or who are able to stage interventions if and when these are needed. For the case of addiction in particular, in the literature a 'sponsor' is distinguished from a mere enabler, as comforting and pleasing as the latter may be, and the kind of friend that can best support someone struggling with addiction is recognised as being one that has the resources to help them re-forge a new conception of the good life without the addictive substance, thus effecting a lifestyle change that may run against their existing preferences at first.[755] But to accept this is of course to concede that our Subjective Values cannot be the focal point or Final End of our deliberations at the level of Life Planning.[756]

In this section, we have focused on challenges that may always be encountered by any adult human being, as part of life in general, and how an ongoing openness to accepting help from supportive others and especially our friends is important for protecting our status as Practical Reasoners from the threats these challenges pose. In the following, final section, I develop the argument by considering the heightened vulnerabilities experienced in physical and mental illness and in old age.

[753]See again Rubin, Shmilovitz, and Weiss, 1993; see also the end of Section 3.1 above, and again MacIntyre, 2009b [1999], p. 96 on independence of mind.

[754]Cf. Aristotle, *NE* X.3, 1173b32-1174a1 on friends versus mere flatterers, and Scanlon, 1993, p. 194 on our benefactor's conceptions of our well-being needing to extend beyond what we happen to find pleasuring.

[755]See Henry and Robinson, 1978 for some early data on this, drawn from participants in 'Alcoholics Anonymous'.

[756]Cf. Section 1.7 above.

3.6 Physical Infirmity, Mental Illness, and Old Age

This final section will be about three broad types of adverse circumstances in which human beings may find themselves. These are: the debilitating effects of physical injury and disease; the threats to our rationality that may accompany serious mental health problems; and the inevitable impositions of old age. Under these conditions of increased vulnerability, we shall see that our need to draw on the support of others in order to maintain or restore our powers of Practical Reason becomes even more acute. We will address ways that human beings can attempt to navigate themselves through the pitfalls that arise here: both in a preventative mode, and once we have already fallen into such a fragile condition – which may occasion a return to the radical dependence characteristic of early childhood.

One recurring observation will be the converse of a central claim of the previous section: that even if our conceptions of what is good and best for ourselves are initially viable, when we are exposed to such adversity, we may find that these same ideals are now poorly matched to our new situation; that our values or Life Plans as we have construed them hitherto are now unrealistic and cannot be effectively put into practice. Moreover, we will see that such threats to Practical Reason cannot always be met through appealing to our existing Subjective Values themselves for guidance; indeed, attempting to do so may even hasten the decline of our powers.

Contrary to Subjectivism, I shall argue sometimes when our aspirations and expectations for life are thus put in check by changing circumstances, our Subjective Values themselves must adapt and reform in response if we are to find a way of moving forward as independent Practical Reasoners. And as in previous sections, I shall again emphasise our need to be open and receptive to the influence and advice of others here – with the focus now extending beyond our family and friends to a variety of medical professionals who can offer us a different kind of support and from whom we can also learn.

Since many of us are fortunate enough to remain free from such states of increased vulnerability for much or even all of our adult lives, the Subjectivist may again try to sidestep this argument by explicitly restricting the application of their account so that agents facing these

challenges fall outside its scope.[757] Yet with the frailty of human life being what it is, this condition is nevertheless one to which we are all liable to be reduced at any time. There is therefore always an ongoing potential for any of us to require intensive support from our surrounding community:

> ... those who are not yet disabled by age recognize in the old what they are moving towards becoming and that those who are not ill or injured recognize in the ill and injured what they often have been and will be and always may be.[758]

Moreover, I shall further argue that for an agent to fully acknowledge the possibility of and adequately prepare for such eventualities, they must not commit to identifying their good with either what they desire or what brings them pleasure. But to understand this is again to recognise the inadequacy of our Subjectivist accounts more generally.

3.6.1 Physical Impairments

We begin by considering problems with our bodily health – including disease, injury, and disability.

Despite the impressive advances of modern medicine, many humans still suffer serious illness or other physical impairments at least once in their lives. If this occurs, then we may also become so distracted by pain or fatigue that our capacity for rational deliberation at the level of Action Selection is severely impaired. And such eventualities may also undermine our capacity to put into practice our Life Plan as conceived hitherto – a broken leg leading us to abandon a promising career as a sprinter, for instance. However, the chances of many kinds of serious illnesses can be reduced by preventative factors, such as a good diet and adequate exercise. In particular, we have a higher chance of remaining healthy if we listen to the advice of doctors, dieticians, and other experts about what responsible living consists in.

For some fortunate individuals, there may be no tension between what they must do to remain physically healthy and what they desire or take pleasure from. Others, however, may hold a view of their own good that is quite different from that of their doctors – a sedentary one, perhaps, or one of overindulgence in alcohol and cigarettes or drugs, or

[757] Cf. MacIntyre, 2009b [1999], pp. 1-2.
[758] MacIntyre, 2009b [1999], p. 146.

centred on the consumption of lavish food. And where this is the case, engagement with such medical professions thereby becomes a potential source of conflict.[759] Sometimes such differences of opinion and resulting quarrels may be ineluctable and exist as part of an ongoing functional relationship: for instance, I may find it hard to care about the health of my teeth quite as much as my dentist thinks I should, given the centrality of this concern to her professional life, but nevertheless be moved further in this direction by her reproaches. However, with more radical divergences of opinion this may not be so, and the question of whose authority is to prevail when our preferences thus come into conflict with our physical health becomes an important one.

For the Subjectivist, of course, our individual Subjective Values have final authority and our mode of engagement with medical professionals can be rational only insofar as this ultimately leads to more effective satisfaction of our desires or affective preferences.[760] Yet in practice, a shock encounter with a serious health problem that is primarily a consequence of these same preferences can forcibly bring an individual to acknowledge that their existing conception of what is good for them is inadequate and needs revising. Such an experience – commonly referred to as a 'wake-up call' – may indeed feel as unequivocal a falsification of their conception of the good as can be found for any hypothesis of the natural sciences.[761]

Later on, such individuals may understandably wish that this revelation could have occurred less dramatically, before such a decisive and momentous point had been reached. Yet this is possible only if we allow ourselves to be teachable to at least some extent. In particular, if we are to be open to learning from the relevant professionals that other, more positive conceptions of how to live will serve us better, and to making rationally-grounded adjustments to our habits accordingly, then we cannot take our existing Subjective Values to be constitutive of what is good for us.

The Subjectivist theorist may appeal to our standard response here: that if we are to be rational in coming to recognise that a lifestyle we are implicitly motivated towards is bad for us, this can only be

[759]Cf. MacIntyre, 1979, p. 47: 'their relationship embodies or may embody *an argument* about sickness, health, expertise, drugs and many other topics.'; see also Plato, *Gorgias*, 505a.

[760]For a contrary view, see again MacIntyre, 2009b [1999], p. 71.

[761]Cf. MacIntyre, 2016, p. 40 on conceptions of the good life being falsified through experience.

because we have discovered that it turns out to contradict certain deeper preferences at a more abstract level; for instance, our desire for the general state of 'being healthy'. But for this line of response to be convincing, it must be more than a possible *post hoc* interpretation designed to maintain the thesis at all costs.[762] The only preferences of import here are those which are actually available for guiding the agent's deliberations – and in practice, there are indeed cases where an agent appears to have no interest in moving towards this general healthful condition, whereas dangerous levels of overindulgence in drink or overly rich food may be an essential part of the good life as they presently conceive it. So, although such individuals might later adopt desires and affective preferences that are more conducive to maintaining their powers of Practical Reason, if these are not available prior to this learning having occurred and are at this stage merely hypothetical, they cannot yet be Action Guiding. And following the preferences they presently do have will instead lead systematically to a decline of their rational powers.

Sadly, even if the best lifestyle advice is followed to a tee, it is still possible for even a fit, healthy, and temperate person to contract a serious disease, and even if we are careful debilitating injuries may still befall us at any time. Though minor physical impairments may allow us to keep our powers of Practical Reason more or less intact, often there is a partial loss of independence at this stage as we find we must increasingly depend upon others to support us. In particular, although recovery may be possible, when this is the case we may find that in order to thus regain our powers we must set aside our current desires and preferences and openly listen to the advice of professional carers about what we must do to get better – even if this does not coincide with what we would most like to do or would take most pleasure from doing: for instance, abstaining from foods that we enjoy, taking unpleasant medicine, or undergoing painful operations.

Unfortunately, with the motivational power of our Subjective Values being what they are, in practice, these precepts are not always followed, and many patients do not recover due to not following their doctors' directives.[763] Moreover, in other cases, the possibility of getting better is sadly not available regardless of what course of treatment we follow

[762]Cf. Section 1.7.

[763]Corr and Plagnol, 2019, p. 157; see also Kimmel *et al.*, 2012 for discussion and a suggested remedy (specifically, a randomised monetary payout) for individuals with poor adherence, in the context of anti-coagulant (blood thinning) medicine.

or lifestyle adjustments we make. Hence, many patients are confronted with the ordeal of living with chronic illness. In learning to cope here, we may again find that we need to modify and change our preferences and expectations – not only to later feel happy once more, but also in order to embrace and maintain the measure of independence we are capable of in our new situation:

> Circumstances beyond individuals' control may have dealt them a cruel blow, but they can retain dignity as self-determining agents capable of responsible choice in directing and retaining control over their lives within the limits that their new circumstances permit.[764]

However, this process of change and adaptation cannot be rationally negotiated by an agent who takes their existing preferences to be constitutive of their own good.

3.6.2 Mental Illness

We now discuss debilitating ailments that are mental rather than physical in nature.

As with physical illness, mental health problems may also engender a loss of our Practical Reason – either leading us to develop new conceptions of our good that are unrealistic or even incoherent, or else undermining our ability to bring about our existing ideals in practice. Here I will focus on the role of therapists and counsellors in enabling us to remain or once more become capable of independent Practical Reasoning in the face of such threats.[765] I will emphasise how proper engagement with the therapy process requires a recognition of our need to learn from these professionals and an openness to permitting the education and transformation of our present preferences through the Relationship Goods they can provide. If our condition is to improve, we must therefore be guided by resources beyond our Subjective Values; moreover, since none of us is ever immune from coming to need this kind of help, we will find here further reasons why Subjectivist approaches are inadequate in general.

[764] Brock, 1993, p. 125; cf. Etkind *et al.*, 2018 on the impact of aiming at 'getting back to normal' versus 'finding a new normal' following illness and the role of a supportive social environment in forming new preferences.

[765] See here Waller, 2022 on talk therapy improving reason-responsiveness and powers of agency in general; and Biegler, 2019 on 'autonomy' and effective decision-making as a key goal of psychotherapy.

We begin by considering how engagement in therapy must appear on a Subjectivist view. Here a client can rationally choose to participate in therapy only as a means to satisfying their existing Subjective Values, by whose lights they must ultimately be guided – even if for contingent reasons the expression of these preferences is currently problematic in ways that they hope the therapy process to remedy.[766]

Of the major existing approaches to therapy, such a conception fits best with the school deriving from Carl Rogers, variously called 'person-centred therapy', 'client-centred therapy', or 'non-directive therapy'.[767] Within this optimistic approach, clients are encouraged to 'make use of their own internal feelings' rather than being 'guided by externally defined sets of beliefs and attitudes', echoing Rogers' faith that these inner values will in general be accurate and deserving of trust.[768] Ultimately, it aims to move individuals towards having a so-called 'internal locus of evaluation' and an 'internally-directed mode of conduct', which correspond roughly to adopting the Subjectivist standpoint in its Appraisive and Action-Guiding guises, respectively. Therapists, meanwhile, are dissuaded from giving substantive life advice, and it is argued that they should instead play only a minor facilitating role, enabling the client's resolution of their own problems through reliance on these inner resources, or 'strengths'.[769] Rogers famously stated a number of core conditions that characterise a productive therapeutic relationship thus construed – such as the therapist having 'empathetic understanding', and 'unconditional positive regard' for the client – and argued that these are not only necessary but also sufficient in the sense that as long as they are met the progress of the client is virtually inevitable.[770]

The remarkable congruence between this approach to therapy and our Subjectivist theories is less surprising if we consider that both are the product of the same cultural forces and contemporary Western values.[771]

[766]Cf. Sripada, 2022 on mental disorders as problems in our regulation of 'spontaneous' behaviour that cause our actions to be misaligned with our preferences.

[767]For an influential early statement, see Rogers, 1951; for a more recent presentation, see Mearns and Thorne, 2013.

[768]McLeod, 2009, p. 176.

[769]Mearns and Thorne, 2013, p. 7-8; p. 13; McLeod, 2009, p. 168; for a broadly Subjectivist person-centred view of what makes a good counsellor, see Combs, 1989. The term 'counsellor' was invented by Rogers for licensing reasons; I will use the term 'therapist' throughout what follows, even though the two roles are somewhat different today.

[770]Rogers, 1956; see McLeod, 2009, pp. 178-179 and p. 181 for further details and discussion.

[771]See McDougall, 2002 for the view that this approach is tied to contemporary

Yet despite this agreement, the approach's ambition to effect positive change using only the inner resources of the client is in some cases quite unrealisable. Ultimately, this was demonstrated in undeniable fashion. In the five-year 'Wisconsin Project', beginning in the late 1950s, Rogers and his colleagues attempted to apply the person-centred approach to individuals with schizophrenia at the Mendota State Hospital – and the disappointing results made it clear that the optimistic view that clients always possess inner resources sufficient to overcome their own mental health problems and merely need someone to 'unlock' and mobilise these is wholly inappropriate when these problems are sufficiently profound.[772] And indeed, following this incident, it was not long before this once-popular approach went into decline.[773]

For our present purposes, the importance of Rogers' experience with these people with schizophrenia is that it reveals the objective nature of serious mental health conditions and the problems in living they generate. We have seen that on a Subjectivist view of Practical Reasoning, it can only be rational to engage in therapy for Instrumental reasons – to get what we presently want, or improve how we presently feel.[774] However, if we attend to the 'case formulations' produced by practitioners of psychotherapy and counselling, which detail the challenges clients face and the goals their therapy sessions should set out to achieve, often these are not the kinds of problems we find described.[775] Rather, many case formulations instead describe intrinsic problems with clients' desires and affective responses themselves.[776] And where this is the case, a return to the status of independent Practical Reasoners may not be possible unless these wayward Subjective Values first undergo modification.[777] Yet this

American values and ways of life, and consequence warnings about its lack of relevance to other social milieus.

[772]Rogers later described this as 'without doubt the most painful and anguished episode of my whole professional life' – Burton, 1972, p. 62; see also Gendlin, 1966 for a general account of the research programme.

[773]See Quinn, 2015, pp. 82-96 for a broader history of the decline of the tradition, both theoretically and within clinical practice.

[774]See again Sripada, 2022.

[775]For an influential CBT approach to case formulation in terms of a list of the client's *problems* and symptoms; hypothesised underlying *mechanisms*; current *precipitants* that activate these mechanisms, and the *origins* of these contingencies, see Persons, 2008, Chapter 6; see also *ibid.*, Chapter 7 on moving from here to a concrete treatment plan and identifying potential obstacles to success.

[776]See Ellis, 1994 [1962] for a classic statement of the view that the behavioural problems therapy treats stem from emotional problems caused by 'crooked thinking' (p. 377) – a notion somewhat similar to the concept of 'distorted preferences' introduced in the previous section.

[777]Indeed, some clients may doggedly hold on to the view that asking for help is

is a process which will typically run counter to the satisfaction of these preferences in their present form – as reflected in the common adage amongst therapists that there is a big difference between 'feeling better' and 'getting better.'[778] Hence, it cannot be entirely guided by these same Subjective Values – and indeed, if the client understands their own good in terms of the satisfaction of these, then this generates resistance to the process of healing and change. For these outcomes can neither be rationally pursued nor even adequately comprehended by one whose practical deliberations are centred on Subjectivist ends. Moreover, serious mental health problems often also bring with them systematic failures in the client's cognitive processes that severely hamper their overall ability to direct this re-education of their preferences themselves.[779]

When such problems in functioning are present, I shall argue that the best way for clients to preserve or restore their powers of Practical Reason is not to rely solely on the 'closed loop' of their own distorted preferences and judgements, but rather, to instead place their trust in a capable therapist, and acknowledge and make use of the Relationship Goods they have to offer.[780]

By definition, someone in need of therapy has a problem in functioning that they cannot overcome entirely by themselves: the therapist must supply *something* of value here, since otherwise, there would be no need for a therapist at all.[781] And this is indeed recognised by the person-centred school's insistence that a supportive environment is needed for the client to achieve progress. Yet for the same reasons that the 'strengths' model turns out to be too optimistic as a means of guiding this progress, Rogers' corresponding minimal conception of the therapist's contribution has also proved inadequate in the case of sufficiently profound impairments:

> It was largely as a result of the Wisconsin project, during

itself a sign of weakness; yet this is a preference that must be overcome for progress to be made.

[778] For a book-length treatment, see Ellis, 2001; see also Keyes and Annas, 2009 on the Aristotelean distinction between 'feeling' versus 'functioning' and its importance for clinical practice.

[779] Classic examples are overgeneralisation, dichotomous thinking, and personalisation; see Kovacs and Beck, 1978 for some early insights into cognitive distortions induced by depression; and see also Westbrook, Kennerley, and Kirt, 2007, pp. 59-60 on limitations in clients' understandings and resulting problems for case formulation.

[780] Cf. McLeod, 2009 p. 160 on the need of clients who currently lack their own 'coping mechanisms' to rely on the therapist.

[781] See here McLeod, 2009, p. 19.

> which Rogers, Shlien, Gendlin and their colleagues struggled to find ways of communicating with deeply withdrawn schizophrenic inpatients, that it became apparent that the therapist's contribution to the process, his or her ability to use self in the service of the relationship, was crucial to the success of therapy.[782]

In recognition of this, in the intervening period, a greater emphasis has now come to be placed on therapists' substantive contributions to their clients' progress. In particular, in opposition to the person-centred view that the client must be 'regarded as the expert on his or her own life and problems',[783] it is now generally acknowledged that the therapist can offer insight into the client's own life that they are not presently able to see for themselves – echoing the rival psychodynamic approach's long-held insistence that we may not be consciously aware of the true causes of our behaviour.[784] Moreover, alongside recognition of the difference between 'feeling better' and 'getting better' has also come an acknowledgement amongst practitioners that an exclusive focus on achieving what a client happens to presently want or would presently be most pleased by may be to ignore the objective 'problems in living' they are actually facing but may not yet fully understand.[785]

Whilst therapy is ongoing, then, what is needed is not a pure inner directedness, but rather a collaborative partnership, wherein through shared reflection the resources supplied by both parties may be put to work on identifying and resolving the client's problems: what has come to be termed the 'therapeutic alliance'.[786] And following this recognition, over the past few decades academic interest in the nature of this client-therapist relationship itself has also risen.[787] Moreover, one

[782]McLeod, 2009, p. 187. See also Sollod, 1978 for an early argument against 'Rogers' extreme position against authority'; Frankel and Sommerbeck, 2005 for the view that Rogers radically changed his conception and practice of therapy as a result of the encounter; and Truax, 1966 for an early empirical study analysing tape recordings of a case handled by Rogers to argue that so-called 'non-directive therapy' does in fact involve direction after all, via differential reinforcement of the client's responses.

[783]McLeod, 2009, p. 168.

[784]McLeod, 2009, p. 94.

[785]See again Ellis, 2001 for a thorough treatment.

[786]See Bordin, 1979 for a pioneering early contribution; Muran and Barber, 2010 for some recent work on the therapeutic alliance and guidance on its implementation within clinical practice; and Martin, Garske, and Davis, 2000 for a review of the empirical literature on its decisive impact on client outcomes.

[787]McLeod, 2009, p. 189; p. 200; for a detailed discussion by an influential practitioner, see Persons, 2008, Chapter 8; for another discussion of empirical

core conclusion drawn in this literature is that – in keeping with the above remarks, but *pace* Rogers' individualistic approach – trust in the therapist and openness to listening to and learning from the substantive advice they offer are indeed key parts of a functioning therapeutic relationship; that 'empowerment comes from being prepared to relinquish power and control, to trust the psychologist and follow her instructions'.[788]

To fully participate in such a partnership in the required manner, then, the patient must therefore see the contributions of the therapist as Relationship Goods that are of more than merely Instrumental Value. Yet this is something which a committed Subjectivist at the level of Action Selection can never acknowledge.[789] And the Subjectivist cannot respond by retreating to the level of Life Planning and arguing that clients should decide to engage in this committed way only if it is a successful means to achieving Subjectivist ends in the long run. For amongst those most in need of therapy, the perspicacity required for deliberating in this sophisticated manner is again generally lacking.[790] Moreover, if the relationship with the therapist is honest and open enough to be effective, then if this stratagem remains active then it too will be brought to light and examined, and ultimately exposed to possible evaluation and change.

We close the section by briefly looking at a different therapeutic approach in action: cognitive behavioural therapy (CBT).[791] Amongst the most empirically supported approaches to talk therapy, CBT is a practically oriented, 'problem-focused' form of treatment that has been influenced by Stoicism through its founders Albert Ellis and Aaron Beck.[792] Rather than following person-centred therapy in holding an 'uncondi-

evidence on the importance of the therapeutic relationship to successful outcomes, see Patterson, 1984.

[788]McLeod, 2009, p. 150; cf. Fairbairn, 1958, p. 380: 'The aim of psychoanalytic treatment is to effect breaches of the closed system which constitutes the patient's inner world, and thus to make this world accessible to the influence of outer reality'.

[789]Indeed, one of Roger's core conditions is that the client is in close 'psychological contact' with the therapist; yet this in itself would seem to preclude a purely Instrumental relationship; see Rogers, 1956. For a book-length treatment beginning from a person-centred perspective but diverging from this in the direction of the present argument, see again Barrett-Lennard, 2013.

[790]Cf. the similar arguments in the preceding four sections; especially 3.3.

[791]On other, newer approaches to therapy – such as relational, systemic, feminist, and narrative – which have moved away from the Subjectivism inherent to some older traditions, see McLeod, 2009, p. 219: '[These approaches] have each in their own way taken up the challenge of a systemic, relational philosophy, and have applied it with differing emphases, but with the same implicit understanding that in the end individualism is not an adequate basis for living the good life.'

[792]See Robertson, 2018 for a philosophical discussion.

tional positive regard' for the client and their present Subjective Values, CBT instead invites them to critically reflect on their current behaviours, emotions, and thoughts – often through supportive but potentially challenging Socratic dialogues – and to thereby bring them to an acknowledgement of the need to modify these where they have become dysfunctional.[793] In such cases, CBT therapy then aims to effect deliberate changes in the client's preferences, rather than taking them as statically given. Hence, like physicians, CBT practitioners do not merely enable their clients by allowing them to do whatever they happen to want to or feel like doing – and the kind of productive relationship envisioned between client and therapist therefore goes beyond mere facilitation.[794]

As an example of the mechanics of CBT treatment, we consider some practical techniques for curing phobias – such as an intense fear of even otherwise harmless spiders.[795] Such phobias are of special interest to the present inquiry, since the overwhelming fear response they provoke may have a debilitating effect on our cognitive control of our own behaviour, or else pose a serious threat to our ability to make rationally-grounded Practical Judgements about what is good for us.[796]

When a client is thus faced with an irrational fear, a core CBT technique is to arrange for them to be exposed to a series of related stimuli that are gradually more intense, as indicated by an ascending 'fear hierarchy' that they have helped construct. Then, relaxation techniques are used to allow the client's aversive reactions to become 'extinct' at each step. For instance, beginning with viewing a drawing of a cartoon spider, the client may graduate to holding a detailed photograph of a spider, and from there perhaps to viewing a moving spider on live video, with these progressively more challenging tasks culminating in them holding a real spider. This process is known as 'systematic desensitisation'.[797]

[793]McLeod, 2009, p. 154; see also Carey and Mullan, 2004 on Socratic questioning as an aspect of therapy, and Shlien, 1992, p. 1083-1084, who quotes Rogers' later tentatively skeptical remarks about unconditional positive regard following the Wisconsin Project: 'Very tentatively it appears to me at the present time that, in dealing with the extremely immature or regressed individual, a conditional regard may be more effective in getting a relationship under way, than an unconditional positive regard'.

[794]See again Plato, *Gorgias*, 505a.

[795]For a review of some CBT treatment guidelines, see Rothbaum *et al.*, 2000; for some recent evidence of its effectiveness in the context of PTSD, see McLean and Foa, 2014.

[796]Again, for a recent argument that treatment enables patients to 'respond better to reasons', see Waller, 2022.

[797]The practice originates with Joseph Wolpe; see Wolpe, 1958.

Such treatments have proved highly effective against phobias and the threats to Practical Reason they mount; yet because they intentionally effect changes in the client's responses, the goal of these techniques cannot be merely the satisfaction of their existing Subjective Values. Rather, following the prescribed course of action will, in general, go against the direction of the client's preferences in their initial phobic state – so that permitting themselves to be led through this often-terrifying series of steps also requires a solid and trusting relationship with the therapist.[798]

In order to rationally choose to engage in this process, then, and thereby ultimately become more rational in their responses when encountering the feared stimulus in ordinary life, clients must recognise and value the Relationship Goods that can be achieved through submitting themselves to the direction of the therapist. Yet doing so requires them to make a conceptual distinction between what they happen to feel like doing or would be most pleased by doing on the one hand, and what an appropriate and rational response to such a situation would really be on the other.[799] Moreover, although there are radical individual differences in how phobias actually affect us in practice, we are all to some extent liable to developing such irrational fears, due to the ongoing possibility of our being exposed to a sufficiently traumatic learning experience. Hence, this distinction is one that all rational agents must draw in order to be able to acknowledge the help they may always need in remaining capable of independent Practical Reasoning.

3.6.3 Old Age

We now come to one final context of practice: when our cognitive and practical capacities diminish with the onset of old age.

Even if an agent manages to avoid the physical illnesses and mental health problems discussed so far, the slow process of growing old typically presents similar challenges to Practical Reason. Although some individuals display remarkable longevity, both mental and physical, in most cases a person who lives to a very advanced age will thus find themselves attempting to preside over what is at best a graceful decline of their abilities.[800] Here I shall again argue that if we are to retain our

[798]Cf. McLeod, 2009, p. 156: 'CBT encourages clients to face the fear directly ... The aim is the replacement of anxiety or fear responses by a learned relaxation response.'

[799]McLeod, 2009, pp. 155-156.

[800]As instances of each type of longevity, the legendary biologist Ernst Mayr

powers of Practical Reasoning as best we can in these circumstances, then we must supplement our existing inner sources of direction by learning from and being open to the influence of those supportive others on whom we may come to depend.

As with physical infirmity, a key limitation of Subjectivist views here is that as we age and move through the different stages of life, there will likely be occasions where our existing preferences no longer match our present circumstances in the way they once did. For instance, changes in our metabolism and physique as we move toward old age may mean that our existing preferences for food and exercise are no longer appropriate for us, and that we can no longer remain healthy whilst continuing with the lifestyle to which we are accustomed. Where this is so, we may again face a choice between either good health and physical functioning or the satisfaction of our current Subjective Values. And as circumstances thereby push back against our existing conceptions of our own good, we may realise that our preferences must be reshaped in response to our current situation if our rational capacities are to be maintained. Hence, the initial task of developing ourselves as Practical Reasoners with Coherent Preferences and Life Plans may always need to be redone at any future point.[801] Yet as by now been made clear, rationally engaging in this process is not possible if we understand what is good for us in terms of the satisfaction of our existing Subjective Values, whatever these may be. Rather, navigating these challenges requires instead a certain graceful humility and a willingness to make concessions when required:

> What mattered was to humble himself, to organize his heart to match the rhythm of the days instead of submitting their rhythm to the curve of human hopes.[802]

A couple of concrete examples will illustrate the point.

First, consider the case of Joe, who is well into his eighties. Sadly, Joe now struggles to walk unassisted; yet in his prime, he was an accomplished mountain explorer. Despite his present condition, what Joe desires more than anything is to climb Everest one last time; he talks of little

published a well-received book after his 100th birthday – Mayr, 2004; and the late Stanisław Kowalski competed in the 100m, shot put, and discus events at the 2015 Polish Veterans Championships at age 105.

[801]MacIntyre, 2016, p. 75: 'these tasks of disciplining and transforming our feelings have to be undertaken again and again at different stages in our lives.'

[802]Camus, 1972; cf. Section 3.3 above on having a 'growth mindset'.

else, and does not appear to recognise his present physical limitations.[803] Moreover, an alternative specification of his desires – given in terms of achieving any great accomplishment, say – may not be recognised by him; his aims are tied to it being Everest specifically which is conquered once more. And it may also be true that the climb would fill Joe with lasting pleasure – *if* it were successful. But it does not follow from such facts about Joe's Subjective Values that the climb should be attempted. What may now be best for Joe is to instead revise his expectations and seek out new and hitherto unrecognised possibilities of living a satisfying existence that lie within the constraints imposed by his advancing age; of achievement on a more modest scale, or in a different direction – perhaps connected with mountain climbing, but in an advisory role.[804]

Consider now a quite different kind of case. Joan, an older woman also in her eighties, has for many years built her life around her husband Brian, who has now died – a loss which has brought her profound sadness and left her skeptical about the possibility of rediscovering happiness.[805] In finding something new to live for and deciding how to make the best use of her remaining time, Joan may draw on a number of sources: supportive family, friends with similar experiences, or even imaginative literature. But it may be that she is not able to find these answers simply by looking within herself and consulting her existing Subjective Values. Indeed, in the absence of her life partner, whom they have been formed around and in partnership with, these may even no longer make coherent sense in their present state.

In situations such as these, our Life Plans seem to have become stalled, and our existing inner values fail to provide us with the resources we need to direct us forward. But the discussion of the latter case also raised another possible source of inspiration here; one which brings us back to the social context of practical agency. For as with the initial development of our powers, our working out a way to move forward is part of a larger journey that is to a great extent social. In particular, just as we initially need others to form Coherent Preferences at the beginning of our lives, so too may we also need to draw on certain Relationship Goods in order to acquire new and more appropriate Subjective Values

[803] The current record-holder, Yuichiro Miura, who achieved the feat at age 80, reports that the mountain 'calls to him', and reportedly plans to attempt the climb again at age 90: see Henderson, 2016.

[804] Cf. MacIntyre, 2016, p. 7 on finding new and hitherto unforeseen possibilities for how to live out one's life.

[805] For a discussion of resilience and identify reformation in widows, see Evans, 2001.

as time goes on. However, as we saw with some forms of mental illness, in some cases this lack of adequate inner direction in old age may also be combined with a general cognitive decline wherein this need to outgrow their preferences is not realised by the agent themselves. As with young children, then, the required Relationship Goods cannot be viewed in merely Instrumental terms here: being thereby inaccessible through explicit Means-Ends Reasoning, they will only be available to those older people who hold a general ongoing attitude of openness and receptivity to those who care for them.

In cases of extremely advanced age, the agent's capacity to carry out their own decisions may diminish to such an extent that the third-person conception of their good now becomes regnant once again. This is something of a regress to the position of dependence of an infant or young child, and the final stages of life may therefore in some ways resemble the earliest ones.[806] There is a vast literature concerning medical professionals' decision-making on behalf of those in a condition of incapacity due to senescence, which I cannot hope to adequately summarise here.[807] But what is clear from this research is that the patient's present preferences, whilst important, cannot be the sole consideration to which everything else is subordinate and from which all decisions about how to care for them may be derived. For they may provide only a partial, contradictory, or indeterminate answer here; or else no answer whatsoever.[808]

Although once we are actually in this severely diminished condition we may not be able to exercise Practical Reason at all, prior to reaching this point we do have the opportunity to make adequate preparations for our potential future need for such extensive care – including entering into personal or familial relationships that are sufficiently intimate that these

[806]Cf. Shakespeare, *As You Like It*, Act 2 Scene 7; see also Bowlby, 1997 [1969], p. 207 on attachment behaviour in old age.

[807]For a recent review on the processes by which GPs and other healthcare professionals working in primary care make decisions on behalf of incapacitated patients in light of their 'best interests', see Ogden, Huxtable and Ives, 2020; see also Toomey, 2021 on seniors with dementia and supported decision-making and our understanding of their agentic capacities; Jaworska, 1999 for a philosophical discussion of basing decisions on current versus past preferences of Alzheimer's patients, as championed by Rebecca Dresser and Ronald Dworkin, respectively; and Hawkins, 2021 on more complex cases where significant mental capacities other than decision-making are retained but past preferences are no longer appropriate guides to care.

[808]On the inconsistency over time of preferences in end-of-life care in the elderly, see Fried *et al.*, 2007; the authors find that 'the frequency of inconsistent trajectories even among those with stable health states suggests that preferences are influenced by transient factors rather than representing stable core values' – p. 1007.

others are later able to take on the role of our proxies if we become unable to speak for ourselves. But since their discharging this role effectively requires these others to understand our deepest principles and to care about us for our own sake, which in turn requires that our association with them must have a mutual closeness that moves it beyond the merely Instrumental sphere; and since we must also permit such carers to act on substantive values that may extend beyond what can be deciphered about our later wishes, we cannot enter into such relationships whilst fully committing ourselves to a Subjectivist view of the Human Good.[809]

Finally, on a Subjectivist account, getting older must appear to the young to be unequivocally bad, because of the consequent incapacity to satisfy their current preferences. However, if we are able to make apt use of the Relationship Goods provided by a supportive social context and thus effectively revise our preferences in line with our changing circumstances, old age may be felt as far less of a burden; indeed, in reaching this point of maturity, we may even perceive certain benefits which one who understands their good in Subjectivist terms must fail to recognise:

> For in very truth there comes to old age a great tranquillity in such matters and a blessed release. When the fierce tensions of the passions and desires relax ... we are rid of many and mad masters ... [810]

3.6.4 Subjectivism and Vulnerability

In summary, human beings are always vulnerable to both physical and mental ailments, and the diminishing effects of old age cannot be avoided entirely. Moreover, these ordeals are not rare exceptions, but something that most of us will, sooner or later, be brought to face. And in making the adjustments to our Subjective Values necessary to meet these challenges to our status as independent Practical Reasoners, we will often need to draw on supportive others within our community. Underlying an apparent similarity to butterflies, we therefore see a commonality with

[809]See MacIntyre, 2009b [1999], Chapter 12 for detailed discussion of the ethical importance of proxies and the attitudes we must hold earlier in life in order to form the relationships required for this role to later be properly discharged; see also Section 3.2 above on intimate relationships and their incompatibility with holding an Instrumental view of others.

[810]Plato, *Republic* I, 329b-d.

the sapling: as human beings, our ongoing potential to need substantive support and care continues throughout our entire lives.

Though we are all to varying degrees liable to fall into such a condition, we have also seen that we cannot always readily diagnose when we have done so. It is therefore important for us to acknowledge that we are being prone to be deceived here, to adopt an ongoing openness and responsiveness to what those around us have to say, and to have a willingness to change as a result of this advice – rather than taking the attitude of stubborn self-assertion that results from Subjectivism's insistence that we cannot rationally permit such social influence to occur. Hence, it follows from our ongoing frailty and potential to fall into a position of enhanced dependence that we must not consider what we presently desire or take pleasure in to be constitutive of the Human Good if we are to adequately sustain ourselves as rational agents capable of Practical Reason.

3.7 Conclusion

If the insights developed in this chapter are even roughly on the right track, then this poses grave problems for our two Subjectivist accounts. The need to place substantive restrictions on the content of our Subjective Values if they are to provide an adequate standard of what is good for us has already been recognised by Appraisive philosophers of well-being. But in addressing these issues from an Action-Guiding perspective – and especially on the developmental timescale of human lives as a whole, beginning in infancy – we saw that the obstacles to doing this with our internal resources alone proved insurmountable, and that the support provided by our social context is essential. In particular, in gaining and maintaining our powers of Practical Reason, there is always a great deal we must learn that can only be learned from others.

Our Subjectivist is therefore caught in a bind. Any Action-Guiding Theory of the Human Good must recognise Practical Reason as a worthwhile objective, on pain of becoming self-defeating.[811] Yet deep aspects of the human condition – our initially-radical vulnerability and dependency, our reliance on learning, and our inherently social natures – entail that the minimal resources we must Intrinsically Value to succeed in gaining and maintaining the status of rational agents are sufficiently extensive

[811] As argued in Section 3.1, above.

to show our Subjectivist theories to be critically incomplete. In working out how to act and how to live our lives, then, we must draw on further resources beyond what Subjectivism in any of its forms can offer. In the present context, such accounts thereby stand defeated as Action-Guiding theories of the Human Good.

Chapter 4

Conclusion

We close with a summary of the ground covered so far, some foreshadowing of future work, and a few brief remarks about the practical importance of the conclusions drawn.

In the first chapter, we developed a framework of enquiry to investigate what I have called 'Action-Guiding theories of the Human Good'; accounts of what is good for us that may be employed in guiding our deliberations, both in our here-and-now choices (Action Selection) and in considering the direction of our lives as a whole (Life Planning). This move constituted a redirection of focus: away from the third-person evaluation of human lives, and towards ideals that enable situated agents to regulate their practical deliberations from a first-person point of view.

Along the way, I outlined the methodology of 'Problem and Solution' – meaning that rival accounts of the Human Good were to be judged solely for their contribution in this context. In particular, no binding authority is claimed by those accounts which happen to reflect our present, intuitively-held views about what the Human Good consists in; this paradigm I called 'Normative Constructivism'.

I also outlined a distinction between Intrinsic and Instrumental value, reinterpreting and adapting these terms to apply within the framework. Intrinsically Good items were defined to be those that a particular theory always recommends as objects of choice, either at the level of Action Selection or at the level of Life Planning; whereas Instrumentally Good items are those that are only to be pursued when they lead to the attainment of goods of the later sort. According to the Subjectivist views we consider, only the satisfaction of our Subjective Values is Intrinsically Good.

Though in the thesis the goal was to apply the framework to our two

Subjectivist accounts – desire-based and hedonist – it was also intended to be neutral between different positive views – and indeed, will be reused in future work.

In the second chapter, we turned to the characterisation of our two Subjectivist views. Both desire-based and hedonistic views were interpreted epistemically, as informing the acting agent what their good is – again, either at the level of Action Selection or within Life Planning. In doing so, we drew heavily on the literature of academic philosophy, and also that of economics and psychology. In particular, we discussed various versions of EUT, the dominant theoretical approach in neo-classical microeconomics and contemporary behavioural economics, and discussed how it might be mobilised as means of deliberation.

One central conclusion was that we were unable to find any purely-quantitive, value-neutral conception of either desires or pleasure that met two conditions: doing full justice to the agent's actual Subjective Values, and enabling them to make their choices in an algorithmic, mechanical way. Rather, if any version of Subjectivism is to be viable, the agent will be required to make quite complex judgements – for instance, between the satisfaction of competing desires, or the pursuit of different sources of pleasure.

In the third chapter, I then offered an argument against both of our Subjectivist conceptions of the Human Good when they are deployed to guide our Practical Reasoning. The argument explained why our desires and affective preferences cannot function as an all-embracing guide to action at the level of Action Selection, nor as our organising Final End at the level of Life Planning. It proceeded by looking at the development of certain cognitive and practical powers that were necessary for the implementation of Subjectivist views in practice, and which any Subjectivist must therefore also recognise as at least Instrumental Goods. These were: a Conception of Self, having Coherent Preferences, being capable of Means-Ends reasoning, acquiring Language Skills, and the aforementioned capacity for Practical Judgement. I then drew on the literature of developmental psychology and related fields to argue that the acquisition and maintenance of these powers could only take place within a particular kind of supportive social environment, with contributions from such others as parents, teachers, instructors, role models, and medical professionals. In particular, for us to become and continue to be capable of independent rational agency at all, we must fully participate

in certain types of formative relationships with others.

Finally, I argued that participating in these relationships in the necessary manner requires us to value certain correlate Relationship Goods in an Intrinsic rather than merely Instrumental Way – and in particular, that the satisfaction of our Subjective Values must at times be set aside if our proper development as rational agents is to be achieved. But this is inconsistent with Subjectivism as understood here, and the theory is thus seen to be self-defeating.

This critical argument seems to me to be sound; however, an important problem remains: I have not yet offered a comprehensive alternative that might replace the Subjectivist views I have criticised. For instance, although I have claimed that sometimes we must listen to the advice of others, I have given no positive account of the particular circumstances when this should occur. This is a valid worry, since it leaves open the possibility that the Subjectivist view – whilst flawed – is still the best option available so far.

At the same time as elaborating and criticising the Subjectivist position, I have also often contrasted it with a neo-Aristotelian perspective, and though this has not been developed systematically, it will likely be obvious to the reader by now where my allegiance lies. In a future work, then, I intend to develop these somewhat scattered remarks into a positive Action-Guiding theory. In this, I will argue that the tradition of the virtues is able to supply the resources for deliberation that have been found lacking here – including addressing the crucial question of when we should seek substantive advice from others and whom we should listen to, as well as guiding us in developing, educating, and transforming our Subjective Values themselves, in accordance with a more adequate conception of the good life for human beings.

In order to develop this positive view, however, I will first require a much richer empirical characterisation of human behaviour and action, including how rational deliberation according to explicitly-held values fits into the picture here. In supplying this, I will also explore further the central motivational role of our Subjective Values, and how these relate to the virtues and to our character development as a whole. Moreover, as I will explore these matters in the context of developing a specific positive view, I will no longer have to remain neutral on certain points, and can also now set aside certain other aspects of the Subjectivist view that had to be assumed for the present, critical argument.

Finally, a few remarks on the importance of the conclusions drawn in this thesis for life in today's world. The two Subjectivist conceptions of what is good and best for us that I have criticised here have assumed a great influence in our culture. This is true within a number of academic disciplines, and especially economics – where EUT remains the dominant framework for understanding rational choice even amongst those theorists who strongly criticise the neoclassical paradigm. But it has also become a commonplace within the transactions of everyday life that we should follow our inner voice for guidance and treat it as the highest authority, and this notion is now a staple of contemporary music, cinema, literature, and even advertising. Consequently, exposing the limitations of Subjectivist views has implications for ethical theory and practice: for instance, in welcoming constructive advice offered by our social and familial relationships, and in our attitudes towards experts of various kinds. Moreover, to the extent that we *do* in fact succeed in gaining and maintaining our powers of practical agency, the argument given here reveals that we thereby *cannot* be exclusively following the precepts of Subjectivist theory after all. So, in uncovering the illusion of self-sufficiency in the context of practical deliberation, we have also begun to expose some myths about the self.

Chapter 5

Bibliography

[1] Abulela, Mohammed and William Bart. (2016) "A Proposed Model for Student Engagement and Standards-based Education". *The Egyptian Journal of Psychological Studies* 26, no. 90: 1-20.

[2] Adamson, Peter. (2015) *Philosophy in the Hellenistic and Roman Worlds.* Oxford University Press.

[3] Ainsworth, Mary, Mary Blehar, Everett Waters, and Sally Wall. (1978) *Patterns of Attachment: A Psychological Study of the Strange Situation*. Psychology Press.

[4] Ainsworth, Mary and Silvia Bell. (1970) "Attachment, Exploration, and Separation: Illustrated by the Behavior of One-year-olds in a Strange Situation." *Child Development* 41, no. 1: 49-67.

[5] Alchian, Armen. (1953) "The Meaning of Utility Measurement". *American Economic Review* 43, no. 1: 26-50.

[6] Alexandrova, Anna. (2017) *A Philosophy for the Science of Well-Being*. Oxford University Press.

[7] Allasi, Maurice. (1953) "Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école Américaine". *Econometrica* 21, no. 4: 503-546.

[8] Amineh, Roya Jafari and Hanieh Davatgari Asl. (2015) "Review of Constructivism and Social Constructivism". *Journal of Social Sciences, Literature and Languages* 1, no. 1: 9-16.

[9] Anderson, Elizabeth. (2018) "Dewey's Moral Philosophy". In *The Stanford Encyclopedia of Philosophy*, Winter 2019 edition, edited

by Edward N. Zalta [2005].
<https://plato.stanford.edu/archives/win2019/entries/dewey-moral/>

[10] Anderson, Samantha and Scott Maxwell. (2016) "There's More Than One Way to Conduct a Replication Study: Beyond Statistical Significance". *Psychological Methods* 21, no. 1: 1-12.

[11] Andreou, Chrisoula. (2020) "Dynamic Choice". In *The Stanford Encyclopedia of Philosophy*, Winter 2020 Edition, edited by Edward N. Zalta [2007].
<https://plato.stanford.edu/archives/win2020/entries/dynamic-choice/>

[12] Angier, Tom. (2020) "Happiness as Subjective Well-Being: An Aristotelian Critique." *Revista Portuguesa de Filosofia* 76, no. 1: 149-180.

[13] —–. (2015) "Happiness: Overcoming the Skill Model." *International Philosophical Quarterly* 55, no. 1: 5-23.

[14] Annas, Julia. (2011) *Intelligent Virtue*. Oxford University Press.

[15] —–. (2007) "Virtue Ethics and the Charge of Egoism". In *Morality and Self-Interest*, edited by Paul Bloomfield, 205-221. Oxford University Press.

[16] —–. (1993) *The Morality of Happiness*. Oxford University Press.

[17] Anscombe, G. E. M. (2000) *Intention*. Harvard University Press [1957].

[18] Aristotle. (1984) *The Complete Works of Aristotle*, edited by Jonathan Barnes. Princeton University Press.

[19] Asch, Solomon. (1956) "Studies of Independence and Conformity: I. A Minority of One Against a Unanimous Majority". *Psychological Monographs: General and Applied* 70. no. 9: 1-70.

[20] Backhouse, Roger E. and Steven G. Medema. (2009) "Retrospectives: On the Definition of Economics". *Journal of Economic Perspectives* 23, no. 1: 229-230.

[21] Baltzly, Vaughn Bryan. (2020) “The Interpersonal Variability of Gustatory Sensation and the Prospects for an Alimentary Aesthetics”. *Intervalla* 7, no. 1: 6-16.

[22] Barrett-Lennard, Godfrey. (2013) *The Relationship Paradigm.* Palgrave.

[23] Bauer, Michal, Julie Chytilová, and Barbara Pertold-Gebická. (2012) “Parental Background and Other-Regarding Preferences in Children”. *Experimental Economics* 17, no. 1: 24-46.

[24] Bell, David. (1982) “Regret in Decision Making under Uncertainty”. *Operations Research* 30, no. 5: 961-981.

[25] Bentham, Jeremy. (1970) *An Introduction to the Principles of Morals and Legislation.* Edited by J. H. Burns and H. L. A. Hart. Oxford University Press [1789].

[26] —–. (1834) *Deontology*, Vol. 1. Edited by J. Bowring. Longman, Rees, Orme, Brown, Green, and Longman.

[27] —–. (1825) *The Rationale of Reward.* John and H. L. Hunt.

[28] Benton, David. (2004) “Role of Parents in the Determination of the Food Preferences of Children and the Development of Obesity”. *International Journal of Obesity* 28: 858-869.

[29] Berghänel, Andreas, Oliver Schülkeand, and Julia Ostner. (2015) “Locomotor Play Drives Motor Skill Acquisition at the Expense of Growth: A Life History Trade-off”. *Science Advances* 1, no. 7: e1500451.

[30] Bermúdez, José Luis. (2009) *Decision Theory and Rationality.* Oxford University Press.

[31] Berridge, Kent. (2009) “Wanting and Liking: Observations from the Neuroscience and Psychology Laboratory”. *Inquiry* 52, no. 4: 378-398.

[32] Berridge, Kent and Terry Robinson. (2016) “Liking, Wanting and the Incentive-Sensitization Theory of Addiction”. *American Psychologist* 71, no. 8: 670-679.

[33] Berry, Don. (2019a) "Review of *How to be a Stoic*". *Philosophy Now* 134, October/November.

[34] —–. (2019b) "Should Mathematicians Play Dice?" *Logique Et Analyse* 246, no. 62: 135-160.

[35] —–. (2018) "Proof and the Virtues of Shared Enquiry". *Philosophia Mathematica* 26, no. 1: 112-130.

[36] Biegler, Paul. (2019) "Autonomy as a Goal of Psychotherapy". In *Oxford Handbook of Psychotherapy Ethics*, edited by Manuel Trachsel, 86-98. Oxford University Press.

[37] Bishop, Michael. (2015) *The Good Life: Unifying the Philosophy and Psychology of Well-Being.* Oxford University Press.

[38] Bjork, James, Gang Chen, and Daniel Hommer. (2012) "Psychopathic Tendencies and Mesolimbic Recruitment by Cues for Instrumental and Passively-Obtained Rewards". *Biological Psychology* 89, no. 2: 408-415.

[39] Blanchflower, David and Andrew Oswald. (2011) "International Happiness: A New View on the Measure of Performance". *Academy of Management Perspectives* 25, no. 1: 6-22.

[40] Blount, Sally. (1995) "When Social Outcomes Aren't Fair: The Effect of Causal Attributions on Preferences". *Organizational Behavior and Human Decision Processes* 63, no. 2: 131-144.

[41] Boaler, Jo. (2013) "Ability and Mathematics: the Mindset Revolution that is Reshaping Education". *FORUM* 55, no. 1: 143-152.

[42] Bogin, Barry. (1997) "Evolutionary Hypotheses for Human Childhood". *American Journal of Biological Anthropology* 104, no. 25: 63-89.

[43] Bordin, Edward. (1979) "The Generalizability of the Psychoanalytic Concept of the Working Alliance". *Psychotherapy: Theory, Research & Practice* 16, no. 3: 252-260.

[44] Boissonnet, Niels, Alexis Ghersengorin and Simon Gleyze. (2022) "Revealed Deliberate Preference Change". *SSRN* <https://ssrn.com/abstract=3605717>

[45] Bolker, Ethan D. (1967) "A Simultaneous Axiomatisation of Utility and Subjective Probability". *Philosophy of Science* 34, no. 4: 333-340.

[46] Bolton, Gary and Axel Ockenfels. (2000) "A Theory of Equity, Reciprocity, and Competition." *American Economic Review* 90, no. 1: 166-193.

[47] Bowlby, John. (1997) *Attachment and Loss,* Vol. 1. Pimlico [1969].

[48] —–. (1960) "Grief and Mourning in Infancy and Early Childhood". *The Psychoanalytic Study of the Child* 15, no. 1: 9-52.

[49] —–. (1959) "Separation Anxiety". *International Journal of Psychoanalysis* 41: 69-113.

[50] —–. (1958) "The Nature of the Child's Tie to His Mother". *The International Journal of Psychoanalysis* 39, no. 5: 350-373.

[51] Bowles, Samuel. (2004) *Microeconomics: Behaviour, Institutions, and Evolution.* Princeton University Press.

[52] —–. (1998) "Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions". *Journal of Economic Literature* 36, no. 1: 75-111.

[53] Boyd, Pete and Agnieszka Szplit, eds. (2016) *Student Teachers Learning Through Inquiry: International Perspectives.* Wydawnictwo Attyka.

[54] Boyd, Richard. (2020) "The Early Modern Origins of Behavioral Economics". *Social Philosophy and Policy* 37, no. 1: 30-54.

[55] Braddock, Matthew. (2010) "Constructivist Experimental Philosophy on Well-being and Virtue". *The Southern Journal of Philosophy* 48, no. 3: 295-323.

[56] Bradley, Richard. (2017) *Decision Theory with a Human Face.* Cambridge University Press.

[57] Brake, Elizabeth. (2017) "Fair Care: Elder Care and Distributive Justice". *Politics, Philosophy & Economics* 16, no. 2: 132-151.

[58] Bramble, Ben. (2016a) "The Experience Machine". *Philosophy Compass* 11, no. 3: 136-145.

[59] ——. (2016b) "A New Defense of Hedonism about Well-Being". *Ergo* 3, no. 4: 85-112.

[60] Bretherton, Inge. (1980) "Young Children in Stressful Situations: The Supporting Role of Attachment Figures and Unfamiliar Caregivers". In *Uprooting and Development*, edited by G. V. Coelho & P. J. Ahmen, 179-210. Plenum Press.

[61] Brickman, Philip, Dan Coates and Ronnie Janoff-Bullman. (1978) "Lottery Winners and Accident Victims: Is Happiness Relative?" *Journal of Personality and Social Psychology* 36, no. 8: 917-27.

[62] Bridgman, Percy. (1927) *The Logic of Modern Physics.* MacMillan.

[63] Briggs, R. A. (2019) "Normative Theories of Rational Choice: Expected Utility". In *The Stanford Encyclopedia of Philosophy*, Fall 2019 edition, edited by Edward N. Zalta [2014]. <https://plato.stanford.edu/archives/fall2019/entries/rationality-normative-utility/>

[64] Brighouse, Harry and Adam Swift. (2014) *Family Values: The Ethics of Parent-Child Relationships.* Princeton University Press.

[65] Brock, Dan. (1993) "Quality of Life Measures in Health Care and Medical Ethics". In *The Quality of Life*, edited by Martha Nussbaum and Amartya Sen, 95-132. Clarendon Press.

[66] Broom, Donald. (2013) "Needs of Animals". In *Encyclopedia of Animal Rights and Welfare*, edited by Marc Bekoff, 255-256. Routledge.

[67] Broome, John. (1999) *Ethics out of Economics.* Cambridge University Press.

[68] ——. (1979) "Trying to Value a Life: A Reply". *Journal of Public Economics* 12, no. 2: 259-262.

[69] ——. (1978) "Trying to Value a Life". *Journal of Public Economics* 9, no. 1: 91-100.

[70] Brownlee, Kimberley. (2013) "A Human Right Against Social Deprivation". *The Philosophical Quarterly* 63, no. 251: 199-222.

[71] Browowski, Jacob. (1974) *The Ascent of Man.* Little, Brown and Company.

[72] Brue, Stanley and R. G. Grant. (2012) *The Evolution of Economic Thought*, 8th ed. Cengage Learning.

[73] Bruni, Luigino and Robert Sugden. (2007) "The Road Not Taken: How Psychology Was Removed From Economics, and How it Might be Brought Back". *Economic Journal* 117, no. 516: 146-173.

[74] Buchanan, James. (1989) "The Public-Choice Perspective." In *Essays on the Political Economy.* University of Hawaii Press.

[75] Burckhardt, Jacob. (2002) *History of Greek Culture* [abridged]. Dover Publications [first published in German in two volumes; 1898 and 1902].

[76] Burton, Arthur. (1972) *Twelve Therapists.* Jossey-Bass.

[77] Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler. (1997) "Labor Supply of New York City Cabdrivers: One Day at a Time." *The Quarterly Journal of Economics* 112, no. 2: 407-441.

[78] Camus, Albert. (1972) *A Happy Death.* Translated by Richard Howard. Alfred A. Knopf.

[79] Cappelen, Herman, David Plunkett, and Alexis Burgess, eds. (2020) *Conceptual Engineering and Conceptual Ethics*. Oxford University Press.

[80] Carey, Timothy and Richard Mullan. (2004) "What is Socratic questioning?" *Psychotherapy: Theory, Research, Practice, Training* 41, no. 3: 217-226.

[81] Carr, David. 2005, "Personal and Interpersonal Relationships in Education and Teaching: A Virtue Ethical Perspective". *British Journal of Educational Studies* 53, no. 3: 255-271.

[82] Chugani, Harry, Michael Behen, Otto Muzik, Csaba Juhász, Ferenc Nagy, and Diane Chugani. (2001) "Local Brain Functional Activity Following Early Deprivation: a Study of Postinstitutionalized Romanian Orphans". *Neuroimage* 14, no. 6: 1290-1301.

[83] Churchland, Patricia. (1989) *Neurophilosophy: Toward a Unified Science of the Mind-Brain.* MIT Press.

[84] Clark, Andrew. (2018) "Four Decades of the Economics of Happiness: Where Next?" *Review of Income and Wealth* 64, no. 2: 245-269.

[85] Clarke, Christopher. (2016) "Preferences and Positivist Methodology in Economics". *Philosophy of Science* 83, no. 2: 192-212.

[86] Clarke-Doane, Justin and Kathryn Tabb. (2022) "Addiction and Agency". In *Agency in Mental Disorder: Philosophical Dimensions*, edited by Matt King and Joshua May, 154-168. Oxford University Press.

[87] Coase, Ronald. (1937) "The Nature of the Firm". *Economica* 4, no. 16: 386-405.

[88] Cohen, G. A. (1993) "Equality of What? On Welfare, Goods, and Capabilities". In *The Quality of Life*, edited by Martha Nussbaum and Amartya Sen, 9-29. Oxford: Clarendon Press.

[89] Cohn, Michael and Barbara Fredrickson. (2009) "Positive Emotions". In *Oxford Handbook of Positive Psychology*, 2nd ed., edited by Shane Lopez and C. R. Snyder, 13-24. Oxford University Press.

[90] Colander, David. (2007) "Edgeworth's Hedonimeter and the Quest to Measure Utility". *Journal of Economic Perspectives* 21, no. 2: 215-226.

[91] Collingwood, R. G. (2013) "Autobiography". In *R. G. Collingwood: An Autobiography and Other Writings*, 1-174. Oxford University Press [1939].

[92] —–. (1992) "Goodness, Rightness, Utility". In *The New Leviathan*, edited by David Boucher, 391-479. Oxford University Press [1940].

[93] —–. (1958) *The Principles of Art.* Oxford University Press [1938].

[94] Combs, Arthur. (1986) "What Makes a Good Helper? A Person-Centered Approach". *Person-Centered Review* 1, no. 1: 51-61.

[95] Cooper, John. (2012) *Pursuits of Wisdom: Six Ways of Life in Ancient Philosophy from Socrates to Plotinus.* Princeton University Press.

[96] —–. (1998) "Pleasure and Desire in Epicurus". In *Reason and Emotion*, 485-514. Princeton University Press.

[97] Cordelli, Chiara. (2015) "Justice as Fairness and Relational Resources". *Journal of Political Philosophy* 23, no. 1: 86-110.

[98] Corr, Philip and Anke Plagnol. (2019) *Behaviour Economics: The Basics*. Routledge.

[99] Craig, Edward. (2007) "Genealogies and the State of Nature". In *Bernard Williams: Contemporary Philosophy in Focus*, edited by Alan Thomas, 181-200. Cambridge University Press.

[100] —–. (1990) *Knowledge and the State of Nature*. Clarendon Press.

[101] Crisp, Roger. (2021) "Well-Being". In *The Stanford Encyclopedia of Philosophy*, Winter 2021 edition, edited by Edward N. Zalta [2001]. <https://plato.stanford.edu/archives/win2021/entries/well-being/>

[102] —–. (2006) "Hedonism Reconsidered". *Philosophy and Phenomenological Research* 73, no. 3: 619-645.

[103] Crisp, Roger and Morten Kringelbach. (2018) "Higher and Lower Pleasures Revisited: Evidence from Neuroscience". *Neuroethics* 11, no. 2: 211-215.

[104] Cunningham, Joe. (2021) "The Matter of Motivating Reasons". *Philosophical Studies* 179, no. 5: 1563-1589.

[105] Cunningham, John, Linda Sobell, Mark Sobell, and Janet Gaskin. (1994) "Alcohol and Drug Abusers' Reasons For Seeking Treatment". *Addictive Behaviors* 19, no. 6: 691-696.

[106] Damasio, Antonio. (2008) *Descartes' Error: Emotion, Reason and the Human Brain.* Vintage.

[107] Davidson, Donald. (2004) *Problems of Rationality.* Cambridge Universty Press.

[108] ——. (1963) “Actions, Reasons, and Causes”. *The Journal of Philosophy* 60, no. 23: 685-700.

[109] Davis, Kingsley. (1940) “Extreme Social Isolation of a Child”. *American Journal of Sociology* 45, no. 4: 554-565.

[110] Davis, Ryan and Rachel Finlayson. (2021) “When Should we be Open to Persuasion?” *Ethical Theory and Moral Practice* 24, no. 1: 123-136.

[111] d’Avray, David L. (2010) *Rationalities in History.* Cambridge University Press.

[112] De Brigard, Felipe. (2010) “If You Like It, Does It Matter If It’s Real?” *Philosophical Psychology* 23, no. 1: 43-57.

[113] de Lazari-Radek, Katarzyna, and Peter Singer. (2014) *The Point of View of the Universe: Sidgwick and Contemporary Ethics*. Oxford University Press.

[114] Debreu, Gérard. (1959) *Theory of Value.* Wiley.

[115] ——. (1954) “Representation of a Preference Ordering by a Numerical Function”. In *Decision Processes*, edited by Thrall, Coombs, and Raiffa, 159-167. Wiley.

[116] Dennett, Daniel. (2017) *From Bacteria to Bach and Back.* W. W. Norton & Company.

[117] ——. (1991) *Consciousness Explained.* Little, Brown & Co.

[118] Deutsch, Morton and Harold Gerard. (1955) “A Study of Normative and Informational Social Influences Upon Individual Judgement”. *The Journal of Abnormal and Social Psychology* 51, 3: 629-636.

[119] DeVries, Rheta, Betty Zan, Carolyn Hildebrandt, and Rebecca Edmiaston. (2002) *Developing Constructivist Early Childhood Curriculum: Practical Principles and Activities.* Teachers College Press.

[120] Dewey, John. (2002) *Human Nature and Conduct.* Dover Publications [1922].

[121] —–. (1998) *The Essential Dewey*, 2 vols. Edited by Larry Hickman and Thomas Alexander. Indiana University Press.

[122] Diener, Ed, Richard, Lucas, and Christie Scollon. (2006) "Beyond the Hedonic Treadmill: Revising the Adaptation Theory of Well-Being." *American Psychologist* 61, no. 4. 305-314.

[123] Dietrich, Franz and Christian List. (2016) "Mentalism versus Behaviourism in Economics: A Philosophy-of-Science Perspective". *Economics & Philosophy* 32, no. 2: 249-281.

[124] Ding, Yan-hua, Xiu Xu, Zheng-yan Wang, Hui-rong Li, and Wei-ping Wang. (2014) "The relation of infant attachment to attachment and cognitive and behavioural outcomes in early childhood". *Early Human Development* 90, no. 9: 459-64.

[125] Dollard, John and Neal Miller. (1950) *Personality and Psychotherapy: An Analysis in Terms of Learning, Thinking, and Culture*. McGraw-Hill.

[126] Driver, Julia. (2014) "Mill, Moral Sentimentalism, and the Cultivation of Virtue". In *Cultivating Virtue: Perspectives from Philosophy, Theology, and Psychology*, edited by Nancy E. Snow, 49-64. Oxford University Press.

[127] Dunne, Joseph. (1993) *Back to the Rough Ground: Practical Judgement and the Lure of Technique.* University of Notre Dame Press.

[128] Duska, Ronald and Mariellen Whelan. (1977) *Moral Development: A Guide to Piaget and Kohlberg.* Gill and Macmillan.

[129] Dweck, Carol. (2006) *Mindset: The New Psychology of Success.* Ballantine Books

[130] Edgeworth, Francis. (1879) "The Hedonical Calculus". *Mind* 4, no. 15: 394-408.

[131] Ellis, Albert. (2001) *Feeling Better, Getting Better, Staying Better: Profound Self-help Therapy for Your Emotions.* Impact Publishers.

[132] —–. (1994) *Reason and Emotion in Psychotherapy.* Carol Publishing Group [1962].

[133] Ellsberg, Daniel. (1961) "Risk, Ambiguity, and the Savage Axioms". *Quarterly Journal of Economics* 75, no. 4: 643-669.

[134] Enfield, Nick and Kockelman, Paul. (2017) *Distributed Agency.* Oxford University Press.

[135] Enoch, David. (2006) "Agency, Shmagency: Why Normativity Won't Come from What is Constitutive of Agency". *Philosophical Review* 115, no. 2: 169-198.

[136] Epictetus. (2008) *Discourses and Other Writings.* Translated by Robert Dobbin. Penguin Classics.

[137] Epicurus. (2011) *Letter to Menoikos.* Translated by Peter Saint-Andre. Monadnock Valley Press.
<https://monadnock.net/epicurus/letter.html>.

[138] —. (2008) *Principal Doctrines.* Translated by Peter Saint-Andre. Monadnock Valley Press.
<https://monadnock.net/epicurus/principal-doctrines.html>.

[139] Eriksson, Nicholas, Shirley Wu, Chuong B Do, Amy K Kiefer, Joyce Y Tung, Joanna L Mountain, David A Hinds & Uta Francke (2012) "A Genetic Variant Near Olfactory Receptor Genes Influences Cilantro Preference". *Flavour* 1, no. 22.

[140] Etkind, Simon, Natasha Lovell, Caroline Nicholson, Irene Higginson, and Fliss Murtagh. "Finding a 'New Normal' Following Acute Illness: A Qualitative Study of Influences on Frail Older People's Care Preferences". *Palliative Medicine* 33, no. 3: 301-311.

[141] Evans, Karen. (2001) "Resilient older widows: A Phenomenological Study of Identity Reformulation, Personal Transformation, and Re-creation of Life Meaning Following Loss of a Longtime Spouse". *Dissertation Abstracts International: Section B: The Sciences and Engineering* 62, no. 5-B: 2515.

[142] Fairbairn, W. Ronald D. (1958) "On the Nature and Aims of Psycho-Analytic Treatment". *International Journal of Psychoanalysis* 39, no. 5: 374-385.

[143] Falk, Armin and Urs Fischbacher. (2006) "A Theory of Reciprocity". *Games and Economic Behavior* 54, no. 2: 293-315.

[144] Fehr, Ernst and Klaus Schmidt. (1999) "A Theory of Fairness, Competition, and Cooperation". *The Quarterly Journal of Economics* 114, no. 3: 817-868.

[145] Feldman, Fred. (2010) *What Is This Thing Called Happiness?* Oxford University Press.

[146] ——. (2004) *Pleasure and the Good Life.* Oxford University Press.

[147] ——. (2002) "The Good Life: A Defense of Attitudinal Hedonism". *Philosophy and Phenomenological Research* 65, no. 3: 604-628.

[148] ——. (2000) "Basic Intrinsic Value". *Philosophical Studies* 99, no. 3: 319-346.

[149] ——. (1997) *Utilitarianism, Hedonism, and Desert: Essays in Moral Philosophy.* Cambridge University Press.

[150] Feldman, Ruth. (2012) "Oxytocin and Social Affiliation in Humans". *Hormones and Behavior*, 61, no. 3: 380-391.

[151] ——. (2007) "Parent-Infant Synchrony: Biological Foundations and Developmental Outcomes". *Current Directions in Psychological Science* 16, no. 6: 340-345.

[152] Feldman, Ruth, Esther Bamberger, and Yaniv Kanat-Maymon. (2013) "Parent-Specific Reciprocity From Infancy to Adolescence Shapes Children's Social Competence and Dialogical Skills". *Attachment and Human Development* 15, no. 4: 407-423.

[153] Field, John. (2000) *Lifelong Learning and the New Educational Order.* Trentham Books.

[154] Finnis, John. (2011) *Natural Law and Natural Right*, 2nd ed. Oxford University Press [1980].

[155] Fletcher, Guy. (2016) *The Philosophy of Well-being: An Introduction.* Routledge.

[156] Forsythe, Robert, Joel Horowitz, Nathan Savin, and Martin Sefton. (1994) "Fairness in Simple Bargaining Experiments". *Games and Economic Behavior* 6, no. 3: 347-369.

[157] Frankel, Marvin and Lisbeth Sommerbeck. (2005) "Two Rogers and Congruence: The Emergence of Therapist-Centered Therapy and the Demise of Client-Centered Therapy". In *Embracing Non-directivity: Reassessing Person-Centered Theory and Practice in the 21st Century*, edited by B.E. Levitt, 40-61. PCCS Books.

[158] Frankfurt, Harry. (2006) *Taking Ourselves Seriously & Getting it Right.* Stanford University Press.

[159] ——. (1971) "Freedom of the Will and the Concept of a Person". *The Journal of Philosophy* 68, no. 1: 5-20.

[160] Fraser, David. (2013) *Understanding Animal Welfare.* Wiley-Blackwell.

[161] Frederick, Shane and George Loewenstein. (1999) "Hedonic adaptation". In *Well-being: Foundations of Hedonic Psychology*, edited by Daniel Kahneman, Ed Diener, and Norbert Schwarz, 302-329. Russell Sage Foundation.

[162] Fredrickson, Barbara and Daniel Kahneman. (1993) "Duration neglect in retrospective evaluations of affective episodes". *Journal of Personality and Social Psychology* 65, no. 1: 45-55.

[163] Fried, Terri, John O'Leary, Peter Van Ness, and Liana Fraenkel. (2007) "Inconsistency Over Time in the Preferences of Older Persons with Advanced Illness For Life-Sustaining Treatment". *Journal of the American Geriatrics Society* 55, no. 7: 1007-1014.

[164] Friedman, Milton. (1953) "The Methodology of Positive Economics". In *Essays in Positive Economics*, 3-16. University of Chicago Press.

[165] Frijda, Nico. (1986) *The Emotions.* Cambridge University Press.

[166] Galbraith, John. (1958) *The Affluent Society.* Houghton Mifflin.

[167] Gauthier, David. (1987) *Morals by Agreement.* Oxford University Press.

[168] Geach, Peter. (1977) *The Virtues: The Stanton Lectures 1973-74.* Cambridge University Press.

[169] Gendlin, Eugene. (1966) "Research in Psychotherapy with Schizophrenic Patients and the Nature of That "Illness."" *American Journal of Psychotherapy* 20, no. 1: 4-16.

[170] Gheaus, Anca. (2018) "Personal Relationship Goods". In *The Stanford Encyclopedia of Philosophy*, Fall 2018 Edition, edited by Edward N. Zalta.

<https://plato.stanford.edu/archives/fall2018/entries/personal-relationship-goods/>.

[171] Ghoshal, Sumantra. (2005) "Bad Management Theories Are Destroying Good Management Practices". *Academy of Management Learning & Education* 4, no. 1: 75-91.

[172] Gilbert, Daniel, Elizabeth Pinel, Timothy Wilson, Stephen Blumberg, and Thalia Wheatley. (1998) "Immune Neglect: A Source of Durability Bias in Affective Forecasting". *Journal of Personality and Social Psychology* 75, no. 3: 617-638.

[173] Gilbert, Daniel and Timothy Wilson. (2000) "Miswanting: Some Problems in the Forecasting of Future Emotional States". In *Thinking and feeling: The Role of Affect in Social Cognition*, edited by J. Forgas, 178-197. Cambridge University Press.

[174] Gneezy, Uri and Jan Potters. (1997) "An Experiment on Risk Taking and Evaluation Periods". *The Quarterly Journal of Economics* 112, no. 2: 631-645.

[175] Goffman, Erving. (1974) *Frame Analysis.* Harvard University Press.

[176] Gould, Stephen J. (2008) "A Biological Homage to Mickey Mouse". *Ecotone* 4, no. 1: 333-340.

[177] Green, Leonard, Nathanael Fristoe, and Joel Myerson. (1994) "Temporal Discounting and Preference Reversals in Choice Between Delayed Outcomes". *Psychonomic Bulletin & Review* 1, no. 3: 383-389.

[178] Greene, Robert. (2014) *Mastery*, concise edition. Profile Books.

[179] Griffin, James. (1993) "Dan Brock: Quality of Life Measures in Health Care and Medical Ethics" [Commentary]. In *The Quality of Life*, edited by Martha Nussbaum and Amartya Sen, 133-139. Clarendon Press.

[180] —–. (1986) *Well-Being.* Clarendon Press.

[181] Groh, Ashley, R. M. Pasco Fearon, Marinus van IJzendoorn, Marian J. Bakermans-Kranenburg, and Glenn I. Roisman. (2017) "Attachment in the Early Life Course: Meta-Analytic Evidence

for its Role in Socioemotional Development". *Child Development Perspectives* 11, no. 1: 7-76.

[182] Griffin, Gary and Harry Harlow. (1966) "Effects of Three Months of Total Social Deprivation on Social Adjustment and Learning in the Rhesus Monkey". *Child Development* 37, no. 3: 533-547.

[183] Grinde, Bjørn. (2002) *Darwinian Happiness*. The Darwin Press.

[184] Gul, Frauk and and Wolfgang Pesendorfer. (2008) "The Case for Mindless Economics". In *The Foundations of Positive and Normative Economics: A Handbook*, edited by Andrew Caplin and Andrew Schotter, 2-40. Cambridge University Press.

[185] Güth, Werner, Rolf Schmittberger, and Bernd Schwarze. (1982) "An Experimental Analysis of Ultimatum Bargaining". *Journal of Economic Behavior & Organization* 3. no. 4: 367-388.

[186] Halevy, Yoram. (2015) "Time Consistency: Stationarity and Time Invariance". *Econometrica* 83, no. 1: 335-352.

[187] Hamre, Bridget and Robert Pianta. (2006) "Student-Teacher Relationships". In *Children's needs III: Development, prevention, and intervention*, edited by G. G. Bear and K. M. Minke, 59-71. National Association of School Psychologists.

[188] Hands, D. Wade. (2012) "Normative Rational Choice Theory: Past, Present, and Future". *Voprosy ėkonomiki* 10: 52-75.

[189] Hansson, Sven Ove. (2018) "Risk". In *The Stanford Encyclopedia of Philosophy*, Fall 2018 edition, edited by Edward N. Zalta [2007]. <https://plato.stanford.edu/archives/fall2018/entries/risk/>

[190] Hansson, Sven Ove and Till Grüne-Yanoff. (2022) "Preferences". In *The Stanford Encyclopedia of Philosophy*, Spring 2022 edition, edited by Edward N. Zalta [2006]. <https://plato.stanford.edu/archives/spr2022/entries/preferences/>.

[191] Haidt, Jonathan. (2006) *The Happiness Hypothesis*. Arrow Books.

[192] Hardie, W. F. (1965) "The Final Good in Aristotle's Ethics". *Philosophy* 40, no. 154: 277-295.

[193] Harlow, Harry and Robert Zimmermann. (1958) "The Development of Affective Responsiveness in Infant Monkeys". *Proceedings of the American Philosophical Society* 102, no. 5: 501-509.

[194] Harris, Judith. (1998) *The Nurture Assumption: Why Children Turn Out the Way They Do*. The Free Press.

[195] Harsanyi, John. (1977) "Morality and the Theory of Rational Behavior". *Social Research* 44, no. 4: 623-656.

[196] Hausman, Daniel. (2015) *Valuing Health*. Oxford University Press.

[197] —. (2011) *Preference, Value, Choice, and Welfare*. Cambridge University Press.

[198] Hausman, Daniel and Michael McPherson. (2009) "Preference Satisfaction and Welfare Economics". *Economics & Philosophy* 25, no. 1: 1-25.

[199] Hawkins, Jennifer. (2021) "Further Reflections: Surrogate Decision Making When Significant Mental Capacities are Retained". *Cambridge Quarterly of Healthcare Ethics* 30, no. 1: 192-198.

[200] Haybron, Daniel. (2016) "The Philosophical Basis of Eudaimonic Psychology". In *Handbook of Eudaimonic Well-Being*, edited by Joar Vittersø, 27-53. Springer.

[201] Hazan, Cindy and Phillip Shaver. (1987) "Romantic Love Conceptualized as an Attachment Process". *Journal of Personality and Social Psychology* 52, no. 3: 511-524.

[202] Hazan, Marcella. (2011) *The Essentials of Classic Italian Cooking*. Boxtree.

[203] Hearne, Viki. (1987) *Adam's Task: Calling Animals by Name*. Vintage Books.

[204] Heathwood, Chris. (2005) "The Problem of Defective Desires". *Australasian Journal of Philosophy* 83, no. 4: 487-504.

[205] Heilbron, Johan. (2009) "Social Thought and Natural Science". In *The Cambridge History of Science*, Vol. 7, edited by Theodore Porter and Dorothy Ross, 40-56. Cambridge University Press.

[206] Held, Virginia. (1993) *Feminist Morality: Transforming Culture, Society and Politics*. University of Chicago Press.

[207] Helm, Bennett. (2021) "Friendship". In *The Stanford Encyclopedia of Philosophy*, Fall 2021 Edition, edited by Edward N. Zalta [2005]. <https://plato.stanford.edu/archives/fall2021/entries/friendship/>

[208] Henderson, Barney. (2016) "Meet Yuichiro Miura, the Man Planning to Conquer Everest at 90". *The Telegraph*, 1st January. <https://www.telegraph.co.uk/health-fitness/body/meet-yuichiro-miura-the-man-planning-to-conquer-everest-at-90/>

[209] Henry, Stuart and David Robinson. (1978) "Understanding Alcoholics Anonymous: Results from a Survey in England and Wales". *The Lancet* 311, no. 8060: 372-375.

[210] Hicks, John and Roy Allen. (1934) "A Reconsideration of the Theory of Value". *Economica* 1, no. 1: 52-76.

[211] Hills, Alison. (2007) "Practical Reason, Value and Action". *Journal of Moral Philosophy* 4, no. 3: 375-392.

[212] Hobbes, Thomas. (1949) *De cive, or, The Citizen.* Appleton-Century-Crofts, Inc. [1642].

[213] Hsee, Christopher. (2000) "Attribute Evaluability and its Implications for Joint-Separate Evaluation Reversals and Beyond". In *Choices, Values, and Frames*, edited by Daniel Kahneman and Amos Tversky, 543-563. Cambridge University Press.

[214] Hull, David. (1986) "On Human Nature". *Proceedings of the Biennial Meeting of the Philosophy of Science Association* 1986, no. 2: 3-13.

[215] Hume, David. (2000) *A Treatise of Human Nature.* Oxford University Press [1739-1740].

[216] —–. (1994) *David Hume: Political Writings.* Edited by Stuart Warner and Donald Livingston. Hackett.

[217] Hurka, Thomas. (1987) "'Good' and 'Good For"'. *Mind* 96, no. 381: 71-73.

[218] Hursthouse, Rosalind. (1999) *On Virtue Ethics*. Oxford University Press.

[219] Ingrao, Bruna and Giorgio Israel. (1990) *The Invisible Hand.* Translated by Ian McGilvray. MIT Press.

[220] IJzendoorn, Marinus van, Carlo Schuengel, and Marian Bakermans-Kranenburg. (1999) "Disorganized Attachment in Early Childhood: Meta-Analysis of Precursors, Concomitants, and Sequelae". *Development and Psychopathology* 11, no. 2: 225-249.

[221] Islam, Gazi. (2022) "Business Ethics and Quantification: Towards an Ethics of Numbers". *Journal of Business Ethics* 176, no. 2: 195-211.

[222] Jaffe, Joseph, Daniel Stern, and J. Craig Peery. (1973) "'Conversational' Coupling of Gaze Behavior in Prelinguistic Human Development". J*ournal of Psycholinguistic Research* 2, no. 4: 321-329.

[223] Jarden, Aaron. (2011) "An Interview with Daniel Kahneman". *The International Journal of Wellbeing* 1, no. 1: 186-188.

[224] Jaworska, Agnieszka. (1999) "Respecting the Margins of Agency: Alzheimer's Patients and the Capacity to Value". *Philosophy & Public Affairs* 28, no. 2: 105-138.

[225] Jeffrey, Richard. (1965) *The Logic of Decision.* McGraw-Hill.

[226] Jevons, William. (1965) *The Theory of Political Economy.* A. M. Kelly [1871].

[227] Johnson, Scott, Alan Slater, and Ian Hocking. (2017) "The Scope and Methods of Developmental Psychology". In *An Introduction to Developmental Psychology*, 3rd ed, edited by Alan Slater, and J. Gavin Bremner, 41-78. Wiley.

[228] Jones-Lee, M. W. (1979) "Trying to Value a Life: Why Broome Does Not Sweep Clean". *Journal of Public Economics* 12, no. 2: 249-256.

[229] Kahneman, Daniel. (2011) *Thinking, Fast and Slow.* Penguin.

[230] ——. (2003) "Maps of Bounded Rationality: Psychology for Behavioral Economics". *The American Economic Review* 93, no. 5: 1449-1475.

[231] ——. (1999) "Objective Happiness". In *Well-being: Foundations of Hedonic Psychology*, edited by Daniel Kahneman, Ed Diener, and Norbert Schwarz, 3-25. Russell Sage Foundation.

[232] Kahneman, Daniel and Amos Tversky. (1979) "Prospect Theory: An Analysis of Decision under Risk". *Econometrica* 47, no. 2: 263-291.

[233] Kahneman, Daniel, Barbara Fredrickson, Charles Schreiber, and Donald Redelmeier. (1993) "When More Pain Is Preferred to Less: Adding a Better End". *Psychological Science* 4, no. 6: 401-405.

[234] Kahneman, Daniel, Peter Wakker and Rakesh Sarin. (1997) "Back to Bentham? Explorations of Experienced Utility". *The Quarterly Journal of Economics* 112, no. 2: 375-405.

[235] Kahneman, Daniel and Richard Thaler. (2006) "Utility Maximization and Experienced Utility". *Journal of Economic Perspectives* 20, no. 1: 221-234.

[236] Kagan, Jerome. (1998) "Is There a Self in Infancy?" In *Self-awareness: Its Nature and Development*, edited by M. D. Ferrari and R. J. Sternberg, 137-147. Guilford Press.

[237] Kant, Immanuel. (1996) *Practical Philosophy*. Translated and edited by Mary J. Gregor. Cambridge University Press.

[238] Katz, Leonard. (2016) "Pleasure". In *The Stanford Encyclopedia of Philosophy*, Winter 2016 Edition, edited by Edward N. Zalta [2005]. <https://plato.stanford.edu/archives/win2016/entries/pleasure/>

[239] Kauppinen, Antti. (2008) "Working Hard and Kicking Back: The Case for Diachronic Perfectionism". *Journal of Ethics and Social Philosophy* 3, no. 1: 1-10

[240] Kellman, Philip and Martha Arterberry. (2000) *The Cradle of Knowledge: Development of Perception in Infancy*. MIT Press.

[241] Keyes, Corey and Julia Annas. (2009) "Feeling Good and Functioning Well: Distinctive Concepts in Ancient Philosophy and Contemporary Science". *The Journal of Positive Psychology* 4, no. 3: 197-201.

[242] Kimmel, Stephen, Andrea Troxel, George Loewenstein, Colleen Brensinger, Jane Jaskowiak, Jalpa Doshi, Mitchell Laskin, and Kevin Volpp. (2012) "Randomized Trial of Lottery-Based Incentives to Improve Warfarin Adherence". *American Heart Journal* 164, no. 2: 268-74.

[243] Kirschner, Paul A., John Sweller, and Richard E. Clark. (2006) "Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-based, Experiential, and Inquiry-Based Teaching". *Educational Psychologist* 41, no. 2: 75-86.

[244] Kirton, Andrew. (2020) "Matters of Trust as Matters of Attachment Security". *International Journal of Philosophical Studies Volume* 28, no. 5: 583-602.

[245] Kittay, Eva. (1999) *Love's Labor: Essays on Women, Equality and Dependency*, 2nd ed. Routledge.

[246] Klein, Mélanie (1946) "Notes on some Schizoid Mechanisms". *International Journal of Psycho-Analysis* 27, no. 3-4: 99-110.

[247] Kohut, Heinz. (1972) "Thoughts on Narcissism and Narcissistic Rage". *The Psychoanalytic Study of the Child* 27, no. 1: 360-400.

[248] Konow, James, and Joseph Earley. (2008) "The Hedonistic Paradox: Is *homo economicus* Happier?" *Journal of Public Economics* 92, no. 1-2: 1-33.

[249] Korsgaard, Christine. (1996) *The Sources of Normativity.* Cambridge University Press.

[250] —–. (1983) "Two Distinctions in Goodness". *The Philosophical Review* 92, no. 2: 169-195.

[251] Kovacs, Maria and Aaron Beck. (1978) "Maladaptive Cognitive Structures in Depression". *The American Journal of Psychiatry* 135, no. 5: 525-533.

[252] Kövecses, Zoltán. (1991) "Happiness: A Definitional Effort". *Metaphor and Symbolic Activity* 6, no. 1: 29-46.

[253] Knight, Frank. (1921) *Risk, Uncertainty, and Profit.* Houghton Mifflin Company.

[254] Knight, Kelvin. (1996) "Editor's Introduction". In *The MacIntyre Reader*, edited by Kelvin, Knight, 1-31. Polity Press.

[255] Kraut, Richard. (2011) *Against Absolute Goodness.* Oxford University Press.

[256] —–. (2007) *What is Good and Why.* Harvard University Press.

[257] Krumpal, Ivar. (2013) "Determinants of Social Desirability Bias in Sensitive Surveys: a Literature Review". *Quality & Quantity* 47, no. 4: 2025-2047.

[258] Laërtius, Diogenes. (1991) *Lives of Eminent Philosophers*, Volumes I and II. Translated by R.D. Hicks. Harvard University Press.

[259] Lakoff, George and Mark Johnson. (2003) *Metaphors We Live By.* University of Chicago Press [1980].

[260] Lamb, Michael. (1997) "The Development of Father-Infant Relationships". In *The Role of the Father in Child Development*, edited by Michael Lamb, 104-120. John Wiley & Sons, Inc.

[261] Layard, Richard. (2011) *Happiness: Lessons from a New Science*, revised edition. Penguin.

[262] Leffler, Olof. (2019) "New Shmagency Worries". *Journal of Ethics and Social Philosophy* 15, no. 2: 121-145.

[263] Lemaire, Stéphane. (2020) "Subjectivism Without Idealization and Adaptive Preferences". *Utilitas* 33, no. 1: 85-100.

[264] Lepper, Mark and David Greene. (1978) "Overjustification Research and Beyond: Toward a Means-Ends Analysis of Intrinsic and Extrinsic Motivation". In *The Hidden Costs of Reward*, edited by Mark Lepper and David Greene, 109-148. Psychology Press.

[265] Lerner, Jennifer, Ye Li, Piercarlo Valdesolo, and Karim Kassam. (2015) "Emotion and Decision Making". *Annual Review of Psychology* 66, no. 33: 799-823.

[266] Levi, Isaac. (1986) *Hard Choices: Decision Making Under Unresolved Conflict*. Cambridge University Press.

[267] Levitt, Steven D. and Stephen J. Dubner. (2006) *Freakonomics*, revised and expanded edition. William Morrow.

[268] Lewis, David. (1996) “Desire as Belief II”. *Mind* 105, no. 418: 303-313.

[269] —–. (1988) “Desire as Belief”. *Mind* 97, no. 418: 323-332.

[270] Lockhart, Paul. (2009) *A Mathematicians’ Lament.* Bellevue Literary Press.

[271] Loewenstein, George, Ted O’Donoghue, and Matthew Rabin. (2003) “Projection Bias in Predicting Future Utility”. *The Quarterly Journal of Economics* 118, no. 4: 1209-1248.

[272] Long, A. A. and D. N. Sedley. (1987) *The Hellenistic Philosophers,* Vol. 1. Cambridge University Press.

[273] Lorenz, Konrad. (1935) “Der Kumpan in der Umwelt des Vogels: Der Artgenosse als auslösendes Moment sozialer Verhaltensweisen”. *Journal für Ornithologie* 83, 137-215, 289-413.

[274] Louden, Robert. (1992) *Morality and Moral Theory.* Oxford University Press.

[275] Luce, Robert. D. (2002) “A Psychophysical Theory of Intensity Proportions, Joint Presentations, and Matches”. *Psychological Review* 109, no. 3: 520-532.

[276] Machek, David. (2021) “Mengzi on Nourishing the Heart by Having Few Desires”. *Philosophy East and West* 71, no. 2: 393-413.

[277] MacIntyre, Alasdair. (2016) *Ethics in the Conflicts of Modernity.* Cambridge University Press.

[278] —–. (2012) “Catholic Instead of What?” [Lecture]. Notre Dame, Indianna, November 9. <https://chamberscreek.net/library/macintyre/macintyre2012catholic.html>

[279] —–. (2011) *After Virtue*, 3rd ed. Bloomsbury [2007].

[280] ——. (2009a) "Alasdair MacIntyre: The Illusion of Self-Sufficiency" [Interview with Alex Voorhoeve]. In *Conversations on Ethics,* edited by Alex Voorhoeve, 110-131. Oxford University Press.

[281] ——. (2009b) *Dependent Rational Animals.* Duckworth [1999].

[282] ——. (2008) "Conflicts of Desire". In *Weakness of Will from Plato to the Present*, edited by Tobias Hoffmann, 276-292. Catholic University of America Press.

[283] ——. (2006a) *Edith Stein: A Philosophical Prologue.* Continuum.

[284] ——. (2006b) "The End of Education: The Fragmentation of the American University". *Commonweal* 133, 18: 10-14.

[285] ——. (1998a) "Social Science Methodology as the Ideology of Bureaucratic Authority". In *The MacIntyre Reader*, edited by Kelvin Knight, 53-68. Polity Press.

[286] ——. (1998b) "Interview with Giovanna Borradori". In *The MacIntyre Reader*, edited by Kelvin Knight, 255-266. Polity Press.

[287] ——. (1993) *The Objectivity of Good.* St. Lawrence University.

[288] ——. (1992) "Plain Persons and Moral Philosophy: Rules, Virtues and Goods". *American Catholic Philosophical Quarterly* 66, no. 1: 3-19.

[289] ——. (1990) *Three Rival Versions of Moral Enquiry.* University of Notre Dame Press.

[290] ——. (1988) *Whose Justice? Which Rationality?* University of Notre Dame Press.

[291] ——. (1987) "Practical Rationalities as Forms of Social Structure". *Irish Philosophical Journal* 4, no. 1-2: 3-19.

[292] ——. (1982) "How Moral Agents Became Ghosts, or Why the History of Ethics Divided from that of the Philosophy of Mind". *Synthese* 53, no. 2: 295-312.

[293] ——. (1979) "Social Science Methodology as the Ideology of Bureaucratic Authority". In *Through the Looking Glass: Epistemology and the Conduct of Inquiry*, edited by Maria Falco, 42-58. University Press of America.

[294] Mayr, Ernst. (2004) *What Makes Biology Unique?* Cambridge University Press.

[295] Maialeh, Robin. (2019) "Generalization of Results and Neoclassical Rationality: Unresolved Controversies of Behavioural Economics Methodology". *Quality & Quantity* 53 no. 4: 1743-1761.

[296] Main, Mary and Judith Solomon. (1986) "Discovery of an Insecure-Disorganized/ Disoriented Attachment Pattern". In *Affective Development in Infancy*, edited by T. B. Brazelton and M. W. Yogman, 95-124. Ablex Publishing.

[297] Marshall, Alfred. (1920) *Principles of Economics*, 8th ed. Palgrave.

[298] Martin, Daniel, John Garske, and Katherine Davis. (2000) "Relation of the Therapeutic Alliance with Outcome and Other Variables: A Meta-Analytic Review". *Journal of Consulting and Clinical Psychology* 68, no. 3: 438-450.

[299] Mas-Colell, Andreu, Michael Whinston, and Jerry Green. (1995) *Microeconomic Theory: International Student Edition.* Oxford University Press.

[300] Mayer, Richard. (2004) "Should There Be a Three-Strikes Rule Against Pure Discovery Learning? The case for Guided Methods of Instruction". *American Psychologist* 59, no. 1: 14-19.

[301] McCaskey, John. (2020) "History of 'Temperature': Maturation of a Measurement Concept". *Annals of Science* 77, no. 4: 399-444.

[302] McDougall, Carol. (2002) "Rogers's Person-Centered Approach: Consideration For Use in Multicultural Counseling". *Journal of Humanistic Psychology* 42, no. 2: 48-65.

[303] McHale, James. (2007) "When Infants Grow Up in Multiperson Relationship Systems". *Infant Mental Health Journal* 28, no. 4: 370-392.

[304] McLean, Carmen and Edna Foa. (2011) "Prolonged Exposure Therapy for Post-Traumatic Stress Disorder: a Review of Evidence and Dissemination". *Expert Review of Neurotherapeutics* 11, no. 8: 1151-63.

[305] McLeod, John. (2009) *An Introduction to Counselling*, 4th ed. Open University Press.

[306] McMahan, Ethan and David Estes. (2011a) "Hedonic Versus Eudaimonic Conceptions of Well-being: Evidence of Differential Associations With Self-reported Well-being". *Social Indicators Research* 103, no. 1: 93-108.

[307] —. (2011b) "Measuring Lay Conceptions of Well-being: The Beliefs about Well-being Scale". *Journal of Happiness Studies* 12, no. 2: 267-287.

[308] Mead, George Herbert. (1934) *Mind, Self, and Society*, edited by Charles W. Morris. University of Chicago Press.

[309] Mearns, David and Brian Thorne. (2013) *Person-Centred Counselling in Action*, 4th ed. SAGE Publishing

[310] Meltzoff, Andrew and M. Keith Moore. (1983) "Newborn Infants Imitate Adult Facial Gestures". *Child Development* 54, no. 3: 702-709.

[311] —. (1977) "Imitation of Facial and Manual Gestures by Human Neonates". *Science* 198, no. 4312: 75-78.

[312] Meyer, Hilbert. (2004) *Was ist guter Unterricht?* Cornelsen.

[313] Mill, John Stuart. (2018) *Autobiography.* Oxford University Press [1873].

[314] —. (2000) "On the Definition of Political Economy; and on the Method of Investigation Proper to It". In *Essays on Some Unsettled Questions of Political Economy*, 2nd ed., 86-114. Batoche Books [1844].

[315] —. (1998) *Utilitarianism.* Oxford University Press [1863].

[316] —. (1989) *On Liberty.* In *On Liberty and Other Writings*, edited by Stefan Collini, 1-116. Cambridge University Press [1859].

[317] Mischel, Walter, Yuichi Shoda, and M. Rodriguez. (1989) "Delay of gratification in children". *Science* 244, 4907: 933-938.

[318] Mishan, E. J. (1969) *Cost-Benefit Analysis.* North-Holland.

[319] Moore, Andrew. (2013) “Hedonism”. In *The Stanford Encyclopedia of Philosophy*, Winter 2019 Edition, edited by Edward N. Zalta [2004]. <https://plato.stanford.edu/archives/win2019/entries/hedonism/>

[320] Moore, G. E. (1993) *Principia Ethica*. Cambridge University Press [1903].

[321] Morewedge, Carey, Daniel Gilbert, Kristian Myrseth, Karim Kassam, and Timothy D. Wilson. (2010) “Consuming Experience: Why Affective Forecasters Overestimate Comparative Value”. *Journal of Experimental Social Psychology* 46, no. 6: 986-992.

[322] Morewedge, Carey, Daniel Gilbert, and Timothy Wilson. (2005) “The Least Likely of Times: How Remembering the Past Biases Forecasts of the Future”. *Psychological Science* 16, no. 8: 626-630.

[323] Morgan, Mary. (2003) “Economics”. In *The Cambridge History of Science*, Vol. 7, edited by Theodore Porter and Dorothy Ross, 275-305. Cambridge University Press.

[324] Morris, Hannah, John Larsen, Emma Catterall, Antony Moss, and Stephan Dombrowski. (2020) “Peer Pressure and Alcohol Consumption in Adults Living in the UK: a Systematic Qualitative Rseview”. *BMC Public Health* 20, no. 1: 1014.

[325] Moscati, Ivan. (2019) *Measuring Utility: From the Marginal Revolution to Behavioral Economics*. Oxford University Press.

[326] Motta, Massimo. (2004) *Competition Policy: Theory and Practice*. Cambridge University Press.

[327] Muramatsu, Roberta. (2009) “The Death and Resurrection of ‘Economics with Psychology’: Remarks from a Methodological Standpoint”. *Revista de Economia Política* 29, no. 1: 62-81.

[328] Muran, John and Jacques Barber, eds. (2010) *The Therapeutic Alliance: An Evidence-Based Guide to Practice*. Guilford Press.

[329] Murray, Lynne and Colwyn Trevarthen. (1986) “The infant’s role in mother-infant communications”. *Journal of Child Language* 13, no. 1: 15-29.

[330] Nadel, Jacqueline and George Butterworth, eds. (1999) *Imitation in infancy.* Cambridge University Press.

[331] Nesse, Randolph. (1991) "What Good is Feeling Bad?: The Evolutionary Benefits of Psychic Pain". *The Sciences*, November/December: 30-37.

[332] Neumann, John von, and Oskar Morgenstern. (2004) *Theory of Games and Economic Behavior: Sixtieth-Anniversary Edition*, 3rd ed. Princeton University Press [1953].

[333] Nietzsche, Friedrich. (2005) *Thus Spoke Zarathustra.* Oxford University Press [1883-1885].

[334] —–. (2002) *Beyond Good and Evil.* Edited by Rolf-Peter Horstmann and translated by Judith Norman. Cambridge University Press [1986].

[335] —–. (1997) *Untimely Meditations.* Edited by Daniel Breazeale and translated by R. J. Hollingdale [1873-1876].

[336] Nowell-Smith, Patrick. (1954) *Ethics.* Penguin Books.

[337] Nozick, Robert. (1981) *Philosophical Explanations.* Harvard University Press.

[338] Nussbaum, Martha. (2011) *Creating Capabilities: The Human Development Approach.* Harvard University Press.

[339] —–. (2001) *The Fragility of Goodness*, 2nd ed. Cambridge University Press [1986].

[340] O'Connor, James. (2012) "The Trolley Method of Moral Philosophy". Essays in Philosophy 13, no. 1: 243-256.

[341] O'Dea, Jane. (1993) "Teaching as Entertainment: A Comment on a 20th Century Phenomenon". *Education Canada* 33, no. 3: 24-29.

[342] O'Dea, Rose, Daniel Noble, and Shinichi Nakagawa, 2022, "Unifying Individual Differences in Personality, Predictability and Plasticity: A Practical Guide". *Methods in Ecology and Evolution* 13, no. 2: 278-293.

[343] Ogden, Simon, Richard Huxtable, and Jonathan Ives. (2020) "Protocol For a Scoping Review to Understand What is Known About How GPs Make Decisions With, For and on Behalf of Patients Who Lack Capacity." *BMJ open* 10, no. 10: e038032.

[344] Ok, Efe. (2002) "Utility Representation of an Incomplete Preference Relation". *Journal of Economic Theory* 104, no. 2: 429-449.

[345] Panjabi, Camellia. (1994) *50 Great Curries of India.* Kyle Books.

[346] Panksepp, Jaak. (1998) *Affective Neuroscience.* Oxford University Press.

[347] Parfit, Derek. (1984) *Reasons and Persons.* Oxford University Press.

[348] Patterson, Cecil. (1984) "Empathy, Warmth, and Genuineness in Psychotherapy: A Review of Reviews". *Psychotherapy: Theory, Research, Practice, Training* 21, no. 4: 431-438.

[349] Pereboom, Derk. (2018) "Kant's Transcendental Arguments". In *The Stanford Encyclopedia of Philosophy*, Fall 2021 Edition, edited by Edward N. Zalta [2009]. <https://plato.stanford.edu/archives/spr2019/entries/kant-transcendental/>

[350] Persons, Jacqueline. (2008) *The Case Formulation Approach to Cognitive-Behavior Therapy.* Guilford Publications.

[351] Pettit, Philip. (2018) *The Birth of Ethics: Reconstructing the Role and Nature of Morality.* Oxford University Press.

[352] Pfeifer, Rolf, Josh Bongard, and Don Berry. (2011) *Designing Intelligence.* GRIN Publishing GmbH.

[353] Piantadosi, Steven and Celeste Kidd. (2016) "Extraordinary Intelligence and the Care of Infants". *PNAS* 113, no. 25: 6874-6879.

[354] Pizzagalli, Diego, Allison Jahn, and James Shea. (2005) "Toward an Objective Characterization of an Anhedonic Phenotype: A Signal-Detection Approach". *Biological Psychiatry* 57, no. 4: 319-327.

[355] Porter, Theodore. (1994) "Rigor and Practicality: Rival Ideals of Quantification in Nineteenth-Century Economics". In *Natural*

Images in Economic Thought, edited by Philip Mirowski, 128-170. Cambridge University Press.

[356] Prior, Vivien and Danya Glaser. (2006) *Understanding Attachment and Attachment Disorders: Theory, Evidence and Practice*. Jessica Kingsley Publishers.

[357] Prinzing, Michael. (2021) "How to Study Well-Being: A Proposal for the Integration of Philosophy With Science". *Review of General Psychology* 25, no. 1: 152-162.

[358] Quinn, Adam. (2015) *A Person-centered Approach and the Rogerian Tradition: A Handbook.* CreateSpace Independent Publishing Platform.

[359] Quinn, Warren. (1993) *Morality and Action.* Cambridge University Press.

[360] Rabin, Matthew. "Incorporating Fairness into Game Theory and Economics". *The American Economic Review* 83, no. 5: 1281-1302.

[361] Rabin, Matthew, and Richard Thaler. (2001) "Risk Aversion". *Journal of Economic Perspectives* 15, no. 1: 219-232.

[362] Ramsey, Frank P. (1931) "Truth and Probability". In *The Foundations of Mathematics and other Logical Essays*, edited by R. B. Braithwaite, 156-198. Kegan, Paul, Trench, Trubner & Co [1926].

[363] —–. (1928) "A Mathematical Theory of Saving". *Economic Journal* 38, no. 4: 543-559.

[364] Rawls, John. (1971) *A Theory of Justice*, 1st ed. Harvard University Press.

[365] Raz, Joseph. (2004) "The Role of Well-being". *Philosophical Perspectives* 18, no. 1: 269-294.

[366] —–. (1986) *The Morality of Freedom.* Clarendon Press.

[367] Read, Daniel and Barbara Van Leeuwen. (1998) "Predicting Hunger: The Effects of Appetite and Delay on Choice". *Organizational Behavior and Human Decision Processes* 76, no. 2: 189-205.

[368] Redelmeier, Dondald and Daniel Kahneman. (1996) “Patients’ Memories of Painful Medical Treatments: Real-Time and Retrospective Evaluations of Two Minimally Invasive Procedures”. *Pain* 66, no. 1: 3-8.

[369] Reindl, Vanessa, Christian Gerloff, Wolfgang Scharke, and Kerstin Konrad. (2018) “Brain-to-brain Synchrony in Parent-Child Dyads and the Relationship with Emotion Regulation Revealed by fNIRS-Based Hyperscanning”. *Neuroimage* 178: 493-502.

[370] Resnik, Michael D. (1987) *Choices: An Introduction to Decision Theory.* University of Minnesota Press.

[371] Rieff, Philip. (2008) *The Jew of Culture: Freud, Moses, and Modernity*, Vol. 3. University of Virginia Press.

[372] Rittel, Horst and Melvin Webber. (1973) “Dilemmas in a General Theory of Planning”. *Policy Sciences* 4, no. 2: 155-169.

[373] Robbins, Lionel. (1935) *An Essay on the Nature and Significance of Economic Science*, 2nd ed. Macmillan. [1932].

[374] Robertson, Donald. (2018) *The Philosophy of Cognitive-Behavioural Therapy (CBT): Stoic Philosophy as Rational and Cognitive Psychotherapy.* Taylor & Francis.

[375] Rogers, Carl R. (1957) “The Necessary and Sufficient Conditions of Therapeutic Personality Change”. *Journal of Consulting Psychology* 21, no. 2: 95-103.

[376] —–. (1951) *Client-Centered Therapy: Its Current Practice, Implications, and Theory.* Houghton Mifflin.

[377] Rosati, Connie. (2006) “Preference-Formation and Personal Good”. *Royal Institute of Philosophy Supplement* 81, no. 59: 33-64.

[378] —–. (1995) “Persons, Perspectives, and Full Information Accounts of the Good”. *Ethics* 105, no. 2: 296-325.

[379] Ross, W. D. (2002) *The Right and the Good.* Clarendon Press [1930].

[380] Rothbaum, Barbara, Elizabeth Meadows, Patricia Resick, and David Foy. (2000) “Cognitive-Behavioral Therapy”. In *Effective*

Treatments for PTSD: Practice Guidelines from the International Society for Traumatic Stress Studies, edited by E. B. Foa, T. M. Keane and M. J. Friedman, 320-325. The Guilford Press.

[381] Rubin, Nissan, Carmella Shmilovitz, and Meira Weiss. (1993) "From Fat to Thin: Informal Rites Affirming Identity Change." *Symbolic Interaction* 16, no. 1: 1-17.

[382] Russell, Bertrand. (1999) "What Desires are Politically Important?" In *Nobel Lectures: Literature, 1901-1967*, edited by Horst Frenz, 452-464. World Scientific [1950].

[383] Rutter, Michael. (1998) "Developmental Catch-up, and Deficit, Following Adoption after Severe Global Early Privation". *The Journal of Child Psychology and Psychiatry and Allied Disciplines* 39, no. 4: 465-476.

[384] Sagan, Carl. (1986) *The Dragons of Eden.* Ballantine Books [1977].

[385] Sagi, Abraham, Marinus H. van IJzendoorn, Ora Aviezer, Frank Donnell,Ofra Mayseless. (1994) "Sleeping Out of Home in a Kibbutz Communal Arrangement: It Makes a Difference for Infant-Mother Attachment". *Child Development* 65, no. 4: 992-1004.

[386] Salomon, M., E. D. Aflalo, M. Coll and Y. Lubin. (2015) "Dramatic Histological Changes Preceding Suicidal Maternal Care in the Subsocial Spider *Stegodyphus Lineatus* (Araneae: Eresidae)". *Journal of Arachnology* 43, no. 1: 77-85.

[387] Sapolsky, Robert. (2017) *Behave: The Biology of Humans at Our Best and Worst.* Penguin Press.

[388] Samuelson, Paul. (1938) "A Note on the Pure Theory of Consumers' Behaviour". *Economica* 5, no. 17: 61-71.

[389] Sandel, Michael J. (2013) "Market Reasoning as Moral Reasoning: Why Economists Should Re-engage with Political Philosophy". *Journal of Economic Perspectives* 27, no. 4: 121-140.

[390] Savage, Leonard. (1954) *The Foundations of Statistics.* John Wiley and Sons.

[391] Sawatzky, Alexandra, Ashlee Cunsolo, Sherilee Harper, Inez Shiwak, Michele Wood. (2019) "“We Have Our Own Way”: Exploring Pathways for Wellbeing Among Inuit in Nunatsiavut". In *The Routledge Handbook of Indigenous Well-being*, edited by Christopher Fleming and Matthew Manning, 223-236. Routledge.

[392] Scanlon, Thomas. (1998) *What We Owe to Each Other.* Harvard University Press.

[393] —. (1993) "Value, Desire, and Quality of Life". In *The Quality of Life*, edited by Martha Nussbaum and Amartya Sen, 185-200. Clarendon Press.

[394] Schaffer, Heinz. (2007) *Introducing Child Psychology*. Blackwell.

[395] Schaffer, Heinz and Peggy E Emerson. (1964) "The Development of Social Attachments in Infancy." *Monographs of the Society for Research in Child Development* 29, no. 3: 1-77.

[396] Schirmer, Annett, Merle Fairhurst, and Stefanie Hoehl. (2021) "Being 'In Sync' – Is Interactional Synchrony the Key to Understanding the Social Brain?" *Social Cognitive and Affective Neuroscience* 16, no. 1-2: 1-4.

[397] Schkade, David and Daniel Kahneman. (1998) "Does Living in California Make People Happy? A Focusing Illusion in Judgments of Life Satisfaction". *Psychological Science* 9, no. 5: 340-346.

[398] Schliesser, Eric. (2019) "Synthetic philosophy". *Biology and Philosophy* 34, no. 2: 19.

[399] Schore, Alan. (2000) "Attachment and the Regulation of the Right Brain". *Attachment & Human Development* 2, no. 1: 23-47.

[400] —. (1996) "The Experience-Dependent Maturation of a Regulatory System in the Orbital Prefrontal Cortex and the Origin of Developmental Psychopathology". *Development and Psychopathology* 8, no. 1: 59-87.

[401] Schreiber, Charles and Daniel Kahneman. (2000) "Determinants of the Remembered Utility of Aversive Sounds". *Journal of Experimental Psychology* 129, no. 1: 27-42.

[402] Schroeder, Tim. (2015) "Desire". In *The Stanford Encyclopedia of Philosophy*, Summer 2020 edition, edited by Edward N. Zalta [2009].
<https://plato.stanford.edu/archives/sum2020/entries/desire/>

[403] Schunk, Dale and Ellen Usher. (2012) "Social Cognitive Theory and Motivation". In *The Oxford Handbook of Motivation*, edited by Richard Ryan, 13-27. Oxford University Press.

[404] Sen, Amartya. (1993) "Capability and Well-being". In *The Quality of Life*, edited by Martha Nussbaum and Amartya Sen, 30-53. Clarendon Press.

[405] —–. (1977) "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory". *Philosophy & Public Affairs* 6, no. 4: 317-344.

[406] Sen, Amartya and Bernard Williams, eds. (1982) *Utilitarianism and Beyond*. Cambridge University Press.

[407] Seth, Anil K. (2012) "Optimised Agent-Based Modelling of Action Selection". In *Action Selection,* edited by Anil Seth, Tony Prescott, and Joanna Bryson, 37-60. Cambridge University Press.

[408] Sheehy, Gail. (1996) *New Passages: Mapping Your Life Across Time*. Ballantine Books.

[409] Shiv, Baba and Alexander Fedorikhin. (2002) "Spontaneous versus Controlled Influences of Stimulus-Based Affect on Choice Behavior". *Organizational Behavior and Human Decision Processes* 87, no. 2: 342-370.

[410] Shlien, John. (1992) "Theory as Autobiography: The Man and the Movement". *Contemporary Psycholog*y 37, no. 19: 1082-1084.

[411] Sidgwick, Henry. (1981) *The Methods of Ethics,* 7th ed. Hackett [1907].

[412] Siedentop, Larry. (2014) *Inventing the Individual.* Harvard University Press.

[413] Siegel, Harvey. (1995) "Why Should Educators Care about Argumentation?" *Informal Logic* 17, no. 2: 159-176.

[414] Simon, Herbert. (1955) "A Behavioral Model of Rational Choice". *The Quarterly Journal of Economics* 69, no. 1: 99-118.

[415] —–. (1947) *Administrative Behavior*. Macmillan.

[416] Simonson, Itamar. (1990) "The Effect of Purchase Quantity and Timing on Variety Seeking Behavior". *Journal of Marketing Research* 27, no. 2, 150-162.

[417] Singer, Peter. (1979) *Practical Ethics*. Cambridge University Press.

[418] Skelton, Anthony. (2016) "Children's Well-being: A Philosophical Analysis". In *The Routledge Handbook of the Philosophy of Well-being*, edited by Guy Fletcher, 366-377. Routledge.

[419] Skinner, B. F. (1965) *Science and Human Behaviour.* The Free Press [1953].

[420] Slater, Alan, Scott Johnson, and Darwin Muir. (2017) "The Scope and Methods of Developmental Psychology". In *An Introduction to Developmental Psychology*, 3rd ed, edited by Alan Slater, and J. Gavin Bremner, 3-40. Wiley.

[421] Smart, J. J. C. and Bernard Williams. (1973) *Utilitarianism: For and Against*. Cambridge University Press.

[422] Smith, Joshua. (2004) "The Effects of Student Receptivity on College Achievement and Retention". *Journal of College Student Retention: Research, Theory & Practice* 6, no. 3: 251-271.

[423] Snead, O. Carter. (2020) *What It Means to Be Human: The Case for the Body in Public Bioethics*. Harvard University Press.

[424] Sober, Elliott. (1998) "Innate knowledge". In *Routledge Encyclopedia of Philosophy*, Vol. 4, edited by Edward Craig, 794-797. Routledge.

[425] Sollod, Robert. (1978) "Carl Rogers and the Origins of Client-Centered Therapy". *Professional Psychology* 9, no. 1: 93-104.

[426] Somerville, James. (1989) "Collingwood's Logic of Question and Answer". *The Monist* 72, no. 4: 526-541.

[427] Spinhoven, Philip, Jolijn Drost, Mark de Rooij, Albert van Hemert, and Brenda Penninx. (2014) "A Longitudinal Study of Experiential Avoidance in Emotional Disorders". *Behavior Therapy* 45, no. 6: 840-850.

[428] Smith, Linda and Joseph Renzulli. (1984) "Learning Style Preferences: A Practical Approach for Classroom Teachers". *Theory Into Practice* 23, no. 1: 44-50.

[429] Sripada, Chandra. (2022) "Mental Disorders Involve Limits on Control, not Extreme Preferences". In *Agency in Mental Disorder: Philosophical Dimensions*, edited by Matt King and Joshua May, 169-192. Oxford University Press.

[430] Sroufe, L. Alan. (1996) *Emotional Development: The Organization of Emotional Life in the Early Years*. Cambridge University Press.

[431] Sroufe, L. Alan, Brianna Coffino, and Elizabeth A. Carlson. (2010) "Conceptualizing the Role of Early Experience: Lessons from the Minnesota Longitudinal Study". *Developmental Review* 30, no. 1: 36-51.

[432] Stacy, Alan, Michael Newcomb, and Peter Bentler. (1991) "Social Psychological Influences on Sensation Seeking From Adolescence to Adulthood". *Journal of Personality and Social Psychology* 61, no. 4: 614-628.

[433] Stanley, Thomas. (1656) *The History of Philosophy*, Vol. 2. J. Moseley and T. Dring.

[434] Staples, Megan, Joanna Bartlo, and Eva Thanheiser. (2012) "Justification as a Teaching and Learning Practice: Its (Potential) Multifaceted Role in Middle Grades Mathematics Classrooms". *The Journal of Mathematical Behavior* 31, no. 4: 447-462.

[435] Steffe Leslie P. and Jerry Gale, eds. (1995) *Constructivism in Education*. Routledge.

[436] Stigler, George J. and Becker, Gary S. (1977) "*De Gustibus Non Est Disputandum*". *The American Economic Review* 67, no. 2: 76-90

[437] Steele, Katie. (2014) "Choice models". In *Philosophy of Social Science: A New Introduction*, edited by Nancy Cartwright and Eleonora Montuschi, 185-207. Oxford University Press.

[438] Steele, Katie and H. Orri Stefánsson. (2020) “Decision Theory”. In *The Stanford Encyclopedia of Philosophy*, Winter 2020 edition, edited by Edward N. Zalta [2015]. <https://plato.stanford.edu/archives/win2020/entries/decision-theory/>

[439] Steiner, George F. (2017) *Neanderthals in Plato’s Cave: A Relativistic Approach to Cultural Evolution.* Nova Science Publishers.

[440] Stern, Robert. (2019) “Transcendental Arguments”. In *The Stanford Encyclopedia of Philosophy*, Fall 2021 Edition, edited by Edward N. Zalta [2011]. <https://plato.stanford.edu/archives/fall2021/entries/transcendental-arguments/>

[441] Stewart, Neil. (2009) “Decision by Sampling: The Role of the Decision Environment in Risky Choice”. *The Quarterly Journal of Experimental Psychology* 62, no. 6: 1041-1062.

[442] Stocker, Michael. (1976) “The Schizophrenia of Modern Ethical Theories”. *The Journal of Philosophy* 73, no. 14: 453-466.

[443] Strathern, Marilyn. (2016) “‘Improving Ratings’: Audit in the British University system”. *European Review* 5, no. 3: 305-321.

[444] Sugden, Robert. (2018) *The Community of Advantage: A Behavioural Economist’s Defence of the Market.* Oxford University Press.

[445] Sumner, L. W. (1996) *Welfare, Happiness, and Ethics.* Oxford University Press.

[446] Swanton, Christine. (2014) “Cultivating Virtue: Two Problems for Virtue Ethics”. In *Cultivating Virtue*, edited by Nancy Snow, 111-134. Oxford University Press.

[447] Tal, Eran. (2015) “Measurement in Science”. In *The Stanford Encyclopedia of Philosophy*, Fall 2020 Edition, edited by Edward N. Zalta [2015]. <https://plato.stanford.edu/archives/fall2020/entries/measurement-science/>

[448] Talbot, Brian. (2021) "Why Impossible Options Are Better: Consequentializing Dilemmas". *Utilitas* 33, no. 2: 221-236.

[449] Tartter, Vivien. (1998) *Language Processing in Atypical Populations.* SAGE Publishing.

[450] Tate, Tyler. (2020) "What We Talk About When We Talk About Pediatric Suffering". *Theoretical Medicine and Bioethics* 41, no. 4: 143-163.

[451] Taylor, Charles. (1992) *The Ethics of Authenticity.* Harvard University Press.

[452] —–. (1989) *Sources of the Self: The Making of Modern Identity.* Harvard University Press.

[453] Temkin, Larry. (2012) *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning.* Oxford University Press.

[454] —–. (1996) "A Continuum Argument for Intransitivity". *Philosophy & Public Affairs* 25, no. 3: 175-210.

[455] Tiberius, Valerie. (2013) "Well-being, Wisdom, and Thick Theorizing: On the Division of Labor Between Moral Philosophy and Positive Psychology". In *Thick concepts*, edited by Simon Kirchin, 217-233. Oxford University Press.

[456] Tiffany, Evan. (2012) "Why Be an Agent?" *Australasia Journal of Philosophy* 90, no. 2: 223-233.

[457] Thaler, Richard and Cass Sunstein. (2008) *Nudge: Improving Decisions about Health, Wealth, and Happiness.* Yale University Press.

[458] —–. (2003) "Libertarian Paternalism". *The American Economic Review* 93, no. 2: 175-179.

[459] Thoma, Johanna. (2021) "In Defence of Revealed Preference Theory". *Economics & Philosophy* 37, no. 2: 163-187.

[460] Tomlin, Patrick. (2016) "Saplings or Caterpillars? Trying to Understand Children's Wellbeing". *Journal of Applied Philosophy* 35, no. S1: 29-46.

[461] Toner, Christopher. (2006) "Aristotelian Well-Being: A Response to L. W. Sumner's Critique". *Utilitas* 18, no. 3: 218-231.

[462] Toomey, James. (2021) "Understanding the Perspectives of Seniors on Dementia and Decision-Making". AJOB Empirical Bioethics 2, no. 2: 101-112.

[463] Tremblay, Léon and Wolfram Schultz. (1999) "Relative Reward Preference in Primate Orbitofrontal Cortex". *Nature* 398, no. 6729: 704-708.

[464] Truax, Charles. (1966) "Reinforcement and Nonreinforcement in Rogerian Psychotherapy". *Journal of Abnormal Psychology* 71, no. 1: 1-9.

[465] Tversky, Amos and Daniel Kahneman. (1992) "Advances in Prospect Theory: Cumulative Representation of Uncertainty". *Journal of Risk and Uncertainty* 5, no. 4: 297-323.

[466] Tversky, Amos and Richard Thaler. (1990) "Preference Reversals". *Journal of Economic Perspectives* 4, no. 2: 201-211.

[467] Vallacher, Robin and Daniel Wegner. (1987) "What Do People Think They're Doing? Action Identification and Human Behavior". *Psychological Review*, 94, no. 1: 3-15.

[468] Varian, Hal. (1982) "The Nonparametric Approach to Demand Analysis". *Econometrica* 50, no. 4: 945-973.

[469] Velleman, David. (2009) *How We Get Along*. Cambridge University Press.

[470] Voorhoeve, Alex. (2018) "Epicurus on Pleasure, a Complete Life, and Death: A Defence". *Proceedings of the Aristotelian Society* 118, no. 3: 225-253.

[471] Vredenburgh, Kate. (2020) "A Unificationist Defense of Revealed Preferences". *Economics & Philosophy* 36, no. 1: 149-169.

[472] Vygotsky, Lev. (1978) *Mind in Society: Development of Higher Psychological Processes*. Edited by Michael Cole, Vera John-Steiner, Sylvia Scribner, and Ellen Souberman. Harvard University Press.

[473] Waddington, C. H. (1957) *The Strategy of the Genes.* London: Allen and Unwin.

[474] Walker, Sarah, Kit Double, Hannah Kunst, Michael Zhang, and Carolyn MacCann. "Emotional Intelligence and Attachment in Adulthood: A Meta-Analysis". *Personality and Individual Differences* 184, no. 2: 111174.

[475] Wallace, R. Jay. (2020) "Practical Reason". In *The Stanford Encyclopedia of Philosophy*, Spring 2020 edition, edited by Edward N. Zalta [2003]. <https://plato.stanford.edu/archives/spr2020/entries/practical-reason/>

[476] Waller, Robyn. (2022) "Taking Control with Mechanisms of Psychotherapy". In *Agency in Mental Disorder: Philosophical Dimensions*, edited by Matt King and Joshua May, 86-112. Oxford University Press.

[477] Waters, Everett and E. Mark Cummings. (2000) "A Secure Base From Which to Explore Close Relationships". *Child Development* 71, no. 1: 164-72.

[478] Watts, Alan. (1966) *The Book: On the Taboo Against Knowing Who You Are.* Pantheon Books.

[479] Way, Jonathan, (2017) "Reasons as Premises of Good Reasoning". *Pacific Philosophical Quarterly* 98, no. 2: 251-270.

[480] Weber, Max. (2012) "Objectivity and Understanding in Economics". In *The Philosophy of Economics: An Anthology*, 3rd ed., edited by Daniel Hausman, 59-72. Cambridge University Press [1904].

[481] Weijers, Dan. (2014) "Nozick's Experience Machine is Dead, Long Live the Experience Machine!" *Philosophical Psychology*, 27, no. 1: 513-535.

[482] —–. (2013) "Intuitive Biases in Judgements about Thought Experiments: The Experience Machine Revisited". *Philosophical Writings* 41, no. 1: 17-31.

[483] Weinert, Franz. (2001) "Competencies and Key Competencies: Educational perspective". In *International Encyclopedia of the*

Social and Behavioral Sciences, edited by N. J. Smelser & B. Baltes, 2433-2436. Elsevier.

[484] Westbrook, David, Helen, Kennerley, and Joan Kirk. (2007) *An Introduction to Cognitive Behaviour Therapy: Skills & Applications.* SAGE Publishing.

[485] Whewell, William. (1862) “Additional Lectures on the History of Moral Philosophy”. In *Lectures on the History of Moral Philosophy.* Cambridge University Press.

[486] Wilkes, Kathleen. (1978) “The Good Man and the Good for Man in Aristotle’s Ethics”. *Mind* 87, no. 348: 553-571.

[487] Williams, Bernard. (2006) *Ethics and the Limits of Philosophy.* Routledge [1985].

[488] —–. (1979) “Internal and External Reasons”. In *Rational Action*, edited by Ross Harrison, 17-28. Cambridge University Press.

[489] Wilson, Jan and John Rhodes. (2010) “Student Perspectives on Homework”. *Education* 131, no. 2: 351-358.

[490] Wilson, Timonthy and Daniel Gilbert. (2003) “Affective forecasting”. In *Advances in Experimental Social Psychology*, vol. 35, edited by Mark Zanna, 345-411. Elsevier Academic Press.

[491] Winnicott, Donald. (2002) ‘Saying “No”’. In *On the Child*, 108-121. Da Capo Press.

[492] —–. (1964) “Further Thoughts on Babies as Persons”. In *The Child, the Family, and the Outside World*, 85-92. Penguin Books [1947].

[493] Witter, Robert A., Morris A. Okun, William A. Stock, and Marilyn J. Haring. (1984) “Education and Subjective Well-Being: A Meta-Analysis”. *Educational Evaluation and Policy Analysis* 6, no. 2: 165-173.

[494] Wittgenstein, Ludwig. (2003) *Philosophical Investigations: 50th Anniversary Commemorative Edition.* Blackwell [1953].

[495] Wolpe, Joseph. (1958) *Psychotherapy by Reciprocal Inhibition.* Stanford University Press.

[496] Wood, Allen. (1998) "Hegel on Education". In *Philosophy as Education*, edited by Amélie Rorty, 300-317. Routledge.

[497] Young, Andrew and Mike Burton. (2017) "Recognizing Faces". *Current Directions in Psychological Science* 26, no. 3: 212-217.

[498] Zeveloff, Samuel and Mark S. Boyce. (1982) "Why Human Neonates Are So Altricial". *The American Naturalist* 120, no. 4: 537-542.

[499] Zimmerman, Michael J. and Ben Bradley. (2019) "Intrinsic vs. Extrinsic Value". In *The Stanford Encyclopedia of Philosophy*, Spring 2019 edition, edited by Edward N. Zalta [2002]. <https://plato.stanford.edu/archives/spr2019/entries/value-intrinsic-extrinsic/>

[500] Zuckerman, Martin. (2007) *Sensation Seeking and Risky Behavior.* American Psychological Association.