

Human Motives



Human Motives

*Hedonism, Altruism, and the
Science of Affect*

Peter Carruthers

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

© Peter Carruthers 2024

The moral rights of the author have been asserted

All rights reserved. No part of this publication may be reproduced, stored in
a retrieval system, or transmitted, in any form or by any means, without the
prior permission in writing of Oxford University Press, or as expressly permitted
by law, by licence or under terms agreed with the appropriate reprographics
rights organization. Enquiries concerning reproduction outside the scope of the
above should be sent to the Rights Department, Oxford University Press, at the
address above

You must not circulate this work in any other form
and you must impose this same condition on any acquirer

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2023942445

ISBN 978-0-19-890613-1

DOI: 10.1093/oso/9780198906131.001.0001

Printed and bound in the UK by
Clays Ltd, Elcograf S.p.A.

Links to third party websites are provided by Oxford in good faith and
for information only. Oxford disclaims any responsibility for the materials
contained in any third party website referenced in this work.

For Fred and Blythe
Friends, family, and joy in nature



Contents

<i>Preface</i>	ix
1. Introduction	1
1.1 The Theoretical Terrain	1
1.2 Why It Matters	4
1.3 The Way Ahead	7
2. The Science of Affect	10
2.1 The Scope of Affect	11
2.2 The Input: Appraisals of Value	12
2.3 Output (1): Behavior	15
2.4 Output (2): Arousal	19
2.5 Output (3): Valence	20
2.6 Output (4): Affective Learning	26
2.7 Top-Down Influences on Affect	34
2.8 Conclusion	39
3. The New Hedonism	40
3.1 Prospective Decision-Making	40
3.2 How Valence Works	43
3.3 Future-Directed Pleasure	47
3.4 Concerns about Consciousness	49
3.5 Other Forms of Decision-Making	54
3.6 Difficult Decisions	60
3.7 Conclusion	62
4. Traditional Critiques Critiqued	64
4.1 Evolutionary Arguments	64
4.2 Self-Sacrifice: Spontaneous and Reflective	68
4.3 Desire and Satisfaction	72
4.4 The Experience Machine	74
4.5 Experimental Evidence	76
4.6 Two Forms of Hedonism	81
4.7 Conclusion	83
5. Feelings Versus Representations of Value	85
5.1 Valence as Feeling: An Initial Concern	85
5.2 Valence as Representation of Value	86
5.3 Valence and Consciousness	92
5.4 Theories of Representation	95
5.5 Erroneous Emotions	100

viii Contents

5.6	Erroneous Pleasures and Displeasures	106
5.7	Merely Sub-Personal or Extended Egoism: A Dilemma	109
5.8	Conclusion	113
6.	Indicatives Versus Imperatives	115
6.1	Two Forms of Imperativism	115
6.2	Imperative Versus Indicative Content	121
6.3	Consummatory Pleasure Is about the Present	127
6.4	Valence and Imperatives in the Brain	132
6.5	Alleged Counter-Examples to Evaluativism	134
6.6	Conclusion	138
7.	Pain: A Test Case	140
7.1	The Science of Pain: A Primer	140
7.2	The Sensation of Pain	145
7.3	The Painfulness of Pain	152
7.4	Erroneous Pains	155
7.5	Pain's Transparency	157
7.6	Pains as Reasons	159
7.7	Conclusion	163
8.	Moral Motivation	164
8.1	Kinds of Moral Motivation	164
8.2	Prosociality and Altruism	166
8.3	Equality and Fairness	171
8.4	Social-Group Cognition	177
8.5	Norm Psychology	182
8.6	Conclusion: Scaling Up	188
8.7	Addendum: Weak Humeanism	189
8.8	Summary: The Case against Hedonism	191
	<i>References</i>	195
	<i>Index</i>	219

CF

Preface

CFP1 This book has been a while in the making, even if not nearly so long in the writing. I first began to get interested in the science of affective states like desire and emotion around 2009, when I was writing a book on self-knowledge and needed to say something about our knowledge of our own such states (Carruthers 2011). Luckily, since that project was funded by the National Science Foundation, I was able to have Brendan Ritchie working with me as a research assistant for the year. He knew a lot more about the field than I did. (Brendan had worked as an undergraduate research assistant for Tim Schroeder while the latter was writing his book, *The Three Faces of Desire*.) As a result, he was able to put me right on a lot of points, correcting my many misunderstandings of the field.

CFP2 Thereafter I left the topic aside for the next few years while I worked on other issues, having to do with consciousness, working memory, and conscious thought (Carruthers 2015, 2019). But I then began to do serious work in the area again, focused initially on epistemic emotions like uncertainty (Carruthers 2017) and curiosity (Carruthers 2018d, 2020, 2024). But my reading of the field had led me to think that the nature of positive and negative valence (arguably, pleasure and displeasure by other names), which are described by many as the “common currency” of decision-making, is critical in evaluating the debate between motivational hedonism and its critics. (Motivational hedonism is often called “psychological hedonism” in the philosophical literature; it is the doctrine that all human actions are undertaken to secure good and avoid bad experiences for oneself.) For on some views, it appears that decisions taken to maximize the balance of positive over negative valence are about securing good (and avoiding bad) feelings for oneself; whereas on other views valence is an imperative-like urge to secure or avoid specific types of experience. Both kinds of account seem to lead straight to motivational hedonism. In contrast, I came to feel that valence is best understood representationally, serving to *represent* actions or outcomes as good or bad. This sort of account would warrant motivational pluralism, allowing that people can have many goals besides their own pleasant experiences and avoidance of their own unpleasant ones. The result was a pair of papers (Carruthers 2018c, 2023) which can be regarded as preliminary studies for the present book. Note that these papers were published

x Preface

after a delay of two and three years, respectively, following final acceptance by the journals in question. (That is, they were completed and accepted in 2016 and 2020, respectively.) So they are not as recent as they may seem. As a result, my views have matured and altered quite a bit since then.

CFP3 Work on this book was facilitated by a Scholar's Award from NSF's *Science and Technology Studies* program (award # 2143473) covering the academic year 2022–23, as well as by a sabbatical semester from the University of Maryland in Spring 2022. I am grateful to both institutions for their support. Valuable initial feedback on some of the ideas presented in this book was provided by a group of graduate students at Maryland who attended a seminar I taught on the topic in Fall of 2021. (Special thanks go to Ken Glazer, who wrote a thoughtful critique of some my initial formulations.) In addition, I am grateful to the following individuals for their comments, criticism, and advice on some earlier drafts of this material: Luca Barlassina, Lia Curtis Fine, Joe Gurrola, Chris Masciari, Lixing Miao, Dan Moller, Shen Pan, Aida Roige, Elizabeth Schechter, Nicholas Shea, Julius Schönherr, Rachel Singpurwalla, Moonyoung Song, Louis Trost, and Xiaohui Yu. I am especially grateful to two anonymous reviewers for Oxford University Press for their thoughtful comments on the manuscript. In addition, some of the material in Chapter 6 is drawn from my 2023 paper cited above. I am grateful to the editors and publisher for their permission to make use of it.

C1

1

Introduction

C1P1 This short opening chapter introduces our topic, drawing distinctions and elucidating the main claims that are at stake. It also describes why these issues matter, as well as outlining what I take to be the appropriate methodology for resolving them.

C1S1 **1.1 The Theoretical Terrain**

C1P2 Why do human beings do what they do? What are the ultimate well-springs of human action? For many of us the answers to these questions seem obvious. People do what they do for all sorts of reasons, and different individuals have all kinds of intrinsic goals in life. This is *pluralism* about human motives. Some people want to be rich (for its own sake, not only as a means to other goods). Some want fame. Some want the respect of their colleagues. Many want their children and spouses to be healthy and happy. Some want to save the whale, or to preserve Monarch butterflies; others to protect ancient monuments or wilderness lands. Most pluralists will also allow that people (or at any rate, some people) are intrinsically motivated to do what they believe to be morally right—or morally required—and to avoid doing what they take to be morally wrong or forbidden. Moreover, pluralists will likely agree that many people are also, to some degree, altruistic—wanting to benefit or reduce the suffering of at least some other people, and wanting this for its own sake.

C1P3 Contrasting with motivational pluralism is motivational *egoism*. (Philosophers generally refer to this view as “*psychological egoism*,” although this is somewhat less descriptively accurate.) This is the doctrine that all human actions are undertaken, ultimately, to insure the actor’s own benefit, where benefits can take various forms, including health, wealth, fame, and respect. A more specific form of motivational egoism, however (and the version this book will mostly be concerned with), is motivational *hedonism*. This is the view that all human actions are aimed at securing good experiences and/or avoiding bad experiences for oneself. Hedonism cashes out the “benefit to self” of egoism in experiential terms. Put differently: all actions

2 Human Motives

are selected to achieve one's own pleasure and/or to avoid one's own displeasure—where pleasure and displeasure are understood broadly, to include not just good and bad feelings caused by bodily states (drinking when thirsty, eating when hungry, sensuous touch, orgasm, pain, nausea, physical exhaustion, and so on) but affective states more generally (including joy, cheerfulness, love, humor, pride, enjoyment, boredom, fear, anger, grief, guilt, and sadness).

C1P4 Both egoism and hedonism have an extended history, reaching back at least to the Ancient Greeks, and some form of egoism has arguably been the dominant view in the social sciences throughout much of the twentieth century and beyond (Doris et al. 2020). Moreover, egoism continues to be an influential strand in popular culture, in part through the “objectivist” writings of Ayn Rand. So does hedonism (albeit rarely under that name), with self-help books routinely advising people to “give back” to the community because doing so will make them feel good, for example. But hedonism, in particular, has undergone a resurgence recently among social psychologists, with the rise of theories of decision-making that emphasize prospection of the future together with affective responding to those simulated actions and outcomes (Gilbert & Wilson 2007, 2009; Miloyan & Suddendorf 2015). Indeed, although I will argue that hedonism is false, it will nevertheless emerge as the most defensible form of egoism.

C1P5 It is important to realize that egoism and hedonism are *modal* theses. They claim that human actions *cannot* be motivated by anything other than self-interest (in the case of egoism), or by one's own anticipated hedonic states (in hedonism's case). In fact, they claim to state laws of human psychology. Hence the force of the *cannot* in question is only as modally strong as those of any other special-science laws. These generally admit of exceptions due to breakdowns or variability at lower levels of causal organization (Pietroski & Rey 1995). Relationships among mental states, in particular, are always to some degree stochastic, reflecting random fluctuations in base-rates of neuronal firing in the underlying brain networks (Prather 2014; Bays 2015).

C1P6 Descriptions of someone as “selfish” or “egotistical,” in contrast, are merely factual statements. A selfish person is someone who is *often* or *generally* motivated by self-interest; and a selfish action is one that *is* motivated by self-interest. Likewise, a hedonist in the ordinary colloquial sense is merely someone who generally, or on the whole, pursues pleasures and enjoyments rather than other goods. Motivational hedonism, in contrast, claims that no one, ever (normally), pursues anything other than pleasure and avoidance of displeasure; and that this is true because of the underlying structure and normal functioning of human motivational systems.

C1P7 Since motivational hedonism will loom so large in the discussion that follows, it may be worth pausing here to emphasize its distinctness from hedonism colloquially understood. As already noted, motivational hedonism is a claim about the universal well-springs of human action. It claims that every action, no matter how seemingly selfless or altruistic, is about securing pleasure for oneself, or avoiding displeasure, or both. What one might call “colloquial hedonism,” in contrast, is an optional lifestyle choice, involving avoidance of work, absence of long-term goals, and attempts to maximize short-term enjoyment. A hedonist, in this sense, is someone whose life is devoted to the pursuit of bodily and social forms of pleasure. A hedonist’s life is one that is devoted to “wine, women, and song,” as the old (sexist) adage had it; or in more contemporary terms (the terms used in the 1970s and 80s, anyway), a life of “sex, drugs, and rock-and-roll”—classically represented by Bacchus, the Roman god of food, drink, and fertility. A *motivational* hedonist, in contrast, is someone who endorses a specific theory of human psychology, according to which even Mother Teresa does what she does to secure good feelings for herself (pride, perhaps, or the pleasure that comes from making a difference), or in order to avoid bad ones (such as guilt or shame).

C1P8 A couple of additional points are worth emphasizing. One is that motivational hedonism is quite distinct from hedonic *utilitarianism*. This is the view that only good and bad hedonic states matter morally, and that an overall positive balance of them should be maximized. Indeed, it might be problematic for hedonic utilitarians to endorse motivational hedonism, if one assumes that “ought” implies “can,” for motivational hedonists think that people *can’t* do other than pursue *their own* hedonic good. Of course, one could also claim that hedonic-utility is actually maximized if everyone just pursues their own hedonic maximization. That would be a fortuitous outcome: in acting as they must (motivated by the only thing that *can* motivate them, allegedly) people would also be acting rightly (provided that they do so effectively, of course). Although utilitarians have not generally endorsed such a claim, weaker versions of it are central to classical economics. If the market operates freely, and everyone acts for their own economic interests, then the economy as a whole will flourish. (This is Smith’s [1759] “invisible hand.”) However, even classical economists generally stop short of endorsing either variety of motivational egoism, allowing that people have intrinsic—non-self-interested—concerns for family and friends, for example.

C1P9 Another point to mention is that even if one were to succeed in showing that motivational egoism and hedonism are false, that would still leave intact normative analogs of those views. One could still claim that one *ought* to

4 Human Motives

care only about oneself, or *ought* to act only to secure good and avoid bad hedonic states for oneself. I can, myself, see little that is attractive in such views. But in any case, no attempt will be made to argue against them here. What a successful critique of motivational forms of egoism and hedonism *would* do, however, is underpin the possibility of non-consequentialist forms of moral evaluation of actions and agents, as we will see next. If one thinks, as most of us do, that duties should generally be performed irrespective of one's own interests or enjoyment, and that intrinsic concern for other people is required in order to be a good person, then one needs it to be the case that motivational egoism and hedonism are false. Happily, they are. Or so I think the science of affect demonstrates when properly interpreted.

C1S2 **1.2 Why It Matters**

C1P10 The theoretical stakes surrounding these issues are quite high. For if either motivational egoism or motivational hedonism is true then arguably there are no morally good actions, and nor are there any virtues of character or morally good people. For most of us think that intent is critical for each. Morally worthy actions should aim at worthy ends for their own sakes, not merely out of self-interest or to make oneself feel good. And good people are people who not only value the right things, but reliably act in order to secure those values, and not (or not just) to secure their own benefit or pleasure. All we will be left with are consequentialist notions of good action and good character (that is, actions and people that give rise to good consequences, whatever their underlying motives might be).

C1P11 Consider someone who very publicly makes a large donation to charity, but who does so only because it will increase her chances of winning an upcoming election. (Perhaps she is caught on tape, in an unguarded moment, saying exactly this.) While we would evaluate the consequences of the action as good—assuming the charity is a worthy one, it is better that the money be given than not—we would be unlikely to think any better of the agent for it. On the contrary, knowing that the gift was entirely self-interested, we are likely to think *worse* of her because of it. Likewise, consider someone who repays a large loan to an acquaintance, not because justice requires it, but merely to avoid the inconvenience of having to contest it—if he could have gotten away without having to repay, then he would have done so. (Again, perhaps he is caught on tape saying just this.) We would surely think worse of him as a result.

C1P12 Admittedly, the heavy reliance that people in developed cultures (especially in the West) place on agents' intentions when morally evaluating them appears not to be a cultural universal. Although most small-scale societies, too, place some reliance on the intention behind an action in at least some contexts, this seems not to be true of all; and the differences in moral evaluation between intentional and accidental harms tend to be much smaller than they are in large-scale societies (Barrett et al. 2016).

C1P13 It does appear, however, that modulating one's social evaluations of others by their intentions might be the default setting in human development. This is because infants as young as eight months prefer an agent who tries, but fails, to help another over one who tries, but fails, to hinder the goal of another (Hamlin 2013; Kanakogi et al. 2017). Note that the outcome in these cases is the same. Likewise, eight-month-old infants prefer an agent who tries to help another but fails (because acting on a false belief) over an agent who accidentally provides such help (Woo & Spelke 2023); and infants as young as ten months prefer an agent who intentionally, versus accidentally, helps another; and conversely, they also prefer an agent who accidentally, versus intentionally, harms another (Woo et al. 2017).¹ It may be that these are initial innate tendencies that can be weakened by later cultural input. Whether or not that is true, it remains the case that the vast majority of the people in the world now think that moral evaluation depends critically on the agent's intention when acting.

C1P14 Accepting either egoism or hedonism would come with steep theoretical costs for most of us, then. But in practice those costs could be limited (for motivational hedonists, at least). This is because arguably the underlying function of evaluating others in terms of intent is to predict future patterns of behavior, enabling us to select reliable friends and cooperators while avoiding those who are likely to hurt or betray us. And this function could just as well be served by paying attention to people's proximal, or instrumental, intentions when acting. Even if everyone aims ultimately to secure their own pleasure and avoid their own displeasure, it makes a good deal of difference in practice whether what gives someone pleasure is helping others versus becoming rich, or whether what makes someone feel bad is breaking a promise rather than mere loss of reputation. Motivational hedonists, especially, can thus retain many of the practical benefits of evaluating actions

¹ Admittedly, these findings are only known to hold true among infants in the West. Such studies have not yet been attempted in small-scale societies. However, infant false-belief studies *have* been validated cross-culturally (Barrett et al. 2013).

6 Human Motives

and agents in terms of intentions by focusing on proximal rather than ultimate intentions.²

C1P15 That motivational egoism and hedonism come burdened with theoretical costs is not in itself a strike against those views. For each makes an empirical claim about the psychological roots of human motivation—one that may now seem scientifically grounded, as we will see. So our ordinary intuitions should carry no weight in this debate. Indeed, whenever folk-beliefs come into conflict with established science, it is the former that should give way. Scientific findings do need to be interpreted, however, and scientists themselves sometimes over-interpret or misinterpret their own data.

C1P16 Moreover, the project of this book is to understand what really, truly, motivates people, not to describe and regularize our ordinary concepts and beliefs (which is what most philosophical treatments are concerned to do). So, let the chips fall where they may. If the best scientific account turns out to support hedonism, then we will have to live with the consequences and adjust our beliefs accordingly. Happily, however, I think the best interpretation of the science supports motivational pluralism. Or so I will argue over the course of this book.

C1P17 Egoism and hedonism don't just provide challenges to our ordinary moral beliefs; they are also socially damaging. This is because theories of human psychology (especially motivation-relevant psychology) are apt to be self-fulfilling. For example, believing that intelligence is malleable rather than a fixed quantity can cause people to do better in school; and interventions that induce beliefs in malleability can improve performance (Blackwell et al. 2007). Moreover, people who believe willpower to be a depleting resource—rather than unlimited—can perform worse in a series of attentionally demanding effortful tasks; and interventions to induce unlimited beliefs can likewise improve performance (Job et al. 2015). So we can be confident that people who believe that all actions are ultimately selfish will be apt to behave in a more selfish manner; and conversely, that work persuading people of the falsity of motivational egoism and hedonism will increase the likelihood of prosocial action. Although to the best of my knowledge this specific prediction has not been directly tested, a closely related one has. Economists believe that people are mostly motivated by self-interest, and Frank et al. (1993) found that taking a class in economics causes people to

² Indeed, as we will see in Chapter 2, it is possible for people to have other-regarding values and cares while at the same time all of their decisions are about maximizing their own hedonic utility, as motivational hedonists claim. So it would be possible to understand good *agents* to be those who possess the right *values*, even though all actions are nevertheless hedonically motivated and self-interested. This would pull apart the evaluation of agents from the evaluation of actions, of course, which would itself be a theoretical cost for most of us.

become *more* self-interested. Convincing people of the falsity of motivational egoism and hedonism, then, should lead them to become *less* self-interested.

C1S3 1.3 The Way Ahead

C1P18 Egoism and hedonism are perennially tempting, and have proven remarkably difficult to refute. Those who have attempted to critique them have mostly focused on motivational hedonism, regarding this as the most plausible (and most widely endorsed) form of egoism (Butler 1726; Broad 1952; Nagel 1970; Feinberg 1984; Batson 1991, 2019; Sober & Wilson 1998). Moreover, the most recent iterations of motivational egoism to emerge out of scientific psychology have likewise taken a hedonic form. Indeed, we will see in Chapter 2 that non-hedonic forms of egoism are scientifically deeply problematic. So I, too, will focus on the case for and against motivational hedonism in this book.

C1P19 Chapter 2 will lay out the basics of the recent science of affect, and Chapter 3 will show how it can be (and has been) taken to support motivational hedonism. These two chapters will also, at the same time, demonstrate the scientific basis for at least a weak form of Humeanism about motivation, according to which all of our motives are grounded in affect (“the passions”), thereby resembling views that Blackburn (1999) has defended on other grounds. Chapter 4 will then show that the resulting form of hedonism is immune to the critiques that have been mounted by philosophers in the past, as well as being immune to the experimental findings in psychology that have been intended to refute hedonism.

C1P20 Chapter 5 then begins consideration of the nature of the positive and negative valences (pleasure and displeasure) that are common to all affective states, examining two types of theory. It will contrast an intrinsic-feeling account (which would support motivational hedonism) with an account that construes valence as a nonconceptual representation of value (which entails pluralism). According to this latter view, when we select an option because of anticipatory joy or guilt, we are choosing it because that option seems good to us (in a quasi-perceptual manner), or seems bad, not because it will make us *feel good* or *feel bad*. Chapter 6 then contrasts this representational view with an imperativist account (which sees valence as an imperative-like urge to do something or to forebear from doing something). In some versions this, too, entails motivational hedonism. The value-representing view of valence will emerge as the overall winner from these comparisons.

8 Human Motives

C1P21 Chapter 7 will then examine a test case for the value-representing account of valence, specifically the case of physical pain, about which much has been written by philosophers lately. Since it is natural to think that pain just feels intrinsically bad, rather than represents anything as bad, it might seem to support an intrinsic-feeling account of negative valence in particular. (Similar points apply to simple physical pleasures, like the taste of a ripe strawberry.) Indeed, the case of pain is taken by many to present a challenge for representational accounts of conscious experience generally. Chapter 7 will show that these challenges can be met.

C1P22 Finally, Chapter 8 will examine moral motivation, and the psychological roots of prosociality, through the lens of the theory of valence-based decision-making outlined in previous chapters. The upshot is that human beings are indeed equipped to want to do the right thing for its own sake (some of the time), and to help (some) others for their own sakes. However, the main focus of the chapter will be to argue that some of these motivations are innate or innately channeled. The final two sections then conclude by pulling together the main ideas and arguments of the book.

C1P23 I should stress that my arguments throughout will be grounded in science, debating how it should best be interpreted. There will be no reliance on the kinds of thought-experiments and appeals to intuition that are the bread-and-butter of most philosophical writing. When it comes to understanding the mind—and understanding the sources of human motivation, in particular—such approaches are worse than useless. They provide an illusion of knowledge and understanding, whereas in fact they just reflect folk theories and common-sense conceptions. It is as if philosophers were to continue to advance claims about the physical world by studying Aristotle (whose views, arguably, capture much of our intuitive folk-physics; Clement 1982; McCloskey 1983; Keleman et al. 2013).

C1P24 This book will also devote much more space to animal models of motivation than is usual in philosophy. Indeed, many philosophers of mind in the latter half of the last century were suspicious of—not to say contemptuous of—all talk of reward, punishment, and reward-based learning.³ This was grounded in their rejection of behaviorism, where such language originated, and in their embrace of the cognitive revolution. But the science of evaluative learning has moved on since then. While the language of “reward” and “punishment” is still used, these terms are embedded in sophisticated computational models of conditioning and reward-based learning. In fact, such

³ One of the few philosophers to take the science of affect seriously is Morillo (1995), who is led to endorse motivational hedonism as a result.

models are routinely interpreted realistically, as well, entailing that there are explicit computational and representational processes in the minds of the animals doing the learning. Moreover, while there are, of course, important differences between humans and other animals—resulting especially from gene-culture co-evolution (Henrich 2017)—it is now widely accepted in the affective sciences that the underlying foundations are the same. We ignore them at our peril. As a result, the general approach taken in this book (as in most science) is “bottom-up”—starting with simple mechanisms before building toward an account of more complex cases.

C1P25 The book will thus devote correspondingly less space to considering philosophical treatments of motivation and decision-making. This is because, as noted above, what most such treatments are really treatments of are common-sense theories of motivation, perhaps combined with what people take to be the deliverances of introspection. While it might be fine to take common sense as a starting point if there were no well-developed science in the area, there is now an immense amount of well-established scientific work on the topic. One might wonder, then, what business a philosopher has in addressing these issues at all. Two points can be made in reply. One is that scientific results still need to be interpreted. The other is that scientific work can sometimes be overly narrow. There is a good deal of mileage to be gained from stepping back and looking at the bigger scientific picture in the way that I do here, integrating results across a range of related subdisciplines.

C1P26 Overall, this book has two main goals. One is to make findings from the new science of affect accessible to a wider academic audience. My hope, in particular, is that many long-standing philosophical debates can be both rejuvenated and illuminated if those engaging in them have more knowledge of the science. The other main goal is to refute motivational hedonism. But in doing so, I take hedonism much more seriously than most other critics have done. In fact, I try to develop the strongest possible form of motivational hedonism *before* showing that it fails. I believe Wittgenstein once said something to the effect that good military generals attack their opponents on their weakest sides, whereas good philosophers attack their opponents on their strongest. For our goal is truth and understanding, not point-scoring or quick victory. Since the resulting kind of argumentative good practice is, in effect, the converse of the well-known straw-man fallacy, it is sometimes referred to using a related name. It can be called “a steel-man argument,” “steel-man validity,” or, perhaps better, “the steel-man strategy.” So my second major goal can be described as an attempt to refute motivational hedonism once-and-for-all through use of the steel-man strategy.