

1-1-1980

## Pleasure and intrinsic goodness.

Earl Brink Conee  
*University of Massachusetts Amherst*

Follow this and additional works at: [https://scholarworks.umass.edu/dissertations\\_1](https://scholarworks.umass.edu/dissertations_1)

---

### Recommended Citation

Conee, Earl Brink, "Pleasure and intrinsic goodness." (1980). *Doctoral Dissertations 1896 - February 2014*. 1829.

<https://doi.org/10.7275/svtn-sx30> [https://scholarworks.umass.edu/dissertations\\_1/1829](https://scholarworks.umass.edu/dissertations_1/1829)

This Open Access Dissertation is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Doctoral Dissertations 1896 - February 2014 by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact [scholarworks@library.umass.edu](mailto:scholarworks@library.umass.edu).





312066 0298 4326 0

**FIVE COLLEGE  
DEPOSITORY**



PLEASURE AND INTRINSIC GOODNESS

A Dissertation Presented

By

EARL BRINK CONEE

Submitted to the Graduate School of the  
University of Massachusetts in partial fulfillment  
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 1980

Philosophy

PLEASURE AND INTRINSIC GOODNESS

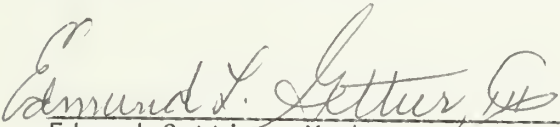
A Dissertation Presented

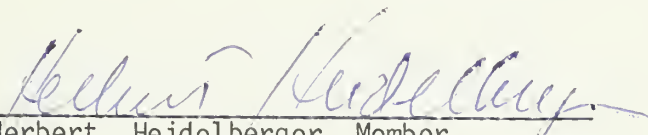
By

EARL BRINK CONEE


Approved as to style and content by:

  
Fred Feldman, Chairperson of Committee

  
Edmund Gettier, Member

  
Herbert Heidelberger, Member

  
Emmon Bach, Member

  
Edmund Gettier, Department Head  
Philosophy

## ACKNOWLEDGMENT

Readers familiar with the philosophical work of Fred Feldman, Edmund Gettier, and Herbert Heidelberger will see that I am deeply indebted to each of them, and that I share their debt to Roderick Chisholm.

## ABSTRACT

### Pleasure and Intrinsic Goodness

September 1980

Earl Conee, B.A., Trinity College  
PH.D., University of Massachusetts

Directed by: Professor Fred Feldman

The three major philosophical connections between concepts are analysis, epistemic mark, and general theoretical link. Each has been thought to relate pleasure to intrinsic goodness. The Introduction consists in a brief study of the nature of these connections. The body of the thesis examines whether pleasure and intrinsic goodness are so related.

In Chapter III the question is whether the concept of pleasure enters into an analysis of the concept of being intrinsically good. This topic is approached through writings by Franz Brentano. It is argued that no such analysis is possible.

Chapter III is concerned with whether pleasure can help us to identify the bearers of intrinsic goodness. Again, some of Brentano's work is central to the topic, as is work by Charles Baylis. It is argued that pleasure is of no special help in discovering the intrinsic goods.

Chapter IV begins with an attempt to give a clear and complete formulation of hedonism--the theory according to which only pleasure is intrinsically good. The formulation builds upon the efforts of Warren

Quinn and Edward Oldfield. Then an argument against hedonism by Brentano is critically assessed. C. D. Broad's objection to the effect that the pleasure of malice is not intrinsically good is then evaluated, as is G. E. Moore's objection to the effect that beauty is also intrinsically good. It is contended that hedonism does not succumb to any of these objections.

The Appendix considers whether instrumental value of any sort depends upon intrinsic goodness. It is argued that most familiar sorts do not so depend, but one can be shown to do so by a kind of First Cause argument.

## TABLE OF CONTENTS

ABSTRACT . . . . .	iv
Chapter	
I    INTRODUCTION . . . . .	1
A. Some Presuppositions . . . . .	1
B. The Goals of Philosophical Analysis . . . . .	3
C. The Form of Epistemic Principles . . . . .	12
D. Philosophical Theories . . . . .	15
E. The Contents of the Appendix . . . . .	19
II   PLEASURE AND THE ANALYSIS OF INTRINSIC GOODNESS . . . . .	23
A. Initial Readings of Brentano's Analysis . . . . .	24
B. Revised Interpretations of Brentano's Analysis . . . . .	29
C. Analysis by Use of Brentano's Criterion . . . . .	33
D. The Requirement Interpretation of Correctness . . . . .	43
III  PLEASURE AND THE DISCOVERY OF INTRINSIC GOODNESS . . . . .	50
A. Tests for Epistemic Principles . . . . .	50
B. Brentano's Criterion . . . . .	51
C. Baylis's Criterion . . . . .	56
D. Eliminating Remaining Apparent Disanalogies . . . . .	59
IV   PLEASURE AND THE POSSESSION OF INTRINSIC GOODNESS . . . . .	62
A. First Attempts to Formulate Hedonism . . . . .	62
B. The Quinn-Oldfield Approach . . . . .	67
C. Hedonism Formulated . . . . .	77
D. Brentano's Argument Against Hedonism . . . . .	85
E. Broad's Malice Objection . . . . .	92
F. Moore's Beauty Objection . . . . .	97
APPENDIX: INSTRUMENTAL VALUE WITHOUT INTRINSIC VALUE? . . . . .	103
A. The Argument From Definition . . . . .	103
B. Beardsley, Taylor and Harman on First Cause Arguments . . . . .	106
C. A First Cause Argument Refuted . . . . .	115
D. A First Cause Argument Upheld . . . . .	117
BIBLIOGRAPHY . . . . .	125



# CHAPTER I

## INTRODUCTION

Broadly speaking, this dissertation is a search for philosophically significant ties between pleasure and intrinsic goodness. We consider whether pleasure enters into an analysis of intrinsic goodness, whether pleasure can be used in a criterion of intrinsic goodness, and whether pleasure and intrinsic goodness are co-extensive. The purpose of this introduction is to explain these topics. First, though, we announce some presuppositions.

### A. Some Presuppositions

Much that has been controverted is taken for granted here. I might as well set out the main assumptions right here at the outset. I assume that there is such a thing as intrinsic goodness (called by some philosophers "inherent goodness," and by others "intrinsic desirability"). I assume that there are intrinsic goods. I do not assume any particular value theory, but I do take this much for granted: Evaluative hedonism--very roughly the view that only pleasure is intrinsically good and only pain is intrinsically bad--is not analytically or criterially false. That is, I assume that it cannot be deduced just from an analysis or an epistemic criterion of intrinsic goodness and uncontroversial premises that evaluative hedonism is false. This gives us something we must have to get anywhere: a test for proposed analyses and criteria. And it does so without presupposing the truth, or even the

possible truth, of any theory of value. Given how basic are the disagreements about the existence, nature and application of intrinsic goodness, it is good to be able to travel this lightly. It is hard to see how we might make any important progress in this investigation using only lighter normative baggage.

I assume that it does not affect the normative questions at issue to take state of affairs or events to be the bearers of intrinsic goodness. I further assume that it does not hurt to take pleasure to be a relation--the taking of pleasure--that can hold between a person and a state of affairs. I attempt to answer relevant objections to this as they arise. In each case, I am not maintaining the truth of the ontological claim. I merely take it that we do proposals about pleasure and intrinsic goodness no special harm by making the topic definite in this way. (The reader is encouraged to verify as we go along that the ontology makes no difference where it is not defended.)

Finally, it is assumed here that pleasure and pain come in degrees, and numerical comparisons within and between the two are possible.<sup>1</sup> Outside of philosophy this is often assumed, as when someone says, "I was pleased when Smith left, but I was twice as pleased when Jones left," or "The pleasure of dining with them balanced off the pain of waiting for them," or "Disease has caused a greater quantity of human misery than famine." I know of no plausible grounds for doubting that such quantitative relationships exist.<sup>2</sup> All doubts that make sense here seem to reduce to doubts about the feasibility of finding out the relevant amounts. I do not assume that that can be done.

Of course the above list does not include everything of philosophical substance that is assumed below. But it summarizes the really big presuppositions--the ones that some philosophers would find shocking or appalling. They are fairly forewarned.

### B. The Goals of Philosophical Analysis

We now proceed to examine in some detail the sorts of connections between pleasure and intrinsic goodness that are to be studied here. First, there is the question whether the concept of pleasure enters into an analysis of intrinsic goodness. What would that entry amount to? Without trouble, we can work any concept into a necessary equivalence with intrinsic goodness:

Pl.1  $p$  is intrinsically good iff  $p$  is intrinsically good and either  $p$  is intrinsically good or \_\_\_\_,

where the blank can be filled in with an expression for any concept, and a truth results. At the other extreme, it is reasonably clear that the concept of pleasure is not part of what is meant by "intrinsically good" in English. The following example can make that plausible: A child is raised in an anti-hedonistic sub-culture. The child is kept in complete ignorance on the topic of pleasure, never having any himself and never getting any evidence about what it is like. The child is taught the concept of intrinsic goodness, though. The examples used are cases of things taken by the anti-hedonists to be good on their own: hard work and perseverance in the face of misfortune. It is insisted that these be sought for their own sake.<sup>3</sup>

Thus we should not expect to find a phrase synonymous with "intrinsically good" in which the concept of pleasure is expressed. And we can find every number of uninteresting necessary equivalences involving pleasure and intrinsic goodness. But there is something in between of considerable importance to philosophers: analysis or philosophical definition. What does it take to have one of those? This is not the place for a full study of analysis. But it will be worth our while to consider a recent proposal by Roderick Chisholm.<sup>4</sup> To understand the proposed analysis of analysis, first we need:

- D1.1 p implies G = df. p is necessarily such that, if it obtains, then something has G.
- D1.2 p involves q = df. p is necessarily such that whoever entertains it entertains q.
- D1.3 p involves G = df. p involves a q which is necessarily such that it obtains iff something has G, and G is not an all or nothing property (i.e., possibly, something has G and something does not).

In terms of these defined expressions we are given:

- D1.4 p analyzes q = df. (i) p is logically equivalent to q; and (ii) both p and q imply a property that p involves but q does not involve; and (iii) p involves every state of affairs that q involves.

The example:

- e1 (a) Something is a male and a sibling, analyzes (b) Something is a brother,



can be used to see how D1.4 works. Clearly (a) is necessarily equivalent to (b). Is there a property that both require to be instantiated, while only (a) requires entertaining something logically equivalent to its instantiation? Yes, the property of being a sibling. There must be a sibling when (b) is true. But we must think of something being a sibling (or an equivalent) when thinking of (a), and not (b). That, at any rate, is how the example is supposed to work.

However well we judge D1.4 to work in cases like e1, other cases make it look much too restrictive. We do not want to set our goal of an analytic connection between the concepts of pleasure and intrinsic goodness so high that it is not met by:

- e2 (c) Something is a closed, 3-dimensional figure with every side being a square plane segment, analyses (d) Something is a cube.

E2 is, I think, as lucid, succinct, and illuminating an explanation of a geometrical concept as we can find. Yet (c) does not involve everything (d) involves, for it does not involve (d). Many have to figure out the fact that (c) pertains to cubes. So that is not staring them in the mind's eye when first they entertain (c). Furthermore, it is manifest that almost none of the concept analyses, e.g. of knowledge, which are actually proposed of late include analysans that involve their analysanda. This lack of involvement is not in fact counted as refutation of these proposals. Of course this might be a mistake. But the widespread sustained interest in these proposals suggests that something weaker than a D1.4-analysis would be significant. Surely e2 for instance

states a connection that is more interesting than non-obvious necessary equivalence. Compare e2 with:

- e3 (f) There is no negation-complete formal theory of arithmetic,  
analyses (g) Something is yellow or nothing is yellow.

To explain what some seek from analyses, construction metaphors are often employed. The analysans, it is said, should tell us what are the "building blocks" of the analysandum, and it should say how the blocks are "put together." But not everything that appears to be a successful analysis conforms to that metaphor. Consider:

- e4 (h) Something is a brother or a sister, analyzes (i) Something  
is a sibling.

The construction metaphor would have it that e1 tells us that being a sibling is a building block of being a brother. But e4 seems as satisfactory an analysis as e1, and the metaphor would have it that e4 tells us that being a brother is a building block of being a sibling (albeit one that is "disjunctively joined" to another to make up siblinghood, while siblinghood is "conjunctively joined" in the composition of brotherhood). No building can be made out of blocks related as e1 and e4 would tell us brotherhood and siblinghood are. If one part is made out of two others, neither of the two can be made out of it. Perhaps we will come to be so well justified in accepting some theory of property composition that we become entitled to reject the "construction" indicated by e1 or e4. But that would not solve the problem. There is clearly something which would be of philosophical interest (if these

were philosophically interesting concepts) that e1 and e4 do equally well.

There is another sort of problem for D1.4. Herbert Heidelberger has pointed out<sup>5</sup> that it counts the following as a case of analysis:

e5 (j) Something is a brother and a sibling, analyzes (k) Something is a brother.

But e5 and any other example where the expression of the analysans contains the expression of the analysandum is not a good analysis. At least such circularity is a fatal weakness for some purposes to be served by analysis. e5 and its ilk cannot be used to informatively identify which concept a phrase expresses, or to solve any puzzle about its nature.<sup>6</sup> Yet some aims of analysis are achieved only when the analysans at least involves the analysandum. To show by analysis that "the average tomato" only appears to refer, the analyzing sentences must convey the whole meanings of the ones containing putative reference to an average tomato. Otherwise we have not been shown that we can, without such reference, say everything that can be said with it. Furthermore, D1.4 has a feature that it is important for a notion of analysis to have--it blocks the paradox of analysis.<sup>7</sup> Briefly stated the paradox is this: Some have held that analysans and analysandum are identical. It is therefore found paradoxical that analyses are more informative than what is expressed by an identity statement where the same referring expression appears twice. On D1.4's account the analysans is a different proposition than the analysandum--one which differs by

asserting something which is just implied by the latter. That allows D1.4-analyses to be sufficiently informative.

I think it is a mistake to look for some one kind of equivalence--"analysis"--that accomplishes all these things. As I see it, there are many kinds of valuable necessary equivalences. Different kinds fulfill different philosophical purposes. We can identify these goals and say what it takes to achieve them. The expression "analysis" has been used by philosophers to cover a variety of philosophical objectives. G. E. Moore, the philosopher who did the most to give it currency, used the term in several ways. There is nothing at all obvious that his uses of the term share.<sup>8</sup> I see nothing of philosophical importance that can only be settled when we determine which equivalences it is best to call "analyses."

The sort of ontological reduction already mentioned requires synonymy. To rid ourselves of the need to suppose that a term with clear and definite meaning refers, we must see that its linguistic work can be done without any phrase that gives evidence of denoting what the former would denote if it did. So in the case of this philosophical goal, what is expressed by the reducing and the reduced sentences must enter into the strongest of equivalence relations--identity. Such reduction is best expressed by sentences of the form: "A is a potentially ontologically misleading way to say B." This will be philosophically helpful when it enables us to see that the truth of A only requires us to acknowledge the ontology of B, rather than the additional thing(s) that A suggests.



Another implication of identity occurs when mere synonymy is expressed, e.g. with expressions of the form: "A means the same as B." I can see no philosophical aim that is always achieved by truths of that form, even when A and B are different expressions for philosophically significant propositions. Many such sentences are useful in conveying concepts. That is no philosophical enterprise (though it is part of some). There are two sorts of philosophical work that can be accomplished by sentences of that form. But neither requires identity.

One such task is the elucidation of philosophical concepts. It is quite difficult to say in general what that is. This is partly because different theories of properties and propositions impose quite different constraints. Suppose a theory that appeals to the construction metaphor is right. Then, assuming that, e.g. brotherhood is a complex "structure," it can be fully elucidated by saying what its simple "parts" are and how they are "put together." Perhaps e1 makes these elements and their relationship obvious. And perhaps (a) in e1 means the same as (b). Then "'Something is a brother,' means the same as 'Something is a male and a sibling'" would be a tolerably clear way to impart the details of what goes into the concept of being a brother. But even better would be an expression that is explicit about the construction, such as:

- e6 (1) Something falling under the conjunction of the concepts maleness and siblinghood is identical with (m) something being a brother.

If e6 is true, it is informative. It is also clear that (l) and (m) are not synonymous, so this informativeness should not seem paradoxical.

One way the sentence e1 could be informative while (a) means the same as (b) is by making evident by its syntax the "construction" of brotherhood, i.e., what e6 tells us explicitly. But if all that is assumed for the moment, then how is e4 informative? It cannot be by displaying the composition of brotherhood.

Whether some concepts are built out of others or not, plainly it illuminates a concept to be told inexplicit implications of it. That is, if we provide an equivalent that involves a mere implication of the concept's application, we show something of what goes along with that application. This is a service that every D1.4-analysis does provide. And regardless of whether "constructivists" are correct, we are free to say that that is something that e4 accomplishes (and e1 too, if Chisholm is right that (b) does not involve (a)). It is worth noting that circularity does not ruin a proposal for this purpose.<sup>9</sup> Note also that this task does not require that one equivalent have all the involvements of the other. That condition of D1.4 (perhaps intended to provide for reduction by "analysis") should not be imposed on all explanations of concepts. I have no full account of concept explanation. One good question is whether any involvementsal overlap is needed. I neither see how to argue that it is, nor see a clear case that lacks it. In any event, this work of bringing out implications is fundamental to one sort of concept explanation.

The other philosophical task that an assertion of a meaning identity may accomplish is what is sometimes called "clarification," sometimes "explication" of concepts. That aim usually arises like this: An expression is said to express a concept having certain remarkable properties or covering just a certain range of cases. Some philosophers doubt whether any concept has those features, or wonder which qualities are present in just those cases. By expressing in clear terms a concept which undoubtedly has those features or applies in just the right cases, the "clarification" or "explication" is accomplished. This might be done by offering a synonymous expression that uses sufficiently clear terminology to make the relevant qualities or extension manifest. But providing another concept (usually, an equivalent is required) that has those features would do as well. I take that to be what is actually being attempted in the name of "analyzing knowledge," for example. Also, even a complicated equivalent often can supply a better understanding of the initial concept. For the target concept is often the psychologically easiest or most prominent way to conceive of the extension in question. E2 gives an excellent example of how this sort of thing works (though there is nothing philosophically doubtful or fascinating about cubicality in the first place.) (c) employs basic geometrical concepts in straightforward relationships that result in a condition for which the intuitively simplest concept is that of being a cube. Here it is worth pointing out that circularity does spoil a proposal for this endeavor. E4, for instance, cannot perform this task. Questions about whether there is a concept having certain properties or curiosity about

what distinguishes a certain extension is not legitimately met by use of the very expression or presupposition of the very conceptual means in question to the extension.

So much by way of brief indication of what might be sought from "analysis." Why should we consider whether pleasure and intrinsic goodness are related in some such way? Well, apart from the intrinsic interest of the question, we can point to these things: Brentano seems to have thought that there is such an equivalence, perhaps an identity, involving the two concepts.<sup>10</sup> On the basis of it he proposed an epistemic criterion of intrinsic goodness<sup>11</sup> and an argument for the proposition that if pleasure is intrinsically good, then so is something else.<sup>12</sup> Further, some philosophers have had doubts about the nature or application of intrinsic goodness.<sup>13</sup> So what we have called "clarification" by appeal to the concept of pleasure would be worthwhile.

### C. The Form of Epistemic Principles

Our next objective is an epistemic criterion of intrinsic goodness by use of pleasure. What does it take to have one of those? It is unilluminating and misleading to say that it takes "a way of finding out what is intrinsically good." That is too broad and too narrow. A "way" that consists in consulting someone who knows does not count, and a proposal that merely improves our evidence about what is intrinsically good does. The general situation is this: There is an ordering of categories of propositions according to the quality of the evidence we can have for propositions in the category. Philosophers dispute what stands



where, but it is widely acknowledged that the contents of sense experience, physical object statements, and cosmological hypotheses are at increasingly distant locations in this ordering. This introduction is not the place to defend, or even to propose, a ranking, but we can say this much: Any true proposal maintaining that evidence from one level in the ordering lends favorable epistemic status to a proposition further out is a candidate for being a significant criterion. It is significant if the subject matter is of interest. In fact, it suffices for the proposal to have some interest if it so relates propositions that have been reasonably thought to be at such different levels.

There are major differences in strength among these proposals. For our purposes, the best we could find would be a criterion that identifies some maximally good evidence that depends somehow upon pleasure for the proposition that a certain state is intrinsically good. That might seem to amount to any true principle of the form:

C1.1 Necessarily, if S considers p and S has (such-and-such a proposition involving pleasure) as evidence, then nothing is more reasonable for S than that p is intrinsically good.

But C1.1 is just a strict conditional. Compare:

C1.2 Necessarily, if Jones has Smith's testimony that there is life on Mars as evidence, then nothing is more reasonable for Jones than the proposition that he exists.

The pleasure in a principle of the form of C1.1 may play no greater role than the testimony in C1.2. We want the evidence to be what supports the intrinsic goodness claim, not a mere sufficient condition for its

having support. Can we say what it is for one proposition to lend credence to another when the one is had as evidence? I believe that we can. In "Propositional Justification,"<sup>14</sup> I offer a definition of the weak propositional support relation of "tending to confirm." I make use of the preliminary notion:

S has minimal evidence for h = df (i) h is evident for S, and (ii) there is a p such that p is evident for S, p entails each thing evident for S, and this is not possible;: there is a q such that p entails but is not entailed by q, q entails each thing evident for S, and q entails h.

On the basis of that, I say:

e tends to confirm h = df necessarily, for any S, if S has minimal evidence for e, then believing h is more reasonable for S than believing not h.

By use of the same devices, we can isolate stronger evidential relationships:

D1.5 e justifies h = df. Necessarily, if S has minimal evidence for e, the h is evident for S.

D1.6 e proves h = df. Necessarily, if S has minimal evidence for e, then h is known by S.

D1.7 e makes h certain = df. Necessarily, if S has minimal evidence for e, then nothing is more reasonable for S than h.

Unfortunately, propositional justification relations do not immediately yield means to belief justification. Consider an e which does

in fact justify an  $h$ , and is of a more secure epistemic rank than  $h$ . Those assumptions do not imply:

C1.3 Necessarily, if  $S$  has  $e$  as evidence, then  $h$  is evident for  $S$ . Since  $S$  may have  $e$  and some countervailing evidence about  $h$ , the mere possession of  $e$  does not insure that  $h$  is evident. It is one thing for a proposition to favor another's truth to some degree, and quite another for the one to make a person under certain specific conditions reasonable or justified in believing the other. Obviously, the latter also depends upon what other evidence the person then has.

I do not know how to give a precise explanation of the relation of making evident.<sup>15</sup> Once we see this difference, though, it becomes clear that the primary and purely philosophical question for us is whether some proposition involving in a crucial way the concept of pleasure lends some degree of evidential support to the proposition that a state is intrinsically good. Finding an actually employable criterion of the sort will be possible only if there is such a support relation that underlies it and if we can actually get into the situation described by the antecedent. (Note that, if a pleasure proposition can make the intrinsic goodness of a state evident to  $S$  under certain conditions, then  $S$ 's total evidence in those conditions  $D_j$ -justifies that the state is intrinsically good.) So this is a partly empirical matter, and one that implies that a suitable propositional justification relation holds. Thus it is not unreasonable to concentrate on the former kind of question, at least until we discover a successful proposal.

Concerning the philosophical significance of a true proposal of this sort, little need be said. The question whether, and on what basis, we might know or even reasonably believe a thing to have value is among the best examples of a matter of philosophical interest. And there is ample historical justification to search for evidence of intrinsic goodness that relies upon pleasure in particular. Brentano certainly turned to it as a source of knowledge of intrinsic goodness,<sup>16</sup> and Meinong probably did.<sup>17</sup> We should take that as good evidence that there is something of philosophical value to be learned by conducting such a search.

#### D. Philosophical Theories

Finally, we will look into hedonism--the theory that singles out pleasurable experiences as the bearers of intrinsic goodness. This is called a "theory" of intrinsic goodness. What is required for such a theory to succeed? There is a genuine problem here. Suppose we have a D1.4-analysis of intrinsic goodness. Then we can say that the bearers of intrinsic goodness are the states that satisfy its analysans. Why is not that a "theory" of intrinsic goodness? It is an illuminating way to identify what is intrinsically good. And we should not require that such a "theory" give us a characterization of the intrinsic goods that is epistemically helpful. Theories of value qualities, such as Mill's theory of moral rightness, rarely provide any epistemic help. That is what we demand of a criterion (though of course we will take it where we get it). What, then, is peculiar to a "theory"?



In one respect the standard seems to be more lenient than those for what was discussed under the title "analysis" above. For it seems plain that a theory's characterization of the bearers of a property need not share any involvement with that property. As mentioned above, such sharing seems requisite for any "analytic" task. But of course this does not show why, e.g. a D1.4-analysis fails to be a "theory." We must not insist that there be no sharing of involvements. That would unreasonably prejudge, e.g., the question whether the concept of pleasure enters into both a D1.4-analysis and a theory of intrinsic goodness, for example.

The prominent historical philosopher to have attempted to say what we seek from a "theory" of a value property is G. E. Moore.<sup>18</sup> Speaking of a theory of rightness, he said that we are looking for "the reason why an act is right," a property such that acts are right "because" they have it.<sup>19</sup>

Unfortunately, that account is unsatisfactory. Strictly, clear cases of reasons are all considerations by someone for doing something. Unless we accept cosmological views attributed to Descartes according to which someone--God--might really have had a reason for making necessary truths true, this sense of "reason" seems inapplicable here. For it is uncontroversial that intrinsic goods are necessarily so. There is a broader sense of reason that can be very vaguely indicated as: thing in virtue of which such-and-such. But in that sense, the only candidate for a "reason" why intrinsic goods are such would seem to be the property of intrinsic goodness itself. "Because" has the first sense

attributed to "reason." It also expresses causation. But there is no cause for the bearing of an essential property. Appeal to causation here appears to be entirely unhelpful.

Moore asserted a distinction between "natural" and "non-natural" properties. Many philosophers, Ernest Sosa being the most recent,<sup>20</sup> have appealed to a distinction between "evaluative" and "non-evaluative" properties. Equivalences where one side can be taken to be about one of the former kinds of properties while the other side can be taken to be about one of the latter seem to have special philosophical interest. Most regrettably, though, no tolerably clear account of either distinction is available.<sup>21</sup> I have none to offer. Without that, we cannot sensibly investigate the basis for such interest. (Also, this would at best show us what is interesting about certain "theories" of value properties. There is no reason to expect that it would lead to a general account of what is distinctive about philosophical "theories.")

I must leave it an open question why certain equivalences involving concepts of interest to philosophers are deemed "theories" of those concepts.

Our philosophical interest in proposals such as hedonism for identifying the bearers of intrinsic goodness needs no excuse. Because of their paucity of shared involvements with intrinsic goodness, they are informative, if true. What we lack is not reason to be interested, but a general way to mark them off from "analytic" connections on the one hand and equivalences like e3 where neither side even seems to be about the other, much less one being a "theory" of the other.

Theories of intrinsic goodness like hedonism and its rivals seem not to be susceptible to clearcut proof or refutation. How should we expect them to be supported, then? We can reasonably demand these two things: First, a clear statement of the theory. As an example of what needs doing, consider the formulation of hedonism: "All and only states of pleasure are intrinsically good." What does that tell us about:

e6 Smith being pleased to know that Jones is pained.

E6 would appear to be a "state of pleasure." And yet intuitively the hedonist should not count it as good.<sup>22</sup> The second support we can ask of hedonism is a clear account of the replies available on behalf of the theory to objections that have been thought to be telling.<sup>23</sup>

#### E. The Contents of the Appendix

An appendix is appended. This gives the thesis an organic unity. It is not central to our inquiry, since it does not concern any putative connection between pleasure and intrinsic goodness. It is about intrinsic goodness and its relation to instrumental value. The question is whether an event being of instrumental value of some sort implies the occurrence of an intrinsic good. My approach is the maximally straightforward one: I attempt to identify clearly the concepts of instrumental value that are in common use, and those that have been topics in discussions of this question. This done, the concepts virtually speak for themselves as to whether there is such an implication.

## Notes to Chapter I

<sup>1</sup>More precisely, it is not questioned here that there is a magnitude--intensity--that pleasures and pains have in common. It is assumed that we can stipulate an arbitrary pleasure to have one unit of intensity and they gain numbers (minus ones for pains, plus for pleasures) by comparison with that standard. (For example, let one intensity unit equal the highest intensity of pleasure attained by the inventor of the vinyl auto roof in the minute following the realization of the significance of his discovery.) That gives some quantities for a hedonistic theory of value to work with. We do not assume which, if any, function from those numbers to intrinsic value ratings is the correct one. I believe that even the purely quantitative hedonist has choices here. For instance, the intrinsic value might be the square of the intensity of the pleasure, and (minus) the cube of the intensity of the pain. This would give us a way to bear out a certain difficult saying of Moore's:

[P]ain . . . appears to be a far worse evil than pleasure is a good.  
(Principia Ethica, p. 212)

We read in the prefix: "At any given intensity, . . . "

<sup>2</sup>John Bennett in "The Problem of Interpersonal Utility Comparisons" (unpublished manuscript) points out that measurement theory imposes certain constraints on what pleasure and pain can be like if they are to have this sort of quantifiable aspect. I see no grounds to doubt that these constraints are obeyed in the case of the intensity of these attitudes. If, as Bennett contends, we must prove that there is this obedience, I do not know how to do so. I do not know any reason to think this must be shown, however.

<sup>3</sup>We should not get carried away about what we can learn about a concept by seeing what behavior might suffice to teach it, or an expression for it. The important features of this example are those making it believable that the child gains awareness of intrinsic goodness with no awareness of pleasure. It would not suffice to observe that "intrinsically good" can be taught without expressing or referring explicitly to the concept of pleasure. Meanings can be taught by introducing the expressions in situations involving mere cues that make it psychologically normal to bring the appropriate concept to mind. Nothing close to a synonym or conceptual breakdown need be given. So, from our not finding a concept expressed or denoted by anything other than the expression being taught, we cannot validly infer that the concept is not part of the expression's sense.

It is worth looking briefly at an argument that such considerations suggest: It is plausible that the following two means of conveying concepts are exhaustive--explanation in terms of previously acquired concepts, and ostensive explanation. So if the example of the child is



possible, then it seems we can conclude that something other than pleasure is intrinsically good. For the child was not given an explicit definition of "intrinsically good." Thus he must have been shown some examples and he was not shown any cases of pleasure. The problem here is that explaining in terms of other concepts does not reduce to providing a synonymous expression as a definition. Dictionaries rarely provide actual synonyms. The way we acquire concepts by looking up words in them includes being guided by the given definiens (and perhaps some illustrative contexts of use) to latch onto a concept that it is not psychologically natural or normal then to think of, though that concept is not expressed by the definiens. And we can learn a concept by ostensions that do not include even one positive instance of it if, for example, our knowledge of the pointer enables us to think of what he would (perhaps perversely) take the indicated items to share. Very little about what is analytic to a concept can be inferred from the mere fact that certain utterances or gestures can suffice to convey it.

<sup>4</sup>Presented in a seminar at UMass/Amherst, in Spring 1978. Unpublished, I think.

<sup>5</sup>Presented in the aforementioned seminar.

<sup>6</sup>See further on in this section for more on such puzzles, and how they might be solved.

<sup>7</sup>The vast literature on this starts with C. H. Langford, "Moore's Notion of Analysis," from P. A. Schlipp (ed.), The Philosophy of G. E. Moore (New York: Tudor Park (1952), pp. 319-343.

<sup>8</sup>E. Klemke makes this point with documentation in the chapter "Analysis" of The Epistemology of G. E. Moore (Evanston: Northwestern University Press, 1969), pp. 64-91.

<sup>9</sup>If Chisholm is right that (b) does not involve (a), then coming to see their equivalence can make a contribution to our understanding of what is present where brotherhood is instantiated. It may strain the metaphor to say that this "illuminates" the concept itself. More precisely, it illuminates those situations where the concept is instantiated. When the nature of those situations is something we seek to understand about the concept, then it is fair to say that it helps with our understanding of the concept. (Note that that is something a meaning identity claim cannot do for us.)

<sup>10</sup>Franz Brentano, The Origin of Our Knowledge of Right and Wrong, (New York: Humanities Press, 1969), p. 18.

<sup>11</sup>Ibid., p. 22.

<sup>12</sup>Franz Brentano, The Foundation and Construction of Ethics (New York: Humanities Press, 1973), p. 164.

<sup>13</sup>Monroe Beardsley expresses doubts about the existence, or at least the actual application, of intrinsic goodness in "Intrinsic Value" (Philosophy and Phenomenological Research 25 (1965), pp. 1-18). He traces such reservations to writings by John Dewey.

<sup>14</sup>Philosophical Studies, forthcoming.

<sup>15</sup>Mark Pastin has made several efforts in this direction. See, e.g., "Warranting Reconsidered," Synthese 38 (1978), pp. 459-464. For problems with Pastin's approach, see Fred Feldman, "Final Comments on the Analysis of Warranting," Synthese 38 (1978), pp. 465-469.

<sup>16</sup>Brentano, The Origin of Our Knowledge of Right and Wrong, p. 22.

<sup>17</sup>A. Meinong, On Emotional Presentation (Evanston: Northwestern University Press, 1972), p. 121.

<sup>18</sup>G. E. Moore, Ethics (London: Oxford University Press, 1965), pp. 19-20.

<sup>19</sup>G. E. Moore, Principia Ethica (London: Cambridge University Press, 1903), pp. 39-41.

<sup>20</sup>In Ernest Sosa, "The Foundations of Foundationalism," (Nous, forthcoming).

<sup>21</sup>Good criticisms of what Moore says about the "natural/non-natural" property distinction appear in Fred Feldman, Introductory Ethics (Englewood Cliffs: Prentice-Hall, 1978), pp. 203-205. Sosa does not attempt to explain the "evaluative/non-evaluative" property distinction. I know of no good account elsewhere.

<sup>22</sup>D4.9 of Chapter IV below excludes e6 as a basic bearer of hedonic value.

<sup>23</sup>This is attempted in Chapter IV below in sections D-F.

## CHAPTER I I

### PLEASURE AND THE ANALYSIS OF INTRINSIC GOODNESS

Many philosophers have sought analytic connections between positive emotive attitudes and intrinsic goodness.<sup>1</sup> Indeed, proposals that appeal to pleasure in particular continue to be made.<sup>2</sup> As I estimate the situation, such connections can be fully investigated by carefully examining what is said and suggested by certain writings by Franz Brentano on the topic. The fundamental variations and problems among accounts with some initial plausibility all arise quite naturally in a study of Brentano's work. That is how the topic will be approached here.

Methodologically speaking, if the remarks about "analysis" in the Introduction are correct, then there are several sorts of "analytic" relationships which might be discovered. Two factors stand out as appropriate tests for the presence of such connections. First, we have seen that necessary equivalence is requisite for any such link. So a counterexample precludes any such tie. And second, circularity is a major liability. Circular equivalences can bring to light non-obvious implications of the concept at stake. But they cannot accomplish the distinctively "analytic" tasks: removing obscurity, reducing ontology, "constructing" concepts.<sup>3</sup> Thus, only after these two tests have been passed would it be worthwhile to sort out which "analytic" tasks an equivalence performs.

### A. Initial Readings of Brentano's Analysis

Brentano's main purpose in The Origin of Our Knowledge of Right and Wrong<sup>4</sup> is to say how we can gain adequate evidence for judgments concerning the moral status of acts. In his view that evidence partly consists in what we can know to be good and bad. He proposed criteria by means of which we are to be able to acquire such knowledge. Some comments in the essay are clearly intended as some sort of explanations of concepts of goodness. These criteria and comments suggest interesting and intuitively attractive equivalences between conditions involving pleasure and intrinsic goodness. Those suggestions will be developed and assessed here.

The first passage we should look at is this one:

We call a thing good when the love relating to it is correct. In the broadest sense of the term, the good is that which is worthy of love, that which can be loved with a love that is correct.

24 Among the things that please us, we may distinguish between those that are pleasing in themselves and those which are pleasing in virtue of something else. In the latter case, the thing is pleasing in virtue of what it brings about or preserves or makes probable.

Hence we must distinguish between primary and secondary goods-- between what is good in itself and what is good in virtue of something else. The useful is a clear example of the latter type of good.<sup>5</sup>



The part of this citation before section 24 plainly is not intended to be an account of intrinsic goodness. Brentano surely means the distinction between "what is good in itself" and "what is good in virtue of something else" to be the distinction between intrinsic and extrinsic goodness. And while the first part of the citation is said to be about "good" read "in the broadest sense of the term," the citation is followed with the remark that the good in itself is the "'good' in its narrow sense."<sup>6</sup> But it will be seen that on one interpretation Brentano is drawing upon part of his account of the "broadest" sense of "good" when he tells us about the good in itself in section 24. So it is worthwhile to have the whole citation available as one piece.

How shall we understand the cited comments about the good in itself? There is no clinching indication what sort of explanation of the concept of intrinsic goodness is being attempted. The use of "hence" in the passage suggests that the distinction between what is intrinsically good and what is extrinsically good is supposed to follow directly and obviously from the fact that some things are "pleasing in themselves" and others are "pleasing in virtue of something else." So the closer the connection between being pleasing in itself and being intrinsically good on the one hand, and being pleasing in virtue of something else and being extrinsically good on the other, the better justified is the remark. Also, an equivalence between the former pair would give us a dramatically simple link between pleasure and intrinsic goodness. So it behooves us to consider it.

What exactly are these two ways of being pleasing? Brentano offers no explanation. The simplest view seems to be this: We assume without explanation the two place relation of something being an object of pleasure for a person. We also assume the three place relation of something being an object of pleasure for a person because of something that the person takes to be brought about, preserved, or made probable by the first. Then we stipulate:

D2.1 p is pleasing for S in virtue of q = df. p is an object of pleasure for S because S takes it that p brings about, preserves, or makes probable q.

D2.2 p is pleasing in itself for S = df. p is an object of pleasure for S, and there is no q such that p is pleasing for S in virtue of q.<sup>7</sup>

Applying this distinction to the text in the most straightforward way, we get:

D2.3 p is intrinsically good = df. There is someone, S, such that p is pleasing in itself for S.

As pleasing as D2.3 is in virtue of the relatively simple tie it asserts between pleasure and intrinsic goodness, it is unacceptable. Moreover, it should not finally be attributed to Brentano, since he gives what is in effect an excellent counterexample to D2.3 almost immediately after our first citation:

[I]t often happens as a result of habit that what is first desired merely as a means to something else comes to be desired for itself alone. Thus the miser is reduced to heaping up riches irrationally . . .<sup>8</sup>

Thus Brentano describes a case where something intrinsically neutral--the miser acquiring riches--becomes desired (and, we can assume, pleasing) in itself for him.

Another problem for D2.3 is that it is plainly asking too much of the intrinsically good things to require that each actually be the object of someone's pleasure.

Were difficulties of these two sorts the only faults in D2.3, it could be repaired with ease. In cases like that of the miser, the states that become pleasing in themselves do not start out that way. It may be credible that they would only come to be pleasing in themselves as a result of habituation. We might then interpret Brentano to have been proposing that the intrinsic goods are states which are spontaneously pleasing in themselves. To meet the second difficulty, we can read the proposal as explicitly giving only a sufficient condition. The most natural modification to gain a necessary and sufficient condition has intrinsic goods being possible objects of the right sort of pleasure. Those two changes yield:

D2.4 p is intrinsically good = df. Possibly, there is someone, S, and a time, t, such that p is pleasing in itself for S at t, and at no time up to t was there a q such that p was pleasing to S in virtue of q.

Brentano provides in the essay the basic materials for a counterexample to D2.4, too. Soon after our first citation, he suggests that we ". . . imagine now another species quite different from ourselves; . . . they . . . despise insight and love error for its own sake."<sup>9</sup>

We have a counterexample to D2.4 when we supplement the case with the modest further assumptions that the species takes the attitude spontaneously, that error is not intrinsically good, and that such a species is possible. Even if normal humans do not take such attitudes, we cannot plausibly preclude their possibility.

It is worth adding that there are many who are spontaneously pleased by various intrinsically neutral states. Some delight in the thought that the universe had a beginning; others take joy in the thought that it is not the case that the universe had a beginning. Some members of each group do not find their thought about the origin of the universe delightful because of anything else, and never did. Yet we can take it as axiomatic that not both a state and its negation are in-

trinsically good. In this case it is plausible that each is neutral.

The sort of objection raised for D2.4 indicates a basic problem for its approach. D2.4 claims that the intrinsic goodness of something is implied by the possibility of someone addressing a simple kind of psychological attitude toward it. But reflection reveals no credibility to the supposition that the possible extension of any such attitude is restricted in a normatively interesting way. Indeed, they seem virtually unrestricted (perhaps one of the few exceptions is that the most blatant of contradictions cannot be believed). Thus no attitude by itself will isolate the intrinsic goods. When evaluating such a proposal, we have only to remind ourselves how irrational people can be, how willful, and how base. We then see that such attempts are not promising.<sup>10</sup>



And we also see that restrictions upon what other attitudes have been, or are being, held do not significantly improve the account.

### B. Revised Interpretations of Brentano's Analysis

Obviously the possibility remains that non-attitudinal restrictions on pleasures will work. We can approach that question by investigating a way in which we may have misconstrued the initially cited remarks. Looking back at the passage, we see that Brentano distinguishes the two ways of being pleasing right after having spoken of "love" that is "correct." Now it is clear from several comments in the essay that "love" is being used there in a particularly broad manner. In one place, he writes of giving preference to one thing over another as "loving" the one "more" than the other.<sup>11</sup> He also writes that ". . . a natural feeling of pleasure is a higher love . . ."<sup>12</sup> Thus it seems proper to interpret some of his uses of "love" to be his way to express a variety of attitudes involving the taking of pleasure. So it may be that Brentano meant still to be discussing "love" that is "correct" in section 24. The two ways of being pleasing (i.e., "loved") would then be asserted there to identify the two sorts of goods when the pleasure is "correct." That allows us to read Brentano to equate being intrinsically good with being possibly correctly pleasing in itself for someone.

But of course to understand that we need to know what Brentano meant by "correct" in the passage. It is vital that we learn what we can from Brentano about this concept of correctness, anyway. He makes

use of it in his criterion for determining that a thing is intrinsically good, a topic of Chapter III. Unfortunately, we are told much less than it would be helpful to know about the notion. It is not hard to say very roughly what Brentano had in mind. Correctness is to be a property of some emotive attitudes that is closely analogous to the property of truth as it pertains to beliefs.<sup>13</sup> Some feelings toward things are to be "the right attitude to take" given their value, just as belief is the right epistemic attitude to take toward what is true.

The way Brentano introduces the concept of correctness in the essay is as follows. First he distinguishes three categories of psychological acts: having ideas or "presentations," making judgments, and undergoing emotions. The latter two, unlike the former, are said to involve an important type of opposition. In the category of making judgments, this opposition relates affirmation to denial; in the category of undergoing emotions, it relates love to hate, and being pleased to being displeased. No opposition is found in the category of having ideas.<sup>14</sup> Brentano follows those claims with this:

This fact has an important consequence. Psychological acts that belong to the first class [having ideas] cannot be said to be correct or incorrect. But in the case of acts that belong to the second class [making judgments], one of the two opposing modes of relation--affirmation and denial--is correct and the other is incorrect, as logic has taught since ancient times. Naturally, the same thing is true in the third class. Of the two opposing types of feeling--

loving and hating, being pleased and being displeased--in every instance one of them is correct and the other is incorrect.<sup>15</sup>

This passage is immediately followed by the first citation above. We gain a loose but workable grip upon the concept from this account. It is valuable to proceed with this degree of understanding, if only in order to find out where we need more information. I do not see how we can do better on the basis of texts alone. In the last sections of this chapter I try to explain the residual unclarity about correctness.

We can now consider the proposal that to be intrinsically good is to be possibly pleasing in itself, where the pleasure is correctly taken, i.e.,

D2.5 p is intrinsically good = df. Possibly, there is someone, S, such that p is pleasing in itself for S, and that p is an object of pleasure for S is correct.<sup>16</sup>

D2.5 seems to be too weak. For anything that might be overall good seems to be a possible object of correctly taken pleasure, though that cannot be definitely established without a fuller account of correctness. At the beginning of the first passage cited above, Brentano holds that anything "good" in the "broadest" sense is a possible object of correctly taken pleasure. It is not clear what is to be that "broadest" sense: perhaps it is to be overall goodness, perhaps the disjunction of intrinsic and extrinsic goodness. But on either reading, some things that are not intrinsically good would be possible objects of correctly taken pleasure. For on either reading, intrinsically neutral overall goods could be objects of correctly taken pleasure. And

intuitively speaking, if intrinsic goodness makes something a "fitting" or "correct" object of pleasure, then so does overall goodness. At least it seems that any notion of correctness that excludes that possibility must be more completely explained. It is also plausible that an intrinsically indifferent overall good might be pleasing in itself for someone. So it seems that too much satisfies the definiens of D2.5.

A conceivable defense of D2.5 against this sort of charge consists in claiming that there are epistemic restrictions on when pleasure taken in what is somehow good is correct. These requirements might be thought to get in the way of a non-intrinsic, overall good being pleasing in itself for someone in the appropriate epistemic situation.

But what requirements would work? None that I have found. For example, it does not help if the person must know that the neutral is overall good for the pleasure to be correct. Smith might know that there being radar is overall good, but be unmoved emotionally by that fact, and by anything he takes to be brought about, preserved, or made probable by radar. He need be no more than a little bit screwy to feel glad nonetheless that there is radar. He would then pass the conditions for it being correctly pleasing to him, and pleasing in itself.

Another modification might seem better. Perhaps we should require that the pleasing in itself relationship be correct, rather than just the pleasure taken in the object. Then we have:

D2.6 p is intrinsically good = df. Possibly, there is someone, S such that p is pleasing in itself for S, and that p is pleasing in itself for S is correct.



Epistemic considerations seem actually to work against D2.6. Consider someone who has fine evidence concerning something which is not in fact intrinsically good. Suppose the person manages to summon up a pleasurable regard for that thing, and not in virtue of anything else. Intuitively, that appears to be a "fitting" or "correct" attitude for him to take. That would refute D2.6, though.

It might be replied that this last example fails to accommodate Brentano's analogy between correctness and truth. Pleasure under the described circumstances would be analogous to justified false belief, and thus not really correct.

The trouble with this reply is that pleasure in itself toward the intrinsically indifferent overall good seems sufficiently analogous to true belief to refute D2.6. Now I do not wish to deny that there is a way to construe "correctness" where only intrinsic goods are possible objects of a certain sort of "correct" pleasure.<sup>17</sup> But I think we must conclude that, in the absence of a suitable clarification of the nature of correctness, the D2.5-D2.6 approach does not appear to succeed.

### C. Analysis by Use of Brentano's Criterion

Just after the first citation above, Brentano writes, "So much for the concept of the good."<sup>18</sup> But we should not let that discourage us. Brentano proceeds to offer a criterion for knowing something to be intrinsically good. It seems to adapt to yield an account of intrinsic

goodness, one that appears better equipped to cope with the epistemic problems that arose above.

Here is his (terse) formulation of the criterion:

. . . in the case of [pleasure in the clarity of insight] the natural feeling of pleasure is a higher love that is experienced as being correct. When we ourselves experience such a love we notice not only that the object is loved and capable of being loved . . . but also that it is worthy of being loved . . . and therefore that it is good.<sup>19</sup>

Here, then, is Brentano's criterion in his own (translator's) words:

C2.1 If S has a natural feeling of pleasure toward p and S experiences that feeling as being correct, then p is known by S to be intrinsically good.

Of course, to understand C2.1 we have to understand what it is for an emotion to be "experienced as being correct." Brentano tries to explain that notion by analogy with our knowledge of self-warranted propositions. He begins by contrasting "blind" or "impulsive" judgments with "insightful" or "evident" ones, where examples of the latter are to include belief in the law of non-contradiction and our knowledge of our own perceptual states. After noting that the difference between the two kinds of judgments is not a matter of degree of conviction, he writes:

If one were to ask [concerning an immediately evident belief] 'Why do you really believe that?' . . . it would be impossible to find

rational grounds . . . but . . . the clarity of the judgment is such as to enable us to see that the question has no point . . .<sup>21</sup>

Brentano then asserts:

. . . [there is] an analogous distinction between the higher and lower types of activity in the emotional sphere . . . there is a higher mode of being pleased . . . it is the analog of something being evident in the sphere of judgment.<sup>22</sup>

Apparently, Brentano holds that "experiencing pleasure as being correct" amounts to being immediately aware of having a "higher love," where the obvious presence of the emotion parallels the "clarity" of the judgement, and the "height" of the emotion constitutes its correctness. What is this elevated type of pleasure, this "higher love"? Brentano does not further identify it. The best construal seems to be that he means to refer to some species of liking that is the intuitive opposite to "blind" or unreflective favoring, something like contemplative approval. The reading of C2.1 this suggests is:

C2.1a If it is self-evident to S that S is feeling contemplative approval for p, then p is known by S to be intrinsically good.

We can take "self-evident" in an intuitive sense for now, and take it for granted that this sort of approval is self-evident to the one who feels it. It is plain that C2.1a is not a credible criterion, nor does it give us a plausible equivalence with intrinsic goodness. First, a brief consideration of the variety of attitudinal propensities that obtain should convince us that not only intrinsic goods are objects of the "exalted" approval of C2.1a (or any other kind of attitude). And of

course the attitude will be self-evidently present when addressed toward something not intrinsically good, too. Second, on this interpretation the high-toned nature of the feeling constitutes the correctness of the emotion. Yet C2.1a has it that correct (i.e., high-grade) approval attaches only to what is intrinsically good. This makes it very hard to understand how Brentano could consistently say that anything good in the "broadest" sense can be correctly loved, as he does in the first quote above. This strongly suggests that we should seek a better interpretation of "experiencing pleasure as being correct."

Fortunately, a better reading is available. It is given added credence by part of a letter from Brentano to Oscar Kraus. Brentano there attempts to say how we acquire the concept of correctness. He claims that we do it by observing several emotive acts that exemplify it and seeing it as something they have in common. Then he writes:

We know with immediate evidence that certain of our emotive attitudes are correct. . . . We will find that there are others whose emotive attitudes correspond to our own. . . . If their attitudes should happen to be only a matter of habit or instinct we may still say that they are correct but not that they are experienced as being correct. . . . One can never find the criterion of correctness in an adaequatio rei et intellectus vel amoris, it can be found only in those attitudes which we know with immediate evidence to be correct.<sup>23</sup>

The interpretation that is thereby made reasonable reads "experiencing pleasure as being correct" to mean making a self-warranting



judgment that the pleasure is correct. So we get:

C2.1b If S feels contemplative approval for p and it is self-evident for S that it is correct for S to feel that way about p, then p is known by S to be intrinsically good.

We should have a somewhat precise notion of self-evidence to work with. But first let us get before us the thing that is our principal concern in this chapter: the equivalence with intrinsic goodness suggested by the criterion. Of course we must not suppose that each intrinsic good is actually an object of approval. But if we make the reasonable assumption that C2.1b is necessarily true if true at all, then if C2.1b is correct only intrinsic goods can satisfy its conditions. And it seems safe to say that if any can, then they all can. So here is our new analyzing proposal:

D2.7 p is intrinsically good = df. Possibly, there is a person, S, such that S feels contemplative approval for p, and it is self-evident for S that it is correct for S to feel that way about p.

It will be harmless and helpful to use this concept of self-evidence:

D2.8 p is self-evident for S = df if p is true; p is evident for S; and for any q such that q makes p evident for s, q entails<sup>24</sup> p.

Roughly, D2.8 permits only p itself and conjunctions of p with other things evident for S to make p evident when p is self-evident.<sup>25</sup>

There is trouble for D2.7. Recall that being "good" in the "broadest" sense is to be a necessary condition for being an object of correctly taken pleasure. But can what is self-evident for a person ever guarantee that something is "good" in that sense? If not, it could not be self-evident that a state passes this requirement of "broad" goodness for being an object of correct pleasure. And that would be a barrier to the correctness itself being self-evident. "Broadest" goodness cannot be self-evident if the sense of "good" in question is that of being overall good. The state's causal contributions cannot be self-evident, and they must be known to determine its overall value status. If instead the "broadest" sense is that of being either intrinsically or extrinsically good, then there is no obvious problem here. If D2.7 is right, then something that it contends can be self-evident implies the intrinsic goodness of the state at stake. So the necessary condition on this reading of "good" would automatically be met.

But this latter reading creates another problem: Is it really correct to contemplatively approve of something intrinsically good but disastrous in its consequences? If not, there would be no way for the correctness of the approval to be self-evident. And it seems not. I think there is a reading making such approval of an intrinsic good "correct." But it brings to light considerations that make grave trouble for the whole D2.7 approach. So I want to postpone that to try something else that is a reasonable interpretation and does not directly raise that problem. The point right here is that the meager intuitions

that are apparently good ones to use given Brentano's account of correctness make it doubtful that approval can be self-evidently correct.

It may be that the sort of "love" that Brentano intends in his comments about the "broadest" sense of "good" does not include the pleasing in itself variety. Such a consideration led us to D2.6, which seemed to fail essentially because of the possibility of misleading evidence. But the self-evidence requirement of the D2.7 approach might well help with the problematic sort of epistemic situation. In order also to accommodate the notion of "higher love," we can concoct the phrase "approved in itself," where  $p$  is so approved just when contemplatively approved, and not in virtue of anything else.

Now we can have:

D2.9  $p$  is intrinsically good = df. Possibly, there is someone,  $S$ , such that  $p$  is approved in itself by  $S$ , and that  $p$  being so approved is correct is self-evident for  $S$ .

The concept of being approved in itself is supposed to capture the idea of being approved "for its own sake," approved "for what it is, rather than what it does." To the extent that what is intuitive in those phrases is captured, it seems more reasonable in the case of D2.9 than in that of D2.7 to say that such approval of something intrinsically good is "correct" whether or not the thing is overall good. Thus D2.9 seems to make a more credible demand on what can be self-evident than D2.7.

D2.9 is an interesting equivalence. I see no example that clearly refutes it. It is not manifest whether what it requires to be self-

evident really can be. I think determining that goes beyond the rough intuitions about correctness to be gained from Brentano's comments about it. But I see some reasonable temptation to say that we could determine the relevant correctness by inspection alone. What I take to account for that temptation, however, ultimately undermines D2.9.

It is time to improve our understanding of correctness. Brentano's analogy of correctness to truth lead us to try to take it to be a one place property. I believe there is no such concept of correctness. I suggest that what is meant by the kind of use of "correct" intended in D2.9 is a relation.<sup>26</sup> Other locutions, clearly synonymous with "correct" in many contexts, are always expressions for relations. When the topic is the evaluation of actions, it is merely a question of terminology whether we say that an action is "proper" or "correct." Yet a relational sense is intended. An action can be "proper" in one respect and not "proper" in another. If Jones knows that only by breaking rudely into a conversation between Smith and Robinson can Robinson be spared a painful revelation, then it might be that Jones should do it. If so, then the act would be both proper (or correct) and improper (or incorrect). For it would be morally proper and improper etiquette. Such an example can show us that "proper" and "correct" at least often express relations where one term is not explicitly mentioned. Context typically determines what it is.

It might be thought obvious that "correct" sometimes expresses a one place property. We say that a belief is "correct" and seem to mean that it is true. But even there, I think our precise meaning is that



the belief is correct with regard to its truth value. Note that true beliefs are also "incorrect" when they are politically forbidden. I propose that a relation is always expressed.

One immediate benefit of that view is the easy explanation it provides for our trouble in being satisfactorily clear about what Brentano meant by "correct." Since he makes no mention of what it is in relation to which the pleasure is to be correct, and since his analogy with truth stifles any attempt to be guided by context, it is no wonder we had a hard time deciding important questions about its application. The case is analogous to that of being told that there is a notion of usefulness which is much like truth, and then being told that a certain interesting category of things are "useful" in that sense. It is likely that context would help somewhat in determining what purpose the things were to be useful for. But the analogy and the absence of explicit relata would allow us at best only a rough understanding.<sup>27</sup>

If "correct" expresses a relation in D2.9, relative to what is the approval to be self-evidently correct? Familiar terms like moral status and truth value clearly will not do. In the former case, consequentialists have it that there is no way for the moral status of an act of approval to be self-evident. They say it is determined by a feature of the act that contemplation cannot reveal: the value of its consequences versus the value of its alternatives' consequences. It is more difficult to speak for all non-consequentialists on this question, but any plausible moral theory will have it be possibly morally correct for an intrinsic neutral to be approved in itself by someone. After all, it is

a harmless and pleasant mental episode under most conditions. And if it can be self-evidently moral to do this approving in the case of an intrinsic good, it is reasonable to suppose that it can also be so in the case of an intrinsic neutral. But then the neutral satisfies the moral status relativisation of D2.9. So D2.9 fails. The truth value sort of correctness makes D2.9 amount to this: *p* is intrinsically good just when it is possible that it is self-evident for some *S* that *p* is approved in itself by *S*. The point about the unrestricted application of attitudes shows that that is wrong. Any state might be approved in itself, and approving can be self-evidently directed toward non-intrinsic goods just as well as toward intrinsic goods.

A relativisation that does not seem open to counterexamples has it that the approving is self-evidently correct with regard to the value status of the thing approved. Not overall value status. As already mentioned, it cannot be self-evident that one is approving an overall good. D2.9 is most plausibly held to lack counterinstances when the approving is to be self-evidently correct with regard to the intrinsic value status of the approved. Here is a principle to which I have no objection:

P2.1 Necessarily, *p* is intrinsically good iff possibly, *p* is approved in itself by someone, and that approval is correct with regard to the intrinsic value status of *p*.

Perhaps in the case of each intrinsic good it is possible that it be self-evident for someone that that person's approval in itself of the intrinsic good is correct with regard to its intrinsic value status. I

see no good way to argue that that cannot be. If it can, then the following equivalence would seem to be acceptable as well:

P2.2 Necessarily,  $p$  is intrinsically good iff possibly, someone,  $S$ , is such that  $p$  is approved in itself by  $S$ , and that  $p$  being approved in itself by  $S$  is correct with regard to the intrinsic value status of  $p$  is self-evident to  $S$ .

P2.1 and P2.2 merit contemplation. P2.1 embodies the intuition that a certain high-minded liking of a state for its own sake is appropriate to the value status of the intrinsic goods only. P2.2 embodies the additional intuition that this liking can show itself to accord with the value status.

But these cannot be analyses. What is the concept of the intrinsic value status of a thing? It is the concept of its intrinsic goodness, neutrality, or evil. So D2.9 is rendered circular by use of that concept.

#### D. The Requirement Interpretation of Correctness

There is a last interpretation of correctness that we should try. Perhaps the approval is to be self-evidently correct "with respect to its object." This tells us that Jones contemplatively approving of Smith being happy would be correct with respect to Smith being happy itself. Can we make sense of this? I think so, if we take this kind of correctness to be the relation of a thing being fitting for something, or we use its converse--the relation of requirement. At times, Brentano

characterizes correctness in terms of fittingness.<sup>28</sup> Examples by Chisholm help to clarify the idea:

. . . promise-making requires--or calls for--promise-keeping; being virtuous, according to Kant, requires being rewarded; the dominant seventh requires the chord of the tonic; one color in the lower left calls for a complementary color in the upper right.<sup>29</sup>

It has been proposed tht we sometimes "experience" requirements.

Maurice Mandelbaum holds the relation to create the basis for felt moral demands; he holds that we then experience a requirement by a situation for an action.<sup>30</sup> So this sort of correctness might have the requisite possible self-evidence.

Thus we ought to consider this proposition:

D2.10 p is intrinsically good = df. Possibly there is someone, S, to whom it is self-evident that p requires p being contemplatively approved by S.

D2.10 has commendable features. If, as P1.1 says, it is appropriate to the value status of intrinsic goods to contemplatively approve of them, and, as P1.2 says, this might be seen by considering whether such affection is required, so much the better for D2.10 as well as P2.1 and P2.2. Intuition might reveal that overall good intrinsic neutrals can be fittingly favored when seen as such. But it cannot be self-evident that the neutral has the requisite extrinsic credentials to be overall good. So such neutrals seem not to satisfy D2.10.

Finally, though, it seems to me that there is trouble for D2.10. And it is of a sort that appears to undermine any pleasure-based



analysis of intrinsic goodness. The difficulty can be brought out by asking what really makes us think that pleasure is especially appropriately taken in things having positive value. Contingently associated influences like moral indoctrination aside, the degree to which the experience is enjoyable does not depend upon the value of its object. No degree of pleasure logically or psychologically requires a good object. What does make the apparent fit, then?

First, I contend that there is no such fittingness that is present in the case of every sort of liking. Consider:

D2.11 p is intrinsically good = df. Possibly there is someone, S, such that it is self-evident for S that p requires p being lustfully thrilling to S.

D2.11 is not plausible. Why? It is more difficult to thrill lustfully in certain lofty states than to contemplatively approve of them. But I believe that that is a distraction which does not really account for the intuitive difference. In brief, I think what explains the special suitability of the "high" pleasures is that each implies a belief that its object is somehow good. This is clear in the case of taking pride, admiring, and appreciating. I think it is also quite plain in the case of the lofty emotion we have been using--contemplative approval. In fact, approval is always approval out of regard for some (believed) feature of the object of approval, where the object is believed thereby to be somehow good.

Now we should ask--What sort of goodness is ascribed to p on the most credible construal of D2.10? Perhaps some intrinsic goods are

aesthetically good, and they require approval out of aesthetic regard for them. But not all intrinsic goods. Only approval out of regard for intrinsic goodness seems exactly fitting for all and only intrinsic goods. So the precisely stated equivalence that intuition actually favors is:

D2.12 p is intrinsically good = df. Possibly, there is someone, S, such that it is self-evident for S that p requires p being contemplatively approved by S out of regard for p's intrinsic goodness.

If, as I believe, this is the best "requirement" account that is finally intuitively acceptable, then we have not improved upon D2.9 on its best reading after all. Obviously D2.12 too is circular.

As I see it, the general situation concerning pleasure-based analyses of intrinsic goodness is this: No more fit can be seen in the sheer taking of pleasure in an intrinsic good than in the same attitude toward an intrinsic neutral. There is a special appropriateness to certain sorts of pleasure in the intrinsic goods--the sorts with an implication of believed goodness. But when we think narrowly and sharply about just when these attitudes seem precisely called for, we see that it is when they involve recognition of the intrinsic goodness of their objects.<sup>31</sup> To identify exactly the required attitudes in our equivalence, we must use the concept of intrinsic goodness on both sides. Thus such equivalences may be interesting, but as analyses they are circular.

Notes to Chapter II

<sup>1</sup>The Naturalist tradition in meta-ethics is rife with such attempts.

<sup>2</sup>See for example John Bailey, "On Intrinsic Value," Philosophia 10 (1980), pp. 1-8.

<sup>3</sup>See the Introduction above, section B.

<sup>4</sup>Franz Brentano, The Origin of Our Knowledge of Right and Wrong (New York: Humanities Press, 1969), Elizabeth Schneewind and Roderick Chisholm, translators.

<sup>5</sup>Brentano, p. 18.

<sup>6</sup>Ibid.

<sup>7</sup>D2.1 and D2.2 do not allow that something "pleasing in itself" is also "pleasing in virtue of something else." No text I have found is decisive on whether Brentano intended the concepts to be compatible. I believe that nothing below turns on this feature of D2.1 and D2.2.

Fred Feldman has observed that we might take "p is pleasing for S in virtue of q" as primitive, and define "p is pleasing in itself for S" as p is pleasing for S in virtue of p. That would seem to leave open the compatibility. But I fear that that does not increase clarity. I would prefer to take both of Brentano's locutions as primitives. The reader is urged to adapt whichever reading is most intelligible to him.

<sup>8</sup>Brentano, p. 19.

<sup>9</sup>Ibid., p. 22.

<sup>10</sup>Such considerations make trouble for Bailey's account, "x is intrinsically good = df x is or would be liked for its own sake" (Bailey, p. 1).

<sup>11</sup>Brentano, p. 26.

<sup>12</sup>Ibid., p. 20.

<sup>13</sup>Ibid., p. 18.

<sup>14</sup>Ibid., pp. 15-17.

<sup>15</sup>Ibid., pp. 17-18.

<sup>16</sup>It is helpful to assume that correctness applies to "singular propositions" (see below, Chapter IV, Section C) to the effect that S takes pleasure in p or the like. But if these are rejected on metaphysical grounds, we can take it to be a relation that can hold between an individual, the relation of taking pleasure, and the object of pleasure (for now. See the end of section C and section D below).

<sup>17</sup>See the end of section C below.

<sup>18</sup>Brentano, p. 18.

<sup>19</sup>Ibid., p. 22.

<sup>20</sup>Ibid., p. 18.

<sup>21</sup>Ibid., p. 20.

<sup>22</sup>Ibid., pp. 20, 21, and 22.

<sup>23</sup>This letter is translated in full in: Roderick Chisholm, "Brentano's Theory of Correct and Incorrect Emotions," Revue Internationale De Philosophie 20 (1966), pp. 399-400.

<sup>24</sup>p entails q iff necessarily, whoever believes p believes q.

<sup>25</sup>It seems that a proposition could make itself evident and also be made evident by non-entailing evidence. I think that that deviation by D2.8 from the intuitive concept of self-evidence does not matter below.

<sup>26</sup>Instead, correctness may be a property that comes in several varieties--like the property of being a rule. The difference between these alternatives does not affect our topic. Either way, specification is required before we have a single proposal before us.

<sup>27</sup>An analogy of more philosophical interest is the case where we are told that substances are the things that can exist "independently," but we are not given adequate information about what substances are to be able to be independent of.

<sup>28</sup>Brentano, p. 74.

<sup>29</sup>Roderick Chisholm, "The Ethics of Requirement," American Philosophical Quarterly 6 (1969) p. 147.

<sup>30</sup>Maurice Mandelbaum, The Phenomenology of Moral Experience (Baltimore: Johns Hopkins Press, 1966), pp. 67-71.



<sup>31</sup>It is tempting to conclude that the correctness of these partially doxastic mental states is really just the correct truth value of their belief components. We would be justified in this if we could explain away the apparent superiority in plausibility that D2.11 enjoys over:

D2.13 p is intrinsically good = df. Possibly, there is someone, S, such that it is self-evident for S that p is intrinsically good.

I think tht D2.11 seems better (to those to whom it does) largely because it is supposed that the feeling component in approval might give evidence of intrinsic goodness which would enable the requirements of D2.11 to be self-evident under some possible conditions; whereas it is not clear that a thing's intrinsic goodness might be self-evident, as D2.13 calls for. The question whether such feelings might supply such evidence is a topic for Chapter III below.

## CHAPTER III

### PLEASURE AND THE DISCOVERY OF INTRINSIC GOODNESS

The question here is whether pleasure can help us learn which things are intrinsically good. A proposal according to which pleasure does just that is one of the most significant assertions in Brentano's The Origin of Our Knowledge of Right and Wrong, as was mentioned above.<sup>1</sup> We must study that proposal here. Charles Baylis has also offered a pleasure-based criterion of intrinsic goodness that merits our consideration.<sup>2</sup> So we will give it that.

#### A. Tests for Epistemic Principles

How is an epistemic test for intrinsic goodness to be tested? The form of such a proposal seems best understood to be:

p E's q being intrinsically good,  
where "E" stands for expressions of propositional support relations such as tending to confirm and justifying,<sup>3</sup> and "p" expresses some condition involving pleasure. Since we would most like to find a way that intrinsic goods can be discovered in the first place, p should state something that is intuitively possibly evident without any other evidence about intrinsic goodness (in contrast with testimonial evidence from an expert, for example, where we would need other evidence about intrinsic goodness to discern his expertise). And the proposal will not allay doubts about evidence for intrinsic goodness if it appeals to a condition less well justifiable than the proposition that q is

intrinsically good. Finally, such a criterion is unsatisfactory if it is satisfiable by a diversity of q's that are not intrinsically good on any plausible value theory, unless it is supplemented by an account showing how to exclude those. The principle is not shown false by being thusly satisfied, since there is such a thing as misleading evidence. But without a way to narrow down at least to things controversially regarded as intrinsically good, we have no reason to believe that it is even guiding us toward the intrinsic goods.

### B. Brentano's Criterion

Unfortunately, Brentano is quite brief on the topic of a test for intrinsic goodness. The relevant passages were all cited in Chapter II.<sup>4</sup> Here is the statement of the test itself again:

. . . in the case of [pleasure in the clarity of insight] the natural feeling of pleasure is a higher love which is experienced as being correct. When we ourselves experience such a love we notice not only that the object is loved and capable of being loved . . . but also that it is worthy of being loved . . . and therefore that it is [intrinsically] good.<sup>5</sup>

To reiterate briefly the interpretations of Chapter I,<sup>6</sup> "being correct" seems best read to express the relation of fittingness, to "experience [the fit] as being correct" seems best read to assert the fit to be selfevident, and the pleasure which is "a higher love" seems best read as contemplative approval. In pursuit of an analysis of intrinsic goodness, these readings lead us to:

D2.12 p is intrinsically good = df. Possibly, there is someone, S, such that it is self-evident for S that p requires p being contemplatively approved by S out of regard for p's intrinsic goodness.

For all the critical comments of Chapter II show, the left and right sides of D2.12 may be necessarily equivalent. I have no objection to that equivalence. Our question here, though, is whether we can extract from D2.12 evidence involving pleasure for intrinsic goodness. We could approach that question by considering whether the right side of D2.12 evidentially supports the left, i.e.:

E3.1 It being self-evident for some S that p requires p being contemplatively approved by S out of regard for p's intrinsic goodness tends to confirm that p is intrinsically good.

But I think that E3.1 is not the best epistemic principle that can be gotten from D2.12. We seek evidence for intrinsic goodness which can be at least as secure epistemically as the belief that the thing in question is intrinsically good can be. I do not see how to demonstrate this, but it seems clear that the self-evidence of the relevant requirement cannot be as well warranted as the intrinsic goodness alone might be. For the former seems to depend on evidence for the intrinsic goodness and for the presence of the requirement and its self-evidence. Also, recognizing that approval out of regard for intrinsic goodness is called for seems to depend upon having used other evidence to discern the intrinsic goodness. So it appears that seeing the requirement could not be our means for gaining initial warrant for the intrinsic



goodness.<sup>7</sup> I think we do better by turning our attention to:

E3.2 S contemplatively approving of p out of regard for p's intrinsic goodness tends to confirm that p is intrinsically good.

In any case, the trouble attributed to E3.2 below seems at least as serious a problem in the case of E3.1, so we do Brentano no disservice by concentrating on E3.2. E3.2 has the distinct advantage that its justifying condition is plainly capable of the highest possible degree of justification for S. (Recall that the approval is to be out of regard for what is taken to be p's intrinsic goodness.)

We should note concerning E3.2 that its justifying condition uses the very concept the application of which it is to justify. That raises once again the unwholesome spectre of circularity. But use of the concept in question is not an objectionable feature of justification principles. For instance, the use of the concept of being an ovoid egg is innocuous here:

(a) All eggs in a large random sample being ovoid tends to confirm that all eggs are ovoid.

For another example, that I seem to see something white justifies itself. Indeed, the fact of self-justification shows that some justifiers flagrantly require the justified to be evidence for the justifier. So the circularity that ruins an analysis is not an objectionable trait here. The most similar feature that is a flaw in an epistemic principle is its presupposing other evidence for the justified of at least equal strength to the evidence it is supposed to supply. That is the fault that spoils this:

- (b) Someone knowing that some suits are worsted wool tends to confirm that some suits are worsted wool.

E3.2 does not give any sign of being flawed in that way.

On the other hand, whenever S approves of p out of regard for p's (believed) intrinsic goodness, S thinks of the proposition that p is intrinsically good. And it seems that intrinsic goods might be self-evidently so. That is, perhaps:

- E3.3 If p is intrinsically good, then possibly, for some S, it is self-evident for S that p is intrinsically good.

If E3.3 is true, then we ought to doubt that pleasure is playing any epistemically useful role in E3.2. S may be getting all the evidence for p's intrinsic goodness from the thought that p is intrinsically good itself. Pleasure would then have no greater epistemic function in E3.2 than it does in:

- (c) Jones taking pleasure in his seeming to see the door closed tends to confirm that the door is closed.

Clearly, the seeming to see supplies all the epistemic oomf in (c). We can tell that this is so by noting the truth of (c), the falsity of:

- (d) Jones taking pleasure in the door being closed tends to confirm that the door is closed.,

and that the justifier in (c) is no better support for the door being closed than is the seeming to see by itself, i.e.,:

- (e) Jones seeming to see the door closed tends to confirm that the door is closed.

How analogous are E3.2 and E3.3 to (c) and (e)? Roughly speaking,<sup>8</sup> I find E3.2 and E3.3 equally plausible. No consideration seems to differentiate their credibility decisively. If there is no such consideration, then however reasonable E3.2 is, it does not give evidence that pleasure plays a more important role in the discovery of intrinsic goodness than it does in the discovery of closed doors. (It should be acknowledged that this epistemic equivalence between E3.2 and E3.3 seems subject to rough estimate only. It is by no means obvious what, if anything, satisfies either. And this is so despite the fact that neither uses intolerably unclear terms.)

We must not conclude from this equivalence that the pleasure in E3.2 does not help to justify *p* being intrinsically good. That would be to employ a principle relevantly like:

E3.4 If A and B epistemically supports C to exactly the same degree as B does, then A does not support C.

To see the error in E3.4 we need do no more than consider a case where A both contains support for C of the same strength as B, and contains evidence that discounts B. For example, let A = Jones seems to hear a door close and seems to remember the expert telling him that if he seems to see a door close, that is a hallucination; B = Jones seems to hear a door close; and C = A door is closed.

What we are entitled to say at this point about E3.2 and E3.3 is that, given their equal plausibility and the inclusion of E3.3's evidence in E3.2, we have no reason yet to believe that pleasure plays a significant role in the discovery of intrinsic goodness. Pleasure would

be shown to be of special help in determining intrinsic goodness if pleasures not taken in consideration of intrinsic goodness (i.e., pleasures not implying that one think of p being intrinsically good) can be shown to give evidence for intrinsic goodness. Then there would be nothing analogous to (d)'s falsehood to complete the parallel to the exhibition of the irrelevance of the pleasure in (c). Charles Baylis has offered an epistemic principle of just the right sort.

### C. Baylis's Criterion

Baylis seeks an "identifying property" for intrinsic goodness:

. . . a discoverable characteristic the presence of which is a reliable sign of the intrinsic goodness of that thing.<sup>9</sup>

Here is his proposal:

When we judge certain things, e.g., pleasant experiences, to be intrinsically good, the best initial evidence we could have, I submit, is that we find ourselves prizing things of that kind, i.e., liking, approving, desiring, preferring and commending them, for their own qualities (rather than for their relations to other valuable things) in circumstances where to the best of our searching knowledge we are making no mistake in our cognition of them. Such evidence gives us an initial probability that we thus prize is intrinsically good.<sup>10</sup>

If "the best of our searching knowledge" can include evidence about the intrinsic value status of the thing in question, then it may be the "circumstances" rather than the "prizing" that does the epistemic



work. This is clearly not the intended interpretation. Baylis suggests that the principal sorts of errors to avoid are to be evaded by:

- (i) attending explicitly to the thing's believed extrinsic value and discounting it,
- (ii) focusing carefully on the thing and other cases of the same sort, and
- (iii) avoiding risky perceptual conditions like being tired or drunk, impassioned or prejudiced.<sup>11</sup>

Let us call a sincere effort to do these things "taking normal precautions." We do not have to find out anything about intrinsic value to take normal precautions; we will let doing so constitute "the best of our searching knowledge." And we will avoid approving and other sorts of "prizing" that imply believed goodness by appealing to liking only. Finally, we will use the concept of being pleasing in itself from Chapter II<sup>12</sup> as our rendition of ". . . liking . . . for their own qualities." Then Baylis's view becomes:

E3.5  $p$  being pleasing in itself to someone who is taking normal precautions tends to confirm that  $p$  is intrinsically good.

E3.5 is unacceptable. Its justifying condition has no tendency to be met by intrinsic goods in particular. A survey of the things that people spontaneously like would largely turn up undisputed intrinsic neutrals such as the rustling of autumn leaves, tickling sensations, the appearance of a flawless, polished chromium surface in bright light, and the feeling of extreme dizziness. These neutrals are at least as apt to be pleasing in themselves as intrinsic goods, and taking normal

precautions does not affect that. I see no variation on E3.5 that does appreciably better.

Someone might ask, "Would not a closer look at what people in these conditions find pleasing in itself reveal that it is really the pleasure taken, rather than its object? Hence, are not these examples all cases where the hedonist's view of intrinsic goods is confirmed?"

My best efforts at introspecting the phenomena in question tell me that this objection is in error. For instance, I like the taste of fresh ripe peaches. That taste is an elaborate gustatory and olfactory experience which I find I do not like because of anything it causes or makes likely. I enjoy eating peaches because that brings on pleasant experiences. But the taste itself does not please me in virtue of any extrinsic consideration--I just like it.

Perhaps Baylis was misled into thinking that E3.4 has merit by an ambiguity in "pleasant experience." The things that satisfy E3.5's justifying condition are called "pleasant experiences" because they are experiential states in which pleasure is taken. They are objects of pleasure at times, and, significantly, those times include any occasion on which E3.5's justifying condition is met. On all such occasions the other relevant sense of "pleasant experience" also applies to something in the vicinity--the prizings themselves can be called that, too. But they are not what is then satisfying E3.5. Doing that is not even a "reliable sign" that something is a "pleasant experience" in this latter sense (nor that it is anything else credibly taken as intrinsically good).

#### D. Removing Remaining Apparent Disanalogies

We can move closer than E3.5 gets to the justifying condition in E3.2 without using something that involves thinking of *p* being intrinsically good, thus strengthening the analogy to the case of the closed door. We can try:

E3.6 *p* being contemplatively approved by someone out of regard for some believed feature of *p* or other tends to confirm that *p* is intrinsically good.

E3.6 is no better than E3.5, though. Smith saying "Please" in the course of making a request can be contemplatively approved as good manners; Robinson using dynamite as an example of an explosive can be contemplatively approved as a good example. The only ways to modify E3.6 that seem better than it bring back the thought of *p*'s intrinsic goodness as the basis for the approval. So the analogy to the epistemic irrelevance of pleasure in (c) seems to be further borne out.

A last putative disanalogy between pleasure in the intrinsic goodness evidence principle E3.2 and pleasure in the door closure principle (c) should be discussed. There is a lingering temptation to find E3.2 better than E3.3 (while there is no such temptation to prefer (c) to (e)). I think that intuition can be explained in a way that gives pleasure no evidential role in learning of intrinsic goodness. Contemplative approval out of regard for intrinsic goodness is closely correlated with focusing upon and attending to the proposition that the thing in question is intrinsically good in search of intrinsic goodness-conferring features of it. This correlated concentration may provide

evidence for p's intrinsic goodness which is superior to that provided by simply thinking of p being intrinsically good in some casual fashion. But even if so, this is certainly not evidence from pleasure itself. We can replicate any such advantage of E3.2 with:

E3.7 S giving careful consideration to whether p is intrinsically good while taking normal precautions and believing p to be intrinsically good on the basis of such reflections tends to confirm that p is intrinsically good.

E3.7 makes no use of pleasure. Yet it achieves the focusing effect in question at least as well as E3.2. If we use E3.7 instead of E3.3 as the analog to (d) in the case of the closed door, the analogy seems complete. The conclusion that pleasure in E3.2 is epistemically superfluous now looks inescapable.



### Notes to Chapter III

<sup>1</sup>Franz Brentano, The Origin of Our Knowledge of Right and Wrong (New York: Humanities Press, 1969), E. Schneewind and Chisholm, trans., p. 22.

<sup>2</sup>Charles Baylis, "Grading, Values and Choice," Mind 67 (1958), pp. 494-495.

<sup>3</sup>See the Introduction above, at the beginning of section C, and Note 12.

<sup>4</sup>See the Brentano citations in sections A, B and C of Chapter II above.

<sup>5</sup>Brentano, p. 22.

<sup>6</sup>See Chapter II above, sections C and D.

<sup>7</sup>It might be thought that, if we need evidence for intrinsic goodness before we can see the requirement, then the requirement could not be self-evident. But not so. If the intrinsic goodness is self-evident by D2.8, then so might be the requirement as a whole, even though it depends on something other than itself (i.e., its entailment that p is intrinsically good) for evidence of p's intrinsic goodness.

<sup>8</sup>See section D below for a modified version of E3.2 that may be more reasonably held to be epistemically equivalent to E3.3.

<sup>9</sup>Baylis, p. 493.

<sup>10</sup>Ibid., pp. 494-495.

<sup>11</sup>Ibid., p. 495

<sup>12</sup>See Chapter II, D2.2 and Note 9.

## C H A P T E R   I V

### PLEASURE AND THE POSSESSION OF INTRINSIC GOODNESS

This chapter has two major parts. In the first, I discuss in some detail problems that arise in formulating a clear and thorough hedonistic theory. That part naturally divides into (a) problems in identifying the hedonists' basic bearers of value (sections A and C below) and (b) problems in distributing the proper values to indefinitely complex states of affairs on the basis of any given identification of basics (section B below). The second main section of the chapter concerns selected objections to any version of hedonism. The first objection (section D below) is an argument by Brentano for the claim that pleasures cannot be the only intrinsic goods. The other objections considered are Broad's contention that malice shows not all pleasures are intrinsically good (section E below), and Moore's beauty objection (section F below).

#### A. First Attempts to Formulate Hedonism

What is hedonism? That seems easy to answer. We can use Moore's succinct formulation: "Pleasure alone is [intrinsically] good."<sup>1</sup> But which doctrine is that? It seems to say that there is one item--pleasure--which is the sole thing having positive intrinsic value. Surely, though, hedonists wish to hold that any of the many pleasing experiences is intrinsically good. That suggests the statement: All and only pleasant experiences are intrinsically good.<sup>2</sup> That is not quite right either. For many persons, tasting an ice-cream sundae is a

pleasant experience. Yet hedonists do not attribute intrinsic value to events of tasting. They hold that the pleasure of the taste is the intrinsic good in the experience, not the taste that is its object. But still, every case of taking pleasure is to count, not just the pleasure taking relation itself. So instances of taking pleasure--which we shall understand here to be states of affairs--are good candidates for the hedonist's bearers of intrinsic goodness. We can put the view by saying: All and only pleasure-states are intrinsically good. But we must not leave the formulation that vague. We have to attach a clear meaning to "pleasure-states." (The "hedonism" in question is not to be confused with an implausible doctrine about America's sun-and-fun spots--Florida, California, and Hawaii.) Here is a proposal:

D4.1 p is a pleasure-state = df necessarily, if p obtains, then someone is pleased.

D4.1 captures all the episodes that a hedonist wants, and a rich assortment of others. For now,<sup>3</sup> we can assume that hedonists would be after states like:

(a) Jones being pleased,

or

(b) Jones being pleased about Smith's good fortune,

or perhaps

(c) Jones being pleased to intensity 5 for 4 seconds at quality level 6 about Smith's good fortune.

D4.1, though, also embraces states such as:

- (d) Jones being pleased and Smith being displeased,
  - (e) Jones being pleased during a bowling escapade,
- and

(f) Someone being pleased.

(d) is no hedonist's idea of an intrinsic good; (e) and (f) might be counted intrinsically good by a hedonist, but only by deriving their intrinsic value from the fundamental bearers of hedonistic value that we now seek to capture, the "pleasure-states."

A better attempt to select out the right states is:

D4.2 p is a pleasure-state = df (i) possibly, there is exactly one x such that necessarily, p obtains iff (a) x is pleased, and (b) necessarily, for all q, if q strictly implies that x is pleased, then q implies p. (In other words, p implies a certain person to be pleased, and anything that implies that person to be pleased implies p.)

One counterexample to D4.2 is:

(g) Jones being pleased or 7 plus 5 equaling 75.

This objection seems surmountable by appeal to Chisholm's concept of involvement.<sup>4</sup> Intuitively, what we want here is the narrowest state of affairs attributing pleasure to Jones. That seems to be the one which, unlike (g), is involved in all the rest that make that attribution:

D4.3 p is a pleasure-state = df. Possibly, there is exactly one x such that necessarily, p obtains iff (a) x is pleased, and (b) for all q, if q strictly implies that x is pleased, then q involves p.



D4.3 achieves its intended objective only if a certain fairly dubious metaphysical proposition is true. Perhaps there are "individual essences" (e is an individual essence of x iff necessarily, x alone insatiates e). Any entity would have many of these. For example, Jones would have one we can call "iel" and the property of being the iel and such that there might be wolves (or, in the unlikely event that iel is just that property, being the iel or not such that there might be wolves). Now, for any individual essence of Jones, iej, necessarily, if Jones is pleased, then the iej is pleased. Thus if there is a single proposition involved in each that implies him to be pleased, then Jones must have a "core essence": c is a core essence of x iff (i) c is an individual essence of x and (ii) necessarily, if there are S, p, and F, such that S thinks of p and p strictly implies x to have F, then S grasps c. That each individual has a core essence is the dubious metaphysical proposition implied by D4.3. For if not, then there are two propositions such that each attributes being pleasing to the one who has what is in fact an individual essence of Jones, while neither involves the other.

Should we press on in our search for a rigorous account of "pleasure-states"? It might seem a misguided project. Consider this complaint: "Suppose you tell us just which are the 'pleasure-states.' What good will that be in formulating hedonism? Presumably, we will be offered:

(H) p is intrinsically good iff p is a pleasure-state.

The trend exemplified in D4.1-D4.3 tells us that this will be a theory of nearly nugatory interest. Think of what we want a theory of intrinsic value for. At least, we want it to provide a value rating of alternatives for the sake of a thorough formulation of utilitarianism. But any half-decent formulation of such a theory--Feldman's MO, Bergstrom's T4--requires evaluating such things as the causal consequences of whole courses of action or whole possible worlds. It rules all such things to be worthless!"

That vigorously advanced objection is not entirely lacking in rational force. But it is not at all clear that the hedonist must assign intrinsic goodness to complex states of affairs in order to rank them on hedonistic principles. There might be a good way to explain the ranking in terms of pleasure-states "contained" in such wholes. This would be a complicated business, no doubt, in the case of some of the purposes we would like a value theory to serve. For example, we might want a hedonistic theory of which actions are, on a given agent's evidence, in that agent's best interest. That would require appeal to the hedonistic value of what the agent's evidence attests to being the outcome of what that evidence purports to be the agent's available acts. Such outcomes will ordinarily include ineliminable disjunctions such as:

(h) Jones being amused or bemused.

Since the evidence in some cases does not determine which of (h)'s disjuncts will obtain, (h) itself must be evaluated. Unlike conjunctions involving pleasure-states, there is no manifest promising sense to be made of such states "containing" pleasure-states that "sum" to the

hedonistic value of the whole. How, then, can hedonism be expanded to cover that sort of territory?

Recent philosophical developments give grounds for hope about the pleasure-state approach to these things. Warren Quinn initiated the work. His "Theories of Intrinsic Value" offers a means to assign the "right" intrinsic value to complex states on the basis of pre-determining evaluations of "basic" or "atomic" ones.<sup>5</sup> Edward Oldfield provided a much improved version of Quinn's theory in "An Approach to a Theory of Intrinsic Value."<sup>6</sup> That work seems to supply just what is needed here--a way to generate values for indefinitely complex states out of the values for pleasure-states. As will be seen though, even Oldfield's development of the theory is not satisfactory as it stands. The order of business here, then, is to examine that work, see whether the mechanism for assigning value can be perfected, then see whether there are ways to explicate "pleasure-state" that allow various versions of hedonism to "plug into" the resulting machinery.

### B. The Quinn-Oldfield Approach

I think I cannot improve upon Oldfield's summary of Quinn's view: [H]is theory is based on a primitive predicate 'basic proposition.' No basic proposition entails any other basic proposition. Basic propositions come in families, that is, sets whose members are mutually exclusive but jointly exhaustive. A proposition is not indifferent to a family of basics iff it is compatible with some members and not with others. On the basis of the notion of indifference,

value constituents are assigned to every proposition. A value constituent of a proposition  $p$  is a conjunction of propositions which first, is compatible with  $p$  and second, contains one member from every family to which  $p$  is not indifferent. Propositions which have more than one value constituent are indeterminate. All other propositions are determinate. The intrinsic value of a determinate  $p$  is calculated at a world. If  $p$  is not true at  $w$ , its intrinsic value at  $w = 0$ . If  $p$  is true at  $w$ , its intrinsic value at  $w$  is equal to the sum of the intrinsic values of the basic propositions which are conjuncts of that value constituent of  $p$  which is true at  $w$  . . .<sup>7</sup>

Quinn's characterization of which are to be "basic" propositions is promising: A basic proposition ". . . locates a specific sentient individual along an evaluatively relevant dimension such as happiness, virtue, wisdom, etc."<sup>8</sup> That sounds like just what we have been seeking as "pleasure-states." We will leave aside the question of what these really are,<sup>9</sup> and we will follow Quinn and Oldfield in taking expressions such as "J5" and "J-8" to abbreviate basics of five units of positive and eight units of negative intrinsic value respectively.

How is Quinn's theory supposed to work? Consider:

(j) J-1 and K2.

First we must find the families to which (j) is not indifferent. Families, according to Quinn, are sets such as the J's = [...J-2, J-1, J0, J1, J2, ..., J does not exist (this last member to ensure that the J's are exhaustive)]. A little thought enables us to see that (j) is indifferent to all families except the J's and the K's. We are to take it

that the one member of the J's with which J-1 is compatible is J-1. (Recall that families have mutually exclusive members.) Thus it can be seen that the value constituent of (j) is (j) itself, that (j) is determinate, and that  $iv(j)=1$ , which is intuitively right.

Things quickly go very bad, however. Consider:

(k) J1 or K2.

(l) Someone 2.

(k) and (l) are indifferent to every family. (k) is compatible with all the J's, for example, since it can be true when just its second disjunct is. So (k) and (l) are determinate. The sum of the (vacuous) conjunction of their value constituents is thus 0--nowhere near their intuitive intrinsic value.

Oldfield points out these defects and develops a theory that is not subject to them. He begins with the basics, but he also appeals to the much broader class of the basic\*s:

D4.4 (i) If p is basic, p is basic\*;

(ii) if p is basic\*, p's negation is basic\*;

(iii) if S is a set of basic\* propositions and C, the conjunction of the members of S is contingent, then C is basic\*;

(iv) nothing else is basic\*.<sup>10</sup>

Oldfield then constructs a notion of "irrelevance," patterned after Quinn's indifference:

D4.5 p is irrelevant = df. P is contingent and (logically) independent of every basic\*.<sup>11</sup>



The irrelevant propositions are supposed to be those that are without evaluative content.<sup>12</sup> "The next task is to say what is the class of basic propositions in terms of which the intrinsic value of a proposition should be calculated at a world."<sup>13</sup> Here we have the proposal:

D4.6 b is a minimal set for p in w ("Min(b,p,w") = df (i) b is a set of basics true in w; and (ii) there is an irrelevant, c, such that (a) c is true in w, and (b) the conjunction of c and the members of b entails p; and (iii) there is no proper subset of b and an irrelevant which is true in w which satisfy (ii).<sup>14</sup>

Let us look at D4.6 in operation. With respect to (k) and (l) things are simple enough. Consider a world, w<sub>1</sub>, where K<sub>2</sub> and the sky is blue, and M<sub>2</sub> for good measure. What is Min(b,(k),w)? Well, [K<sub>2</sub>] satisfies D3(i). The sky being blue can be seen to be an irrelevant. Its conjunction with K<sub>2</sub> entails (k), and no proper subset of [k<sub>2</sub>] conjoined to any irrelevant does so. So Min ([K<sub>2</sub>],(k),w<sub>1</sub>). And by parallel reasoning, Min([K<sub>2</sub>],(l),w<sub>1</sub>).

Now let us look at another sort of case:

(m) (K<sub>2</sub> and the sky is blue) or (J<sub>1</sub> and grass is green)

In this case, the true irrelevant for D4.6(ii) must be chosen with care. For no conjunction of basics alone entails (m). In w<sub>1</sub> we need the sky being blue for c. But we have it, so (m) seems to go through all right.

The principle of evaluation offered is:

(\*\*) For any world, w, and any proposition, p, true in w, the intrinsic value of p in w is equal to the sum of the

intrinsic values of the members of the union of all the minimal sets for  $p$  in  $w$ .<sup>15</sup>

(\*\*) has its peculiarities. As a world,  $w_2$ , where  $J_1$  and  $K_2$ ,  $iv(k)w_2 = 3$ . This is because both  $[K_2]$  and  $[J_1]$  are minimal sets for  $(k)$  in  $w_2$ , and, following (\*\*), we sum the values of the members of their union to get  $iv(k)w_2$ . This result seems at variance with the powerful intuition that the intrinsic value of a state is a value that the state guarantees in any world where it occurs--a value it has necessarily. Oldfield may not share this intuition.<sup>16</sup> In any event, he offers a principle for "absolute" (i.e., non-world varying) intrinsic value that can be generalized to:

AIV The absolute intrinsic value of  $p$  =

- (i) if  $p$  has world relative values for both sides of 0, then 0;
- (ii) if not, and  $p$  has a world relative value as close to 0 and any that it has, then  $p$  has that value;
- (iii) otherwise undefined.<sup>17</sup>

AIV does yield  $iv(k)=1$ . Unfortunately, though, it determines

(n) Tom and Tom alone 1

to have 0 intrinsic value. Consider  $w_3$  where (n) is true and nine other people -1. What is (n)'s minimal set in  $w_3$ ? It must be the set containing the basic for each of the ten people in  $w_3$ , and the (basic) negative existentials ("NE's") for each other possible person. Only that set, together with an irrelevant (any irrelevant truth of  $w_3$ ), fits the requirements on  $b$  in D4.6. (Note that nothing implying Tom's uniqueness

is irrelevant.) Thus it appears that, by (\*\*),  $iv(n)w_3 = -8$ . Since (n) will also have positive values, by AIV we get the unfortunate result:  $iv(n) = 0$ .

Concerning the intrinsic value assignment of -8 to (n) in  $w_3$ , Oldfield offers an ingenious argument (attributed to Jon Ruttenburg<sup>18</sup>) in defense of it. It assumes the principle that the world relative intrinsic value of necessary equivalents must be the same. Lacking the concept of (varying) world relative intrinsic value, I do not know how to assess that assumption. Anyway, were the argument sound, it would move me to either alter AIV in order to salvage the intuitive (absolute) intrinsic value for (n) of 1, or revise things earlier on to avoid such relative values as -8 for any equivalent of (n). Since I think that acceptable (absolute) intrinsic values can be obtained only when we revise both some preliminary material and AIV, that is what I shall do.

First a preliminary difficulty for (\*\*). Consider:

(o) M10 and Alfred being medium-sized.

There is no minimal set for (o) in any world. That is because (o)'s second conjunct is not irrelevant. It is incompatible with the NE for Alfred. Moreover, there is no set of basics which, together with an irrelevant, entails that conjunct. (Initially likely candidates for the irrelevant typically are implied by some basic\*.) So (\*\*) seems to award (o) the value 0, which is clearly inappropriate.

Since I want to have (o) receive the (absolute) intrinsic value 10 by appeal to "minimal minimal sets," so to speak, I must get the Alfred basics out of some of the minimal sets for (o). I think the best way to

proceed is to render the second conjunct of (o) irrelevant by altering D4.5:

D4.5a p is irrelevant = df. p is contingent, and p is independent of every basic\* which is independent of the conjunction of every proposition to the effect that a certain possible individual exists.

D4.5a makes Alfred being medium-sized irrelevant. That negative existential basic\* to which it is not D4.5-irrelevant is not logically independent of the conjunction of existentials, since of course that conjunction implies it to be false.

What are we to do about things like (n)? I think we should begin by finding a way to adjust AIV to make (n) worth 1, make

(r) Exactly two people-5 or exactly three people-3 worth -9 (its disjunct-closest-to-0's value), and leave without absolute value states such as

(s) Someone being happy to some degree or other.<sup>19</sup>

Such ratings would appeal to one sensible set of intuitions about the intrinsic value of complex states.

I think there is a way to build upon Oldfield's work to accomplish that. Notice that there are worlds where the minimal set for (n) is just [T1, the NE for each other possible individual]. Notice further that T1 is a member of every minimal set for (n). But also notice that there is no basic which is a member of each minimal set for (r). Even in the case of (r), however, there will be various 5-apiece pairs and -3-apiece trios (filled out by NE's for all other individuals) that are

minimal sets for (r) and such that no subset of any of these is a minimal set for (r) anywhere. And those are the minimal sets with the right values in them. All this suggests that we should make use of this concept:

D4.7 s is a least minimal set for p = df (i) there is a w such that s is a minimal set for p in w; and (ii) no proper subset of s is a minimal set for p in any world.

Though [K-10, T1, the NE's] is a minimal set for (n) in some worlds, it is not least minimal set for (n). For the minimal set for (n) [T1, the NE's] is a proper subset of it. [J5, L5, the NE's] is a least minimal set for r, the only minimal sets for (r) other than itself whose membership it exhausts also contain some other basic. So far so good.

By appeal to D4.7 we gain the following notion of (absolute) intrinsic value:

AIV2 The absolute intrinsic value of p =

(i) if there is an n such that n = the sum of the values for a least minimal set for p which is such that no such sum for p is close to 0, then n;

(ii) otherwise undefined.

AIV2 yields in the case of (r) the (absolute) intrinsic value -9. To see this, note that least minimal sets for (r) include [J5, K5, the NE's], [L-3, M-3, N-3, the NE's], and [J-3, M-3, L-3, the NE's]. So no sum of the values of the members of (r)'s least minimals can be closer to 0 than -9.



This accords with a certain conservative intuition about how much value disjunctions have. But it is simple enough to accommodate the intuition that states like (r) are worthless because there is no value, positive or negative, that they guarantee. We make further adjustment in AIV:

AIV3 The absolute intrinsic value of p =

- (i) if there is an n such that n is positive (negative) and no positive (negative) sum of the values of the conjuncts of a least minimal set for p is closer to 0, and none is below (above) 0; then n;
- (ii) if, in the case of p, there are two sums of values that satisfy the positive and negative versions of AIV3(i) respectively, then 0;
- (iii) otherwise undefined.

AIV3 assigns 0 to (r) and its ilk, since having least minimals that straddle 0 brings AIV3(ii) into play. But it leaves e.g. (s) undefined. The reason concerns a problem with applying the "guaranteed value" intuition to such cases. It might be that for each world there is some smallest increment of pleasure which it is psychologically possible for the creatures of that world to experience. But it is not plausible that there is a necessary minimum. So the hedonistic world-relative intrinsic values for (s) are as close to 0 as can be, though never 0. This might inspire attributing 0 intrinsic value to (s), albeit with misgivings in light of the fact that (s) guarantees better than 0 (perhaps with consolation derived from there being nothing better

than 0 that (s) guarantees). But such an attribution is not even that satisfactory in the case of:

(t) Two people being pleased to some degree or other.

(t)'s values approach 0 as a limit, too. But there would seem to be twice as much hedonistic intrinsic good in (t). Such a judgment might derive from the lack of a full appreciation of the nature of infinite descending series. But the going gets very tough in the face of:

(u) Someone being pleased to some degree or other and someone else being twice as pleased.

(u)'s world relative values get just as low as (s)'s, despite its virtually asserting itself to be better from the hedonistic point of view. In light of such examples, some will prefer to leave such states without (absolute) intrinsic values, in the manner of AIV3.

I prefer to assign all those states the absolute intrinsic value 0. I believe that the misgivings in question finally do just rest upon the manifest scope ambiguity in "guaranteed value," and that the right reading is "the value the state guarantees," not "the state guarantees some value." To satisfy my preference here we must have a final formulation of AIV:

AIV4 The absolute intrinsic value of p =

(i) if there is an n such that n is positive (negative) and no positive (negative) sum of the values of the conjuncts of a least minimal set for p is closer to 0, and none is below (above) 0; then n;

- (ii) if, in the case of  $p$ , there are two sums of values that satisfy the positive and negative versions of AIV4(i) respectively, then 0;
- (iii) if there is an  $n$  such that the least minimals for  $p$  approach  $n$  as a limit, and no least minimal for  $p$  is nearer 0, and none is on the other side of 0; then  $n$ .

### C. Hedonism Formulated

If AIV4 determines the value of complex states, which are the hedonist's basics? We would like to accomplish two goals with our basics: First, to make AIV4 work right, conjunctions of basics, in conjunction with irrelevants, must imply every state having hedonistic value. Second, we want the assorted versions of hedonism--e.g., those that do, and those that do not, recognize "quality" differences among pleasures--to be stated readily within the scheme. The first goal first: As already mentioned, it may be that for each world there is a physically necessary least interval of time and intensity of pleasure. Maybe not. But in any case, it is not credible that there are certain minimal intervals that hold for all worlds. Thus a difficulty arises when we consider:

(v) Smith experiencing pleasure of 2 units intensity.

(v) is short for one of the following:

(v') Smith experiencing pleasure of two units of intensity now or then, for some contextually definite time),

(v'') Smith experiencing pleasure of two units intensity for some interval of time or other.

Since pleasure must take time,  $(v')$  and  $(v'')$  are in the same boat with "some intensity or other" states--either intrinsically worthless or confounding to the guiding intuition of guaranteed value.

By the same token, from the vantage point of "qualitative" hedonism, states such as

(w) Smith experiencing pleasure of 2 units intensity for 2 seconds also guarantee vanishingly little intrinsic goodness (except in the dubious circumstance that there is some least unit of "quality"). And we should not assume that hedonists can consistently recognize only intensity, duration, and "quality" as independent, evaluatively relevant, variable properties of pleasures. One way in which criticisms to the effect that hedonism values too much can be blunted is to factor in an admirability rating for pleasures. It can be said with *prima facie* consistency that e.g. detached, intellectual enjoyment of sadistic acts and of magnanimous deeds, of equal quantity and quality, differ so much in admirability that the former is intrinsically inferior to the latter. Perhaps still other parameters must be admitted; perhaps intensity and duration tell the whole story. How can we be sure that, in any case our basics will fit properly into the rest of the theory?

Well, if I had some version of hedonism to advocate, it would be sufficient to identify its basics. But an adequate defense of a particular version of hedonism is the topic of some other dissertation. Here, I am restricted to attempting to say which are the basics relative to any given version of hedonism.

Recall that in section A we left off our try at isolating the "pleasure-states" with the problem of the core essence (an individual essence of  $x$  which is thought of whenever anyone thinks of any individual essence of  $x$ , speaking roughly). I am now prepared to hold that core essences are worse than doubtful--they are lacking entirely. For we must acknowledge such individual essences of A. N. Whitehead as being Russell's co-author of Principia Mathematica in @ (where "@" names this world), and being the author of Process and Reality in @. And there is nothing peculiar to Whitehead that must be thought of when thinking of each. So Whitehead has no core essence.

Because of this consideration, I think that we must abandon the D4.1-D4.3 approach. In fact, the only way I see to turn brings us back into the region of the metaphysically dubious. I think we have to look to what Herbert Heidelberger calls "singular propositions." The doctrine of singular propositions is the view that, for every property (relation) and object(s), there is a proposition--a singular one--which is believed just in case the property (relation) is attributed de re to the objects in a given order. If there are such propositions, some of them fill the bill as pleasure-states. But which ones?

As we have seen, each basic for a version of hedonism has to spell out how the pleasure rates with respect to each factor that the version deems a contributor to intrinsic value. Otherwise the state does not imply the presence of any positive quantity of intrinsic value, by the standard of the version in question. Which factors can versions of hedonism appeal to? Since each is to give us quantities of intrinsic



value that pleasures introduce, each must identify magnitudes that pertain in some way to pleasure, and tell us what function of those magnitudes yields the intrinsic value of the pleasure. For example, classical "quantitative" hedonism as I understand it is the view that the intrinsic value of an instance of someone taking pleasure in something is the product of the intensity and the duration of the pleasure.

This example lands us in the middle of several controversies. They might repay detailed discussion, but it does not belong here. Yet if only one variety of hedonism is coherent, much of this discussion has been superfluous. I think that is not the situation; I shall try to say just enough here about the controversies to make that opinion credible.

The "intensity" of a pleasure might be taken to refer to the felt vigor of the sensation or other mental event that gives rise to it. But I think a version of hedonism that assigns a significant independent contribution to intrinsic value by that variable is beyond belief. It overrates the fairly modest joys of jackhammering and underrates the considerable pleasure that can be got from the feeling of a soft warm breeze. Preferable is the view of C. D. Broad according to whom there is a simple, directly apprehended magnitude--hedonic tone--present in some degree whenever pleasure is taken.<sup>20</sup> When the pleasure is sensory, we speak of this as "how good it feels (tastes, smells, etc.)"; when the pleasure is not primarily sensory, we speak of "how much we like it." This quality seems the most reasonable referent for the "intensity" of the pleasure.

Some have doubted that there can be a way to measure this intensity. It seems that there might be, since we might establish a correlation of its degree with the strength of electrical current in some part of the brain or the like. But I see no reason to believe that the truth or coherence of an appeal to this intensity turns on its being measurable.

Mill appears to have thought that the value of a pleasure also depends upon an independent variable that he called "quality." Three questions: Is that thought consistent with the position that only "more pleasure" supplies more positive intrinsic value?<sup>21</sup> What feature of a pleasure is its "quality"? And is it still hedonism when differences of "quality" are brought in? It seems to me that the first question should concern us only if a certain sort of answer to the third is correct. For unless a view is a version of hedonism only if it affirms the biconditional in the first question, it does not matter to us what the answer to that question is. Similarly, only after we determine what would count as a hedonistic theory can we see what can be consistently taken as an intrinsic value determining "quality" difference.

What makes a value theory hedonistic? The following requirement is uncontroversial: All positive value is to be contributed by pleasure. That is too vague to get us far. What facts about a pleasure can be used? Intuitively speaking, the intensity of the pleasure is okay; the extent to which the society in which it is felt is just, is not okay. What is the difference? My hypothesis is that the only magnitudes that a distinctively hedonistic theory of intrinsic goodness can make relevant are experienced variables in pleasurable experiences. Somewhat more precisely:

D4.8 M is potentially hedonistically significant = df. There is a set, S, of properties, each of which is necessarily such that it is exemplified iff someone exemplifies it, is aware of it, and is pleased; and for some number, n, and duration, d, the degree of M is n iff someone exemplifies, and is aware of, each quality in S during d, and some quality or other in S throughout d.

I think that D4.8 sorts magnitudes appropriately. For example, in the case of intensity, S is the set containing just the quality which is some given degree of hedonic tone and i is a single moment. For duration, the quality is that of taking pleasure in something, and the degree, M, equals the duration of i. The property of being pleased in a society having degree of justice M fails D4.8, since degree of justice does not correlate with any experiential quality; society can be that just while no one is aware of it. The property of being pleased while having existed for M seconds also does not pass D4.8. No experiential qualities which are present just when a person is pleased are exemplified only when that person is a given age. I like those inclusions and exclusions, but I do not claim that D4.8 is the only reasonable proposal for potentially hedonistically relevant magnitudes.

If D4.8 is an acceptable account, what might be a magnitude that determines a pleasure's "quality"? Some passages in Utilitarianism<sup>22</sup> suggest that differences in objects of pleasure--what the pleasure is taken in--determine "quality" differences. "Intellectual" objects are esteemed better than "bodily" ones. It is not obvious what magnitude

is intended. A varying intellectual factor that D4.8 allows a hedonist to count relevant is the degree of concentration on the object of pleasure. How intently we are attending to what we are taking pleasure in is, I think, something we experience. And it could be taken to measure the degree of our intellectual involvement, and thus the intellectual "quality" of the pleasure. It is, at any rate, a magnitude in addition to intensity and duration that D4.8 allows a hedonist to take into account.<sup>23</sup> And finally, concerning our first "quality" question, if we say that intensity and duration together determine "how much" pleasure is had, then a hedonist can with consistency appeal to such things as degree of concentration in denying that the better a pleasure is, the more of it there must be.

As I see it, then, a full-blown version of hedonism must select its magnitudes from the properties that satisfy D4.8. And it must choose its basics from these singular propositions:

- D4.9 B is a potential basic bearer of hedonic intrinsic goodness =  
df. B is a singular proposition such that necessarily, S believes B iff there are degrees and potentially hedonistically significant magnitudes such that S attributes to someone the property of entering into the pleasure taking relation to those degrees of those magnitudes.

The magnitudes that are relevant for a version of hedonism are the ones it deems independently to affect the intrinsic value of any potential basic, i.e.:

D4.10 M1-Mn are the relevant magnitudes for version of hedonism V =  
 df (i) M1-Mn are potentially hedonistically significant; and  
 (i) according to V, for any potential basic hedonistic bearers  
 x and y, necessarily, if x is intrinsically better than y,  
 then at least one of M1-Mn has a greater value for x than for  
 y; and (iii) according to V, for each of M1-Mn, there are po-  
 tential basic hedonistic bearers x and y such that x is in-  
 trinsically better than y, and x equals y for every other mag-  
 nitude from M1-Mn.

Now we can identify the basics for a version of hedonism:

D4.11 B is a basic bearer of intrinsic goodness for version of  
hedonism V = df. B is a potential basic bearer of hedonic  
 value that attributes values to exactly the relevant  
 magnitudes for V.

Finally, a full version of hedonism must say which function of its  
 relevant magnitudes gives the intrinsic values of its basics:

D4.12 F is the intrinsic value function for version of hedonism V =  
 df. According to V, the function,  $F(x, M1, \dots, Mn, y)$ , is such  
 that if x is a basic for V attributing pleasure to someone to  
 degrees M1-Mn of the magnitudes deemed hedonistically relevant  
 by V, then y is the intrinsic value of x.

AIV4 will operate on the basics of a version and the values as-  
 signed by its basic intrinsic value function in yielding the intrinsic  
 values of the complex states of affairs. But it is not quite that sim-  
 ple to have a complete theory. We also have to be told which are the  
 intrinsically bad basics. Until now I have said nothing about pain



here. Traditionally, the doctrine that pleasure is the only intrinsic good has been paired with the view that pain is the only intrinsic evil. Yet it is obvious that the former does not logically imply or analytically contain the latter (nor vice versa). Thus it is intriguing to speculate about the justification for this pairing.<sup>24</sup> Intrinsic goodness is our subject, however. So here I confine myself to noting that utilization of the modified Quinn-Oldfield machinery for assigning value to complex states awaits an account of the basics bearing intrinsic disvalue.<sup>25</sup>

#### D. Brentano's Argument Against Hedonism

As I say, a thorough assessment of the many reasonable versions of hedonism is a thesis-size undertaking by itself. Three arguments affecting the truth of any hedonism will be considered below. The first, by Brentano, is aimed at showing that if pleasure is intrinsically good, then so is something else. I take it up because I believe it is interesting and deserves to be laid out more fully than it is in Brentano's formulation. The second argument we will consider is C. D. Broad's objection to the effect that the evil of malice shows that all forms of hedonism count too much as intrinsically good. And the third case is Moore's example intended to show that hedonism, by leaving out beauty, counts too little as intrinsically good. An accounting of the replies available to the hedonist against the latter two objections goes a long way toward exhibiting hedonism's strengths.

Brentano's objection is modestly complicated. It is stated in a compressed fashion, which must be quoted in full:

To feel pleasure or delight is an emotional act, a taking pleasure or a loving; it always has an object, is necessarily a pleasure in something which we perceive or imagine, have an idea of. For example, sensual pleasure has a certain localized sense quality as its object. Now if nothing other than pleasure could be loved, this would mean that every act of loving had an act of loving as an object; but the beloved act of loving would have in turn to be directed upon an act of loving, and so forth ad infinitum. No; in order for pleasure to exist at all something other than pleasure must be capable of being loved.

But it follows further that pleasure is not the only thing worthy of love. If it were, any pleasure would be pleasure taken in something unworthy of love and hence unworthy of being an object of pleasure. And in that case, pleasure could scarcely be worthy of love; the danger would be that nothing at all was worthy of love.

If pleasure is a good, there must also be other goods.<sup>26</sup>

The best way I see to proceed is to set up my reconstruction of the argument right away, then discuss that. It will, I believe, be clear enough what my textual basis is. It will be helpful to begin with this abbreviation:

D8 S is the sequence (of propositional objects) for pleasure P =  
 df. S is an ordered set which is such that (i) the first member, p of S is a singular proposition consisting in an

individual taking pleasure in something, and (ii) if a member of S is a singular proposition consisting in an individual taking pleasure in something, then its successor is that in which that member states the individual to be taking pleasure.

### Brentano's argument.

- (1) For each pleasure, there is exactly one object in which that pleasure is taken.
- (2) No pleasure can have a sequence having no last member (i.e., "ad infinitum").
- (3) Each pleasure has a finitely long sequence. (1, 2)
- (4) (x) if x is worthy of love, then pleasure taken in x is unworthy of love.

Suppose (5) Only pleasures are worthy of love.

- (6) Each pleasure is either pleasure taken in a pleasure, or pleasure in something unworthy of love. (1, 4, 5)<sup>27</sup>
- (7) (x) if x is worthy of love, then x is a pleasure taken in a pleasure. (4, 5, 6)
- (8) (x) if x is a pleasure, then the next to last member of th sequence for x is unworthy of love. (3, 4, 7)
- (9) (x) if x is a pleasure, then each member of the sequence for x (including x itself) is unworthy of love. (repeated applications of 3, 4 and 7)
- (10) If only pleasures are worthy of love, then nothing is worthy of love. (7, 9, and CP 5-9)

[conclusion over]

(11) If pleasures are worthy of love, then so is something that is not a pleasure. (10)

The first inference is in need of comment. Notice first that it follows from (1) alone that there is a sequence for each pleasure. (2) says that pleasures having sequences all have finite ones. So (3) follows.

I believe that--(1)(2), hence (3)--is a sound argument. (1) is not immediately obvious to all. But apparent counter examples lose all force upon inspection. Certain sensory qualities happen to be spontaneously and powerfully pleasing. This creates some temptation to say that they themselves are pleasures, rather than items that are pleasurable taken. That is, it can seem that some gratifications are just cases of having a one place sensory quality, not relations of subjects to states. but I think that careful phenomenological scrutiny reveals that there are no "pure pleasure qualities." Instead, there are pleasing feelings. With some effort or imagination it is possible to think of having those very feelings while not taking pleasure in them. (For example, suppose they were known to be sure signs of one's imminent, ghastly demise.) So they need not be liked. We say they are "pleasures" because we like our having them, i.e., the feeling is the object of the pleasure-taking relation.

Perhaps (2) is debatable. I suppose that someone with divine cognitive powers could take pleasures having infinite sequences, but it seems best just to leave that consideration aside. (We could scrupulously add the antecedent "If nothing has divine powers, then" to

(2) and the conclusion.) The best way I see to justify (2) is to point out that a pleasure violating it would be not merely "infinitely involved," but also "infinitely involuted":<sup>28</sup>

D4.13 p is infinitely involved = df. Necessarily, whoever thinks of P thinks of infinitely many states of affairs.

D4.14 p is infinitely involuted = df. Each state of affairs involved in p is infinitely involved.

It seems to me that infinitely involuted states are necessarily mind-boggling. So I accept (2), and endorse the argument from (1) and (2) to (3).

Premise (4) is taken straight from the text (from the second cited paragraph). It is where the argument breaks down. With effort, we can read (4) to say merely:

(4a) (x) if taking pleasure in x (i.e., "love" of x) is not of the worthy sort, then pleasure taken in x is not worthy (i.e., "unworthy of love").

But of course hedonists can blithely and consistently accept the nearly empty (4a). (4a), (1), and (5) do not imply (6), nor is there another valid way to gain (11) from (1)-(4a). So (4a) is an inappropriate reading.

On the other hand, it is not difficult to give (4) a sufficiently powerful construal to gain validly an anti-hedonistic conclusion. First, though, we should be clear about what sort of conclusion we have. Even a hedonist accepting Brentano's views about the relation of goodness of various sorts to being worthy of love need not dispute (11).



After all, Brentano only claims that being worthy of love is being "good" in the "broadest" sense.<sup>29</sup> Hedonists can agree that things other than pleasures are "broadly" speaking "good," whether that means overall good or the disjunction of intrinsic and extrinsic goodness. The most efficient way to turn this argument into one addressed straightforwardly against hedonism is to suppose that Brentano meant to appeal to a special sort of "love" (call it "i-love") which is suited to exactly the intrinsic goods, i.e., they are worthy of it, and only they are.<sup>30</sup> Then we can suppose him to have been intending i-love throughout. Now to gain (6) we can turn to:

(4b) (x) if x is not worthy of i-love, then pleasure taken in x is not worthy of i-love.

And now (11) implies the distinctly anti-hedonistic:

(12) If a pleasure is intrinsically good, then so is some non-pleasure.

The implication of (12) by (1), (2) and (4b) is interesting. Let us pause and give it due reflection . . .

The trouble is that it is very hard to see why a hedonist, or anyone else, should accept (4b). (4b) is quite a severe doctrine. It implies that all pleasure in the intrinsically neutral--even when it is known to be overall good--lacks intrinsic goodness. Yet a pleasure taken in the discovery of a wonder drug seems as good as any. (4b) is not plausible on its own, and Brentano does not defend it.

Does (4) have a more credible reading which also has sufficient logical power for the argument at hand? Here is a fairly plausible claim, not as trivial as (4a):

(4c) (x) if x is intrinsically bad, then pleasure taken in x is not worthy of i-love.

Brentano has asserted that his epistemic test for intrinsic value vouches for (4c).<sup>31</sup> Of course the hedonist should insist upon seeing both that claim and the criterion verified, since the denial of hedonism is implied by (4c) together with the uncontroversial:

(13) Some pleasures have been taken in intrinsic evils.

It clearly follows that not all pleasures are intrinsically good. But what is most important in the present context is that (4c) does not support a valid inference to (11) via (1) and (2). This can be readily verified by noting that (1) and (2) only give us that some pleasure has an object which is not a pleasure, and thus, when (5) holds, not intrinsically good. That does not imply what is needed to make use of (4c)--that each pleasure has an intrinsic evil as an object when (5) holds. So even the rather dubious (4c) is too weak for the argument.

Other readings of (4) can be tried. None I have thought of accomplishes Brentano's purposes.

We can summarize the results of this section in this way:

Brentano can show that if pleasure is intrinsically good only when its object is intrinsically good, then pleasures are not the only intrinsic goods. But the antecedent to that that statement needs defense and seems indefensible.

### E. Broad's Malice Objection

C. D. Broad's objection to hedonism is simple and direct. He asks us to give careful consideration to a certain sort of pleasure. He thinks that reflection makes it obvious that this sort of pleasure is bad, not good:

Now consider the state of mind which is called 'malice.' Suppose I perceive or think of the undeserved misfortune of another with pleasure. Is it not perfectly plain that any cognition which has the relational property of being cognition of another's undeserved misfortune and the hedonic quality of pleasantness will be worse as the pleasure is more intense? No doubt malice is a state of mind which on the whole tends to increase human misery. But surely it is clear that we do not regard it as evil merely as a means. Even if we were quite sure all malice would be impotent, it seems clear to me we should condemn it as intrinsically bad.<sup>32</sup>

The only way I know of utterly to quash and refute an objection is to deduce a contradiction from it using only the safest inference rules. That cannot be done in this case. In fact, I think that not even "garden-variety" refutation is possible here. (I mean plain old refutation. I suppose "garden-variety" was once roughly co-extensive with "average quality" among flowers or vegetables. It is so much better now as to make the metaphor misleading.) Indeed, Broad's is an appeal to what is intuitively so. That would seem to put it beyond all reasoned criticism--there is no disputing about taste, as they say (usually in Latin). I suppose not. But there are things that can be said against

objections which are said to derive from immediate intuition. Sometimes what is purported to be intuitive is actually a badly drawn inference. Perhaps the claim that it is intuitive that one is not moving when standing still on solid ground is a good example of this. How are such claims to be combatted with rational force? The best that can be done, I think, is to point to an intuitive truth which is being confused with the asserted intuition, and show that the assertion is not justified by that truth. (In the motion example, it can be observed that what one knows is that one is not moving relative to what is standardly presupposed as the velocity-determining frame of reference--the surface of the earth. That of course leaves open the question of motion in relation to other things.) So our question here ought to be: Can the hedonist do that much in the case of the apparent evil of malice?

It is a start to suggest that we focus only upon the value we find in malice when it has no valuable consequences for good or ill. That will help us attend to its intrinsic value. But clearly Broad is aware of the importance of ignoring consequences, and yet he means to draw our attention to an intuition of negative value stemming from thought about the malicious frame of mind itself. And the hedonist should, I believe, grant that there is such an intuition.

But states which may be the ones Broad is counting as cases of malice are ones in which the hedonist can consistently find some evil. Broad speaks of pleasurable "cognition of another's undeserved misfortune." Perhaps, then, he means such states as:

(x) Robinson taking pleasure in what he knows to be Smith's undeserved misfortune.

Hedonists are not even committed to (x) being intrinsically good. (x) implies some pleasure for Robinson, which they must count good. But it also implies misfortune, which hedonists are logically free to count intrinsically evil. At least intensities and durations must also be specified, before the hedonist is committed to an intrinsic value status for (x).

But more importantly, states such as (x), even where all parameters relevant for the version of hedonism are specified, as in:

(x') Robinson taking pleasure for 2 seconds of 2 units intensity in what he knows to be Smith undergoing 2 seconds of 1 unit intensity pain

must not be just assumed to count as good. Again, the hedonist can with consistency rate  $n$  pain-intensity-times-duration units twice as bad as  $n$  pleasure-intensity-times-duration units is good.<sup>33</sup> So only against a complete intrinsic value theory, of which, say, quantitative hedonism is the "top half," can the intrinsic value of a state such as (x') be tested.

We should consider the other major sort of state which might be called "malice":

(y) Robinson taking pleasure at what he takes to be Smith's undeserved misfortune.

Even here, it would be consistent and not crazy for a hedonist to hold (6) to imply a certain evil disjunction--Robinson having a false belief



or Smith experiencing misfortune. That would not be part of an especially plausible value theory; it is not worth stressing. Better for the hedonist to observe that, since it must be malice we consider, a potentially misleading instance of believed evil must be part of the state considered. If we are inclined to believe that beliefs tend to be true, as most of us unreflectively are, then intuitions about what the state gives evidence for may have a pernicious effect on our attempts to assess its intrinsic value.

Another point the hedonist should make also concerns the difficulty in attending to just the state itself, apart from conditions irrelevant to its intrinsic value. It is hard, perhaps psychologically impossible, to try for long to evaluate a state without giving thought to circumstances that make it true. For (y), these would typically include much that a hedonist can consistently condemn. Where there is malice, there is usually some painful episode prompting the attitude. Also, the individual who feels the malice typically has qualms and misgivings about adopting such a feeling. Additionally, despair and desperation tend to be among the attitudes formed by the one who feels malice. Failing to feel well themselves, they attempt consolation by considering cases that appear to advance their relative status. And lastly, the malicious are typically also displeased by the other's good fortune, especially during the immediate temporal vicinity of feeling malice. Now none of these pains is implied by (y) and its ilk. But to repeat, there is a very strong tendency to consider (y) against a background of facts, and these things are part of the normal setting of

cases of malice. And they do, it seems to me, provide a credible origin for some of what enters into our negative evaluation of malice. It does tend to be part of something sordid. Yet none of these things helps to determine (y)'s intrinsic value. (These accompaniments also help to explain away Broad's view that malice gets intrinsically worse as the pleasure gets more intense. For these associated evils do tend to be worse as the pleasure increases.)

Hedonists can note that malice is evidence for another sort of negative quality--bad character. Perhaps the existence of a person having a bad character, a bad person, is intrinsically bad. Hedonists can count it so. If they do, (y) is in another way a mixed state--the pleasure is intrinsically good, but it tends to confirm the presence of an intrinsic evil. So there is also that way in which we can be misled when evaluating it.

Hedonists need not explain all evaluations in terms of what they say is intrinsically good or bad, though. This is obvious in the case of most good-of-a-kind ratings--good cutlery, good thievery, etc. It is perhaps less clear in the case of good personhood, but I do not see why even the utilitarian hedonist must equate person evaluations with having such-and-such a relation to states with such-and-such intrinsic values. This would be a plausible requirement if being a good person implied some close connection to performing morally right acts. No such implication seems to withstand scrutiny, however. So I believe that even the utilitarian hedonist can consistently maintain that there is an

intuition about badness of character prompted by (y) and the like, while not founding that upon intrinsic values.

There are two final distinctions which hedonists of every persuasion can take note of. It tends to be displeasing to think about cases of malice. And perhaps displeasure under the right conditions is evidence of intrinsic evil, or at least evidence against intrinsic goodness.<sup>34</sup> But the displeasure could be giving evidence for the evil of some of the above mentioned associated states and persons. After all, (y) does involve someone suffering undeserved misfortune, even though it does not imply that. And further, that very displeasure may be the source of the intuition that the malice is intrinsically bad.

Those seem to me to be the principal points that any hedonist can make which tend to undermine Broad's objection. Do states like (y) seem intrinsically bad on intuitive grounds after all these factors have been properly discounted? I think it quite doubtful that they do. The objection is not decisive.

#### F. Moore's Beauty Objection

G. E. Moore tries to persuade us that beauty (a beautiful state of affairs, we would say) has intrinsic goodness:

Let us imagine one world exceedingly beautiful. Imagine it as beautiful as you can; put into it whatever on this earth you most admire--mountains, rivers, the sea; trees and sunsets, stars and moon. Imagine all these combined in exquisite proportions, so that no one jars against the other, but each contributes to increase the

beauty of the whole. And then imagine the ugliest world you can possibly conceive. Imagine it simply one heap of filth, containing whatever is more disgusting to us for whatever reason, and the whole as far as it may be without one redeeming feature. . . . The one thing we are not entitled to imagine is that any human being ever has, or ever by any possibility can . . . see and enjoy the beauty of the one or hate the foulness of the other. Well, even so supposing them quite apart from any possible contemplation by any human being; still, is it irrational to hold that it is better that the beautiful world should exist than the one which is ugly?<sup>35</sup>

These remarks were in fact directed against Sidgwick's view of what it would be rational to aim to produce. But it is clear that the beautiful world is to be intrinsically better. Also, the comparison between a beautiful and an aesthetically (and otherwise) neutral world is better for our interests, since hedonists can attribute disvalue of various sorts to ugliness.

Again, I think the most that can be done in rational opposition to the case is to point out potentially misleading facets of the example. The main thing to draw attention to is an effect of that fact that we are actually contemplating the appearances of the pretty and neutral worlds when we assess them. It is more pleasing to think of the former than the latter. The power of this to distract us from the intrinsic evaluation of them can be easily underestimated.

An analogy might help. Think of the occasion on which you were told the most amusing anecdote that you have heard lately. Now imagine

a world in which loosely speaking nothing exists except a perfect simulation of the visual and auditory features of the story-teller by an unlikely sequence of aggregations of randomly fluctuating atoms--no conscious life is in that world. Compare it to a world that is so to speak a complete blank--an empty void. It is quite difficult not to like the former better, and this makes it remarkably tempting to say that it is a better place. But here, sober further reflection enables us to decline to reach the conclusion that the exemplification of the perceptual qualities of humorous anecdote recitation is good in itself. We realize that we are amused to think of the place, and that it contains an occurrence that powerful tends to provoke enjoyment, but that exhausts its value.

I believe that the same realization is available in the case of the beautiful and neutral worlds. At least it seems that at this point the champion of Moore's objection should say more if it is to persuade us. I do not know what Moore would do, nor do I know what more would do.



Notes to Chapter IV

<sup>1</sup>G. E. Moore, Principia Ethica (London: Cambridge University Press, 1903).

<sup>2</sup>R. Brandt, "Hedonism," in P. Edwards, ed., The Encyclopedia of Philosophy, (New York: Macmillan, 1967) p. 432.

<sup>3</sup>See below, section C.

<sup>4</sup>p involves q = df necessarily, whoever entertains p entertains q.

<sup>5</sup>W. Quinn, "Theories of Intrinsic Value," APQ 11 (1974) pp. 123-132.

<sup>6</sup>E. Oldfield, "An Approach to a Theory of Intrinsic Value," Philosophical Studies 3 (1977) pp. 233-249.

<sup>7</sup>Oldfield, pp. 233-234.

<sup>8</sup>Quinn, p. 131.

<sup>9</sup>In fact, Quinn's basics are not much like pleasure-states, since no two of them can be true for the same individual in the same world. Fortunately, Oldfield's modification of Quinn's theory has the effect of drooping this requirement of exclusivity.

<sup>10</sup>Oldfield, p. 237.

<sup>11</sup>Ibid.

<sup>12</sup>Ibid. But see the discussion of sentence (o) below.

<sup>13</sup>Ibid., p. 238

<sup>14</sup>Ibid., p. 239.

<sup>15</sup>Ibid., p. 240.

<sup>16</sup>This is not quite clear from the text. The paragraph on p. 240-241 suggests that he does not.

<sup>17</sup>Oldfield, p. 241. In fairness it must be acknowledged that AIV goes considerably beyond the text.

<sup>18</sup>Ibid.

<sup>19</sup>See the end of this section for the grounds for this.

<sup>20</sup>C. D. Broad, Five Types of Ethical Theory (London: K. Paul, 1920) pp. 229-231.

<sup>21</sup>A good paper on the topics of this question and "qualitative" hedonism in general is: N. Dahl, "Is Mill's Hedonism Inconsistent?", APQ Monograph (Holland: Reidel, 1973), pp. 33-54.

<sup>22</sup>J. S. Mill, Utilitarianism (Indianapolis: Bobbs-Merrill, 1957) pp. 11-15.

<sup>23</sup>We could have an admirability magnitude, too, e.g.--degree of conscious effort to like something believed good.

<sup>24</sup>If C. D. Broad is right that there is a single sort of quality--hedonic tone--that comes in pleasant and unpleasant forms (Broad, pp. 229-230), then perhaps the intuition is that hedonic tone is necessary and sufficient for intrinsic value.

<sup>25</sup>Just about everything said above can be adapted for traditional pleasure/pain value theories by making the corresponding claims about displeasure. That adaptability includes marking "qualitative" differences among pains of equal intensity and duration--something which Mill does not do.

<sup>26</sup>F. Brentano, "Hedonism," from The Foundation and Construction of Ethics (New York: Humanities Press, 1973) p. 164.

<sup>27</sup>This inference succeeds if "unworthy" is understood to mean "not worthy."

<sup>28</sup>On type-theoretical grounds we can exclude states such as would be expressed by:

(z) Jones taking pleasure in (z).

If there were such states, the inference from (3), (4), and (7) to (8) would fail.

<sup>29</sup>F. Brentano, The Origin of Our Knowledge of Right and Wrong, (New York: Humanities Press, 1969) p. 18.

<sup>30</sup>We have seen that the text (Brentano, *ibid.*) permits this interpretation. See Chapter II above, section C.

<sup>31</sup>Brentano, Origin, p. 22.

<sup>32</sup>Broad, p. 234.

<sup>33</sup>See Note 1. to the Introduction above, and this chapter, section C.

<sup>34</sup>See Chapter III above, where this sort of evidence is critically examined.

<sup>35</sup>Moore, p. 84.

## APPENDIX

### INSTRUMENTAL VALUE WITHOUT INTRINSIC VALUE?

There has been a surprising controversy among axiologists. The disputed question is whether something might be instrumentally valuable while nothing is intrinsically valuable. It is surprising that there has been controversy about this because it would seem to amount to the question whether a thing could possess worth solely as a means to a worthless end. It seems obvious that the answer to that question is no. Yet Paul Taylor,<sup>1</sup> Monroe Beardsley,<sup>2</sup> and Gilbert Harman<sup>3</sup> have denied that instrumental goodness implies intrinsic goodness. Charles Baylis<sup>4</sup> has affirmed the implication. I believe that there is no conflict of basic intuitions here. The purpose of this appendix is to resolve the issue. Various concepts of instrumental value must be distinguished. When we bring each into sharp focus, we gain a clear answer to the question whether it implies intrinsic goodness. In the case of all but one of the concepts of interest, the implication fails. For the one exception, an argument resembling the first cause argument for God's existence establishes the implication.

#### A. The Argument From Definition

Monroe Beardsley discusses an argument from definition for the conclusion that instrumental value implies intrinsic value. Here is the definition at stake:

D5.1 "x has instrumental<sub>1</sub> value" means "x is conducive to something that has intrinsic value,"

where "conducive" is meant to cover the relations of being a necessary part and being a means.<sup>5</sup> Beardsley comments:

Obviously, if we accept [D5.1] we are as committed to the existence of intrinsic value as we are to the existence of instrumental value.<sup>6</sup>

We can take this to be a way to affirm the relevant implication:

P5.1 Necessarily, if something has instrumental<sub>1</sub> value, then something intrinsically good occurs.<sup>7</sup>

Clearly, D5.1 guarantees the truth of P5.1. So clear is this, in fact, that to the best of my knowledge D5.1 has never been disputed by any philosopher. If D5.1 gave the concept of instrumental value at issue, there would be no issue. D5.1 has been thought to define a significant concept of instrumental value. G. E. Moore proposed it as an analysis of "good as a means."<sup>8</sup> I think it can be shown that that is not correct. D5.1 misses the concept with which Moore was concerned there, and that concept does raise a non-trivial question about whether there is an implication from instrumental to intrinsic value. To see that the analysis fails, notice that something might have intrinsic value and extrinsic disvalue in such proportion as to be neutral in total value. If an event is conducive to something like that and to nothing else of any sort of value, then it satisfies D1. But there is no familiar or helpful sense in which it is then "good as a means," or "of



instrumental value." So P5.1 uses a technical concept of no clear interest. Its proof by definition does not advance the issue.

In effect, Beardsley too denies the significance of P5.1. He observes that in the case of instrumental value

. . . it does not matter whether the value [of the thing that confers value upon its instrument] is intrinsic or instrumental.<sup>9</sup>

He maintains that the following definition "should be acceptable":

D5.2 "x has instrumental<sub>2</sub> value" means "x is conducive to something that has value."

Charles Baylis has proposed the result of replacing "that has value" in which D5.2 with "good" as his explanation of being instrumentally good.<sup>10</sup> Neither philosopher specifies what sort of "value" or "goodness" is meant. (We might find in the comments by Beardsley that were just cited the suggestion that he intended the concept of having either intrinsic or instrumental value. But that would render D5.2 circular. And the consideration just raised against D5.1 would show that D5.2 also gave a concept weaker than any intuitive notion of instrumental value.) Our complaint against D5.1 suggests that the right concept to use is what we call being "good on the whole" or "overall good." Roughly speaking, that is the concept of making ethically significant contributions that are positive on balance. In the case of that notion, plainly it "does not matter" whether the impact is positive because of intrinsic or instrumental factors. Also, I believe that it is what is usually intended when something is called "good" or "valuable," where that is not

short for "a good such-and-such." So we construe Beardsley and Baylis to be proposing:

D5.3 "x has instrumental<sub>3</sub> value" means "x is conducive to something overall good."

D5.3 captures what it is to be of some good as a means. And it allows the formulation of the non-trivial:

P5.3 Necessarily, if something has instrumental<sub>3</sub> value, then something intrinsically good occurs.

The argument from definition in this case clearly fails. P5.3 cannot be derived from D5.3 by logic alone. But is it true?

#### B. Beardsley, Taylor, and Harman on First Cause Arguments

P5.3 has been taken to be susceptible to another sort of proof. The overall goodness of a thing is entirely composed of the goodness it has on its own and the goodness it gains from contingent relations to other things. That makes something like the first cause argument look attractive. It is tempting to think that something could have instrumental<sub>3</sub> value only if somewhere along the line of things to which it bears a causal or part-to-whole connection is something that is overall good by virtue of having value on its own, i.e., intrinsic goodness. Here is Beardsley's sketch of such an argument:

'Instrumentally valuable' is a relational concept--x borrows its value from y or y confers its value upon x. If the value y confers is itself instrumental, so that it is merely passed along from z,

where does *z* get its value? In the last analysis, something must possess value in itself, or nothing can get any value.<sup>11</sup>

This seems to be the reasoning that Baylis relies on when he asserts that instrumental value does require intrinsic value.<sup>12</sup>

Beardsley's objection to this version of the argument is somewhat obscure. He says that it cannot be a "pure formal demonstration," since in view of the failure of the argument from definition it is not self-contradictory to assert the existence of instrumental value and deny the existence of intrinsic value.<sup>13</sup> It is not clear what the problem is supposed to be. Apparently, a difficulty is to be brought about by the fact that definition alone does not suffice to prove P5.3. But even in the argument's present rather metaphorical rendering, there is no suggestion that definition does all the work. Nearly enough, it proceeds by moving from the claim that gaining value instrumentally has the same logical structure as borrowing to the conclusion that such gain requires "owned" value, i.e., intrinsic goodness. These may be a priori obvious truths even if they do not follow directly from definitions. That seems to be the status of the proposition that red is a color, for example. So there does not seem to be a decisive objection here. We shall return to this argument below,<sup>14</sup> when we are better equipped to evaluate it.

Beardsley offers a revised version, aimed at avoiding his objection:

Premise 1 We know, or have good reason to believe, that some things are instrumentally valuable.

Premise 2 We could not know this unless we knew some things to be intrinsically valuable.

[Thus,] We know some things to be intrinsically valuable.<sup>15</sup>

Since knowledge of instrumental value may depend upon knowledge of intrinsic value even if there is no existential dependence, this argument offers no direct support for P5.3. But we should consider Beardsley's objection to it, for it serves to bring to our attention an important kind of instrumental value. He objects to Premise 2 by offering what he takes to be a case of knowledge of instrumental value without knowledge of intrinsic value. The example is that of good health. It is claimed that we see it to "retain its eligibility" as an instrumental good by witnessing a variety of situations--seeing that where present it did not "interfere with our pursuits," and where absent it "contributed to the rise of difficult problems and limited our capacity to solve them."<sup>16</sup>

We should agree that these things can be learned without evidence for the presence of any intrinsic good. But recall that the sense of "instrumental value" at stake is given by D5.3 (our reading of Beardsley's wording in D5.2). The mentioned features of good health do not give knowledge that it is conducive to an overall good. Knowing that requires evidence that freedom to pursue our pursuits or freedom from difficult problems is, or conduces to, an overall good. We need to be shown that those freedoms could be seen to be of positive overall worth without evidence about intrinsic goods. And that is not done. So Beardsley does not give a clear case of knowledge of instrumental value in the absence of knowledge of intrinsic value here.

There is a familiar notion of instrumental value that is readily seen to be exemplified by good health because of the features indicated in the example. We can bring out that notion in a rough way with:

D5.4 "x has instrumental<sub>4</sub> value" means "events consisting in instantiations of x sufficiently often make causal contributions to gaining something that is sought."<sup>17</sup>

This is the concept of being a generally effective means to some end, whether or not the end has any sort of value. Certainly we are often concerned with just this kind of instrumental value. It is clear that Beardsley's objection to Premise 2 would be successful if this concept were the one at stake. For it applies to good health because of just the sort of feature noted--utility in the long run. And plainly we can be justified in believing good health to have instrumental<sub>4</sub> value while we have no evidence about intrinsic goods. So perhaps this was what Beardsley was thinking of when he proposed the objection, rather than instrumental<sub>3</sub> value.

We should also note that properties can have instrumental<sub>4</sub> value while no intrinsic good obtains. For example, there are the gloomy possible situations where events which consist in instantiating the property of consulting an authority on painless suicide usually gain people something they seek in conditions entirely unrelieved by intrinsic goods. So it is plain that we have no instrumental to intrinsic value implication in the case of this concept.



P5.3 is not refuted by such examples, however. We have yet to see anything that undermines the intuition that instrumental<sub>3</sub> value must ultimately derive from intrinsic value.

Paul Taylor and Gilbert Harman have also attacked first cause arguments for the sort of implication in question. A careful examination of their work brings out new concepts of instrumental value. Taylor writes:

[T]here is nothing inconsistent in supposing . . . a world where all values [are] extrinsic, the value of one thing depending upon the value of another whose value in turn depends on the value of something else, ad infinitum . . .<sup>18</sup>

The sort of "extrinsic" value which Taylor calls "instrumental goodness" is given this characterization:

. . . creation, furtherance, strengthening, or increase of something intrinsically or extrinsically good, or destruction, hinderance, avoidance, weakening or decrease of something intrinsically disvaluable.<sup>19</sup>

The same considerations that showed fault in D5.1 show that this account is too weak to identify any familiar concept. An event might "further" an intrinsic good which is extrinsically so bad as to be neutral in overall value. There is no intuitive sense in which doing that is sufficient for having "instrumental value." Again, we do better by appealing to overall value. Simplifying while retaining what is crucial, we get:

D5.5 "x has instrumental<sub>5</sub> value" means "x causally contributes to bringing about something overall good or preventing something overall bad."

D5.5 does seem to give us a concept of interest. It is clear that we sometimes take things to have instrumental merit because of what they prevent as well as what they accomplish. The implication in question, then, is:

P5.5 Necessarily, if something has instrumental<sub>5</sub> value, then something intrinsically good occurs.

Taylor tries to describe a world that shows P5.5 to be false: In such a world, things would be judged as good solely as means to ends which were to be good solely as parts of wholes which were to be judged to be good solely as means to ends and so on indefinitely. . . . No one would do anything for its own sake, simply because he found personal enjoyment in it. It would be a world of 'practical people' who know how to get things done but had no reason for getting one thing done rather than another.<sup>20</sup>

Since Taylor holds that only "personal enjoyments" can be intrinsically good,<sup>21</sup> there would be nothing in these worlds that he counts as intrinsically good (assuming that there are no personal enjoyments, and not merely that no one does anything for the sake of them). But why should we say that these worlds include instrumental<sub>5</sub> goods? It must be because the things judged good are only judged good as means and parts. Yet of course that fact alone should not persuade us. We need to see that occurrences in such worlds actually further goods or hinder

evils, not merely that they are judged to do so. Taylor does not indicate how we might see this.

We need not follow Taylor by objecting to P5.5 on the ground that an endless sequence of instrumental goods is possible. For an event can get instrumental<sub>5</sub> value by just keeping something overall bad from coming to pass. For example, an inoculation that prevents the acquisition of a dreadful disease presumably has instrumental<sub>5</sub> value. This might occur in the unfortunate absence of any intrinsic good. So D5.5 provides us with another reading of "instrumental value" in which it is simple to see that this predicate might apply where nothing is intrinsically good. But we should not forget that we are still without any disproof of P5.3.

Like Taylor and Beardsley, Gilbert Harman denies the inconsistency of "... the notion of an infinite series A, B, C, etc. each member of which is desirable only because the next one is."<sup>22</sup> Harman tries to locate the plausibility and the defect in first cause arguments against such series, and he gives us yet another sense of "instrumental value."

The crucial notion in his reconstruction of the argument is that of basic intrinsic value. A "basic intrinsic value function" as Harman defines it is any such that the intrinsic value of a thing is the sum of all the results of applying the function to each entailment of the thing; and such that the total value of a thing, x, is the sum of all the results of applying the function to each thing, y, times the degree to which y is probable on x.<sup>23</sup> Here is the first cause argument

Harman constructs:<sup>24</sup> If there is a basic intrinsic value function, then nothing has positive overall value unless the result of applying the function something--the "basic intrinsic value" of the thing--is a positive number. That is clearly true since unless the function yields a positive number somewhere, each overall value will be the sum of the results of multiplying various degrees of probability times the basic intrinsic value zero. So, if there is a basic intrinsic value function, then:

P5.h Necessarily, if something is overall good, then something has positive basic intrinsic value.

Harman's objection to his first cause argument consists in questioning whether there is a basic intrinsic value function.<sup>25</sup> But without finding out whether there does exist such a function, we can ascertain that this argument does not prove the sort of implication in question here. Notice that P5.h states an implication from overall value to basic intrinsic goodness. There is an important link with P5.3, since if something has instrumental<sub>3</sub> value, then by definition something is overall good. But if there is a basic intrinsic value function, then it is clearly possible for there to be something with instrumental<sub>3</sub> value while no intrinsic good simpliciter obtains. For if there is a basic intrinsic value function, the overall value of, for example, administering a medicine will be positive if it increases the probability of some intrinsic good such as the patient experiencing the joy of recovery. This positive value does not depend upon the joy actually being felt. The existence of a basic intrinsic value function

requires the possibility of overall goods occurring where nothing intrinsically good happens. In those situations, events conducing to such overall goods would have instrumental<sub>3</sub> value in the absence of intrinsic goods. Thus no argument which relies upon the existence of a basic intrinsic value function can help to establish that instrumental<sub>3</sub> value necessitates intrinsic goodness, as P5.3 asserts.

This manner of acquiring overall goodness brings us to another concept of instrumental value:

D5.6 "x has instrumental<sub>6</sub> value" means "there is something which would be overall good if it were to occur and such that its probability on x is greater than its absolute probability."<sup>26</sup>

The corresponding implication principle is:

P5.6 Necessarily, if something has instrumental<sub>6</sub> value, then some intrinsically good occurs.

D5.6 does not seem to me to provide us with any familiar notion of instrumental value. In effect Harman contends that it does:

[W]e must say that S is desirable because of what it is likely to lead to. S may happen in various ways having different consequences. For example, your giving me ten dollars will have varying consequences depending on whether you give me the money as a gift, to hold for purposes of a bet, or to give to someone else. . . . S's being desirable because of what it would lead to must be taken as a special case of S's being desirable because of what it would make more likely.<sup>27</sup>



These considerations are not convincing. Even if your giving me ten dollars is very likely to have beneficial consequences, if the unfortunate fact is that it would have only grave results I see no sense in which it is "desirable" or has "instrumental value," as D5.6 has it. As I see it, instrumental value turns on only what would happen, not what is made likely.

In any event, we should have no trouble assessing P5.6. If we add to our most recent example that the patient's joy of recovery would be overall good, then the administration of the likely cure has instrumental value, since it raises the probability of something that would be overall good. This is so even if the patient fails to recover. No intrinsic good has been implied to obtain. So once again the precise concept of instrumental value in question is one concerning which it ought to be uncontroversial that intrinsic goodness is not implied.

We have seen that when Beardsley, Taylor and Herman criticize first cause argumentation for the sort of implication that we are considering they address themselves to notions of instrumental value which clearly do not require intrinsic goodness. But we have yet to find reason to doubt P5.3 or a genuine problem in the first cause argument for it that Beardsley formulates.

### C. A First Cause Argument Refuted

Can we disprove P5.3 directly? A refutation is at hand, if it is as easy to become overall good as Harman's work suggests. For suppose that a thing can become overall good by just making some would-be

overall good more probable. Then an event can be conducive to such a probability-raising overall good--thereby acquiring instrumental<sub>3</sub> value--while nothing intrinsically good actually occurs. That would do it. But many will find it doubtful that an event might come to be overall good so easily. They will take it as intuitively clear that a thing must be actually, rather than just probably, beneficial to be actually good. That is plausible. Other considerations show P5.3 to be false, however.

P5.3 is about instrumental<sub>3</sub> value. That is gained whenever an event results in something that has ethical influence which is positive on balance. So all we need to refute P5.3 is something which counts as an ethical benefit without introducing any intrinsic good. One way to be good is to prevent evil. The Dutch boy's act of plugging the dike was a valuable deed even if no intrinsic good came of it. It was beneficial simply because it kept a catastrophe from coming to pass. Another way to be beneficial is to make an improvement. Assuming the pain of disease to be an intrinsic evil, treatment that causes the pain to abate is valuable for doing that. There need not be any intrinsic good added to the situation for the treatment to be overall good. The diminution or removal of evil suffices. (Perhaps this too is preventing evil--the evil that would still be there if it were not for the improvement.) Intuitively, the idea is that it is equally creditable to move the world upward some amount on the overall value scale from what would have been, whether or not the change introduces something better than neutral.

The impact of these observations on P5.3 is clear. Possible situations exist in which evil is prevented or improvement is made, though no intrinsic good ever occurs. By being conducive to such prevention or improvement, events gain instrumental<sub>3</sub> value in the absence of intrinsic goods. P5.3 is false.

We are now in a position to assess the original argument for P5.3 which was sketched by Beardsley. Recall that the argument moves from the assumption that instrumental<sub>3</sub> value amounts to "borrowed" value to the conclusion that it implies some "possessor" of value--some intrinsic good. It is now plain that this borrowing metaphor is not apt. An event can have a valuable result, and thereby have instrumental<sub>3</sub> value, without taking out any "loan" from the intrinsic goodness of another thing. It is enough to keep evil from obtaining or to make things better. Of course this does not imply that an endless sequence of value dependents is possible. Rather, it shows that no such series is needed to have instrumental<sub>3</sub> value without intrinsic goodness.

#### D. A First Cause Argument Upheld

This refutation of P5.3 and the first cause argument for it may seem unsatisfactory. There remains an intuition that merely removing or diminishing evil is not good enough to be "truly" instrumentally good. What it really takes, it might be contended, is contributing something that is "positively worthwhile." I believe that there is a point to be made here, but it is not properly construed as a complaint against the objection to P5.3. Instead we should say that we have another concept

of goodness which is more stringent than that of being meritorious on balance--what we have called "overall goodness." Of course one narrower concept is that of being intrinsically good. That cannot be what we are looking for. D5.1 gives the most straightforward instrumental connection to an intrinsic good. And we have seen that it is in a way too weak and in a way too strong. It is too weak in that contributing to something merely intrinsically good is insufficient to insure having any intuitive sort of positive value, and too strong in that it does not matter for instrumental value whether the contributed good is made so by intrinsic or instrumental factors. Those reflections lead us to appeal to overall goodness. Now we are seeking something better. It would be best if it were a concept which does not manifestly either imply, or fail to imply, intrinsic goodness. Such a notion of instrumental value might be at the root of controversy about the implication.

A concept that has the qualities we want can be identified as follows. First, let us assume for the sake of simplicity that the contingent relation by which an event may acquire instrumental value is causal contribution, and that cases of improvement that go from bad to less-than-good can all be counted as cases of prevention. Here is a preliminary (inductive) definition:

D5.7 "q is a causal successor to p" means "either p causally contributes to q or a causal successor p causally contributes to q."

The causal successors to p, then, are just all the things in all the causal chains going through p that come after p. Now we identify a

restricted sort of overall goodness, one where prevention does not count:

D5.8 "p is positively worthwhile" means "the overall value of p is positive even after the prevention-induced value of p and p's causal successors has been discounted."

The notion is roughly that of being overall good either in virtue of being intrinsically good or in virtue of causally contributing to overall goods, where the caused goods themselves are taken to have value in only those two ways. Now for our final concept of instrumental value:

D5.9 "x has instrumental<sub>g</sub> value" means "x causally contributes to something positively worthwhile."

Here is our final implication principle:

P5.9 Necessarily, if something has instrumental<sub>g</sub> value, then something intrinsically good occurs.

Loosely speaking, D5.9 identifies the concept of being a means to an improvement that goes beyond the neutral point. It seems to conform to intuitions about instrumental value that might be the basis for a lingering admiration for P5.3. Perhaps, therefore, that is really misplaced admiration for P5.9. Yet it is not just obvious that P5.9 is true. After all, events can be positively worthwhile on purely causal grounds. Why could not there be an endless sequence of such events, never tied down to any intrinsic good?

Considerations are available that show P5.9 to be worthy of belief. In outline, the reasoning goes as follows. We consider the total reservoir of positive value that can be got by causal contribution in



each possible world. We argue that where the consequent of P5.9 is false, this repository of value is at best neutral. Thus in no such world can P5.5's antecedent be true.

Now for the details. We start by identifying a world's total stock of positive worth:

D5.10 The conjunction of valuable effects in  $w$  (CWE $w$ ) = the conjunction of every event occurring in  $w$  which is positively worthwhile and to which something causally contributes.

The CVE for a world,  $w$ , might be called "the first cause of instrumental value in  $w$ ," since each event of instrumental value in  $w$  helps to cause a conjunct of CWE $w$ . Now we consider an arbitrary world,  $w_1$ , where nothing intrinsically good occurs. A thesis:

(T) The instrumental value of every event in  $w_1$  is no higher than the positive value of CWE $w_1$ .

(T) is entirely credible. As just noted, an event in  $w_1$  has nowhere to go for instrumental value but to some conjunct of CWE $w_1$ . As a conjunction, CWE $w_1$  summarizes the value of all its conjuncts. So the positive worth of CWE $w_1$  is the highest that can be gained by causal contribution in  $w_1$ . Thus, to assess P5.5 we should ask: What is the highest possible value for CWE $w_1$ ? Well, positive worth is achieved only by intrinsic goodness or causal contribution. By hypothesis, there are no intrinsic goods in  $w_1$ , and so none conjoined in the CWE $w_1$ . If the CWE $w_1$  has any positively worthwhile effects of its own, then by definition they are among its conjuncts.<sup>28</sup> And in general, the positive value of a conjunction is determined in part by what its conjuncts help to cause.

Can CVEw1 be better than neutral because of its conjuncts' causal contributions? No. Any conjunct that causes something positively worthwhile causes something which is itself a conjunct of CVEw1, by definition of "CVEw." The vital consequence of this is that any positive worth that a conjunct has in virtue of what it brings about has its appropriate influence on the value of CVEw1 only if it is not counted in the cause's contribution to the value of the conjunction CVEw1, but rather counted as the contribution of the effect alone. To see this clearly, consider the simple case of the conjunction of an intrinsically neutral cause with its sole valuable effect. It is plain that the whole positive value those conjuncts add to that conjunction is the worth contributed by the effect. It would be a mistake also to add the causally derived value of the cause. But now notice that the same goes for each conjunct in the CVEw1--its role in determining the positive value of the CVEw1 is properly taken into account only if any worth it gets by being a cause is ignored, since that worth is fully represented by the contribution its effect makes as a conjunct. Only value as a cause and intrinsic value might give positive worth to these effects. Thus just their intrinsic value is counted toward the positive worth of the CVEw1, and by the assumption that w1 contains no intrinsic good we know that value to be zero at best.

That exhausts the ways that CVEw1 might have come to be positively worthwhile. All have been seen to fail. (T) enables us to infer from this that nothing has instrumental value in w1. Since w1 is merely an arbitrarily chosen world in which nothing intrinsically good occurs,

it follows that instrumentally value necessitates intrinsic goodness, i.e., that P5.9 is true.

We have now examined the full range of prevalent concepts of instrumental value. Most of them are readily shown not to imply intrinsic goodness, but we have just seen that one concept this question brings to mind does have that implication. Thus we have the congenial result that each side in the dispute can be construed to have a foundation in fact. Reflection seems to verify that every consideration for or against the sort of implication at stake in fact pertains to one or another of the notions discussed above. If so, and if the above reasoning has been successful, then the controversy has been resolved.

Notes to Appendix

<sup>1</sup>Paul Taylor, Normal Discourse (Englewood Cliffs: Prentice-Hall, 1961) p. 26.

<sup>2</sup>Monroe Beardsley, "Intrinsic Value," Philosophy and Phenomenological Research 25 (1965) pp. 4-8

<sup>3</sup>Gilbert Harman, "Toward a Theory of Intrinsic Value," Journal of Philosophy 64 (1967), p. 800.

<sup>4</sup>Charles Baylis, "Grading, Values and Choice," Mind 67 (1958), p. 490.

<sup>5</sup>Beardsley, op. cit., p. 4.

<sup>6</sup>Ibid., p. 5.

<sup>7</sup>For the sake of uniformity and definiteness we shall assume here that the bearers of instrumental and intrinsic value are things that occur--events. (One exception is required to accommodate a citation. See below, section B, and note 17.) Since the typical instrumental relation is causation, this is the most natural assumption.

<sup>8</sup>G. E. Moore, Principia Ethica (London: Cambridge University Press, 1903), p. 21. It should be made clear at this point that our topic concerns concepts of having instrumental value, not concepts of being instrumentally good. The distinction is that the former are concepts of being of some good from the instrumental point of view, the latter are concepts of making instrumental contributions that are on balance good. Moore is not explicit about which variety he is after with his account of "good as a means," but brief consideration of it shows that it is not at all suited to identifying a concept of the latter sort. D5.1 just looks at one contribution of value, not the on-balance value of all contributions. (I argue just below that D5.1 does not even suffice to say what it is to be of some good as a means.)

We will be concerned below with concepts of having some positive instrumental value of some sort. As in the case of Moore, philosophers who have done previous work on the topic are best interpreted to be appealing to those concepts. And more importantly, there is no consideration that distinguishes the plausibility of a some-good-instrumentally to intrinsic good implication from that of its nearest on-balance-good-instrumentally to intrinsic good counterpart. So this difference does not affect the issue.

<sup>9</sup>Beardsley, op. cit., p. 5.

<sup>10</sup>Baylis, op. it., p. 490.

<sup>11</sup>Beardsley, op. cit., p. 6.

<sup>12</sup>Baylis, op. cit., pp. 488, 490.

<sup>13</sup>Beardsley, op. cit., p. 7.

<sup>14</sup>See section C below.

<sup>15</sup>Beardsley, op. cit., p. 7.

<sup>16</sup>Ibid., p. 8.

<sup>17</sup>"Good health" seems to refer to a property, not an event. D5.4 is intended to express the simplest intuitive instrumental link between such a property and a goal.

<sup>18</sup>Taylor, op. cit., p. 26.

<sup>19</sup>Ibid., p. 28.

<sup>20</sup>Ibid., p. 32.

<sup>21</sup>Ibid., p. 23.

<sup>22</sup>Harman, op. cit., p. 800.

<sup>23</sup>Ibid., p. 801.

<sup>24</sup>Harman is interested in an argument for the conclusion that something is intrinsically good. Since we are concerned with whether instrumental value of various sorts implies intrinsic goodness, Harman's argument has been adjusted to conclude with P5.4--the conditional that comes closest to what we seek.

<sup>25</sup>Harman, op. cit., pp. 800-804.

<sup>26</sup>Ibid., pp. 796-798. Harman does not formulate D5.6, but some such concept is clearly suggested there.

<sup>27</sup>Ibid., p. 797.

<sup>28</sup>The principle that nothing can cause one of its own conjuncts is intuitively attractive. If it is true, then CVEw1 can have no positively worthwhile effects. The important thing here is that whether this principle is true or not, only the CVEw1's conjuncts might give it causally derived worth.



## BIBLIOGRAPHY

- Bailey, J. "On Intrinsic Value." Philosophia 10 (1980): 1-8.
- Baylis, C. "Grading, Values, and Choice." Mind 67, (1958): 488-504.
- Beardsley, M. "Intrinsic Value." Philosophy and Phenomenological Research 25, (1965): 1-18.
- Brandt, R. "Hedonism," in P. Edwards (ed.) The Encyclopedia of Philosophy. New York: Macmillan, 1967, pp. 432-4367.
- Brentano, F. The Origin of Our Knowledge of Right and Wrong. New York: Humanities Press, 1969.
- \_\_\_\_\_. The Foundation and Construction of Ethics. New York: Humanities Press, 1973.
- Broad, C. Five Types of Ethical Theory. London: K. Paul, 1920.
- Chisholm, R. "Brentano's Theory of Correct and Incorrect Emotions." Revue Internationale De Philosophie 20 (1966): 389-415.
- \_\_\_\_\_. "The Ethics of Requirement." American Philosophical Quarterly, 1 (1964): 147-153.
- Conee, E. "Propositional Justification." Philosophical Studies 38 (1980): 65-68.
- Dahl, N. "Is Mill's Hedonism Consistent?" in American Philosophical Quarterly Monograph 7, Holland: Reidel, 1973, pp. 33-54.
- Feldman, F. "Final Comments on the Analyses of Warranting." Synthese 38 (1978): 465-469.
- \_\_\_\_\_. Introductory Ethics. Englewood Cliffs: Prentice-Hall, 1978.
- Harman, G. "Toward a Theory of Intrinsic Value." Journal of Philosophy 64 (1967): 796-804.
- Klemke, E. The Epistemology of G. E. Moore. Evanston: Northwestern University Press, 1969.
- Langford, C. "Moore's Notion of Analysis," from: P. A. Schilpp (ed.) The Philosophy of G. E. Moore. New York: Tudor Publishing Co., 1952.
- Mandelbaum, M. The Phenomenology of Moral Experience. Baltimore: John Hopkins Press, 1958.

- Meinong, A. On Emotional Presentation. Evanston: Northwestern University Press, 1972.
- Mill, J. Utilitarianism. Indianapolis: Bobbs-Merrill, 1957.
- Moore, G. E. Principia Ethica. London: Cambridge University Press, 1903.
- \_\_\_\_\_. Ethics. London: Oxford University Press, 1965.
- Oldfield. "An Approach to a Theory of Intrinsic Value." Philosophical Studies 3 (1977): 233-249.
- Pastin, M. "Warranting Reconsidered." Synthese 38 (1978): 459-464.
- Quinn, W. "Theories of Intrinsic Value." American Philosophical Quarterly 11 (1974): 123-132.
- Sosa, E. "The Foundations of Foundationalism." Nous 13 (1980).
- Taylor, P. Normative Discourse. Englewood Cliffs: Prentice-Hall, 1961.



