Chapter 1

Desert as Fit: An Axiomatic Analysis

Gustaf Arrhenius

1. Introduction

Total Utilitarianism is the view that an action is right if and only if it maximizes the sum total of people's well-being. A common objection to Total Utilitarianism is that it is insensitive to matters of distributive justice. For example, for a given amount of well-being, Total Utilitarianism is indifferent between an equal distribution and any unequal distribution, and if there would be a tiny gain in well-being by moving from an equal distribution to an unequal, we have a duty to do so. To meet the objection from justice, Fred Feldman has suggested a desert-adjusted version of Total Utilitarianism – 'Justicism' – which in addition to the value of well-being takes into account factors concerning people's desert.[1] Feldman's suggestion is novel and interesting but his theory has been severely criticized as a theory of distributive justice.[2] In the present paper, I shall try to salvage what I think might be a kernel of truth in Feldman's suggestion, or at least a kernel that is worthy of further investigation.

In Feldman's presentation of Justicism, he oscillates between two different ways of taking desert into account: the merit-idea and the fit-idea.[3] According to the former idea, the higher the desert level, the higher the value of pleasure. The latter idea, on the other hand, focuses on the degree of fit between desert and receipt of pleasure. The merit-idea corresponds pretty well with Feldman's explicit formulation of Justicism and with his application of Justicism to issues of distributive justice, whereas the fit-idea does the work in his discussion of certain problems in population axiology.[4] In Arrhenius (2003), I made a partial suggestion of how to formulate the fit-idea more explicitly in the context of population axiology. In the present paper, I am going to develop this idea as an approach to distributive justice.

I shall focus on a combination of the fit-idea with a value function according to which the intrinsic value of a life is determined by the sum of the value of pleasure and the value of the fit between pleasure and the recipient's desert, that is, an additively separable value function. I shall introduce a formalism in which we can state this value function and the fit-idea in an exact manner. This will make the structure and the implications of the theory take a clearer form as compared to previous formulations. I shall then suggest that the core of the fit-idea can be reduced to two central principles. The combination of these two principles and the

value function forms a theory that I call Additively Separable Fit Justicism. Lastly, I shall prove that this theory implies a number of auxiliary principles that capture some important intuitions about desert and distributive justice quite well and which will enable us to meet some of the objections directed against Justicism as a theory of distributive justice. First, however, I shall describe Feldman's original proposal.

2. Feldman's Desert-Adjusted Utilitarianism

In hedonism, the value of an episode of pleasure or pain is a function of its hedonic level and duration. In Justicism, the value of such an episode is determined not only by the hedonic level and duration but also by the recipient's desert level: '… the intrinsic value of an episode of pleasure or pain is a function of two variables: (i) the amount of pleasure or pain the recipient *receives* in that episode, and (ii) the amount of pleasure or pain the recipient *deserves* in that episode' (Feldman, 1997, pp. 162 f., emphasis in original).[5] A person's desert level is determined by factors such as her excessive or deficient past receipt of pleasure or pain, her moral worthiness, her rights and legitimate claims, her past conscientious efforts, and so forth.[6] A person is said to have 'positive desert' if she deserves some pleasure, 'negative desert' if she deserves some pain, and 'neutral desert' if she deserves neither pleasure nor pain. Feldman partly describes the relationship between pleasure, pain, desert and intrinsic value with the following six principles:[7]

M1. Positive desert enhances the intrinsic goodness of pleasure.
M2. Negative desert mitigates the intrinsic goodness of pleasure.
M3. Neutral desert neither enhances nor mitigates the intrinsic goodness of pleasure.
M4. Positive desert enhances the intrinsic badness of pain.
M5. Negative desert mitigates the intrinsic badness of pain.
M6. Neutral desert neither enhances nor mitigates the intrinsic badness of pain.

In his discussion Feldman does not consistently abide by his own principles. He claims that 'receipt of much less good than you deserve is not good for the world' and that the intrinsic value of a life led by a person who deserves 100 units of pleasure but receives only one unit is -49.[8] These claims are clearly inconsistent with M1.[9] As Ingmar Persson has pointed out, Feldman oscillates between two ideas: the merit-idea and the fit-idea.[10] According to the former, the higher the desert level, the higher the value of pleasure. The latter idea, on the other hand, focuses on the degree of fit between desert and receipt. The merit-idea corresponds pretty well with M1-M6 above, whereas the fit-idea corresponds with the statements above of the value of deserved and undeserved pleasure.

How should we formulate the fit-idea more exactly? In Arrhenius (2003), I suggested that we could formulate it in terms of enhancement and mitigation of

deserved, under-deserved and over-deserved pleasure and pain. A person's pleasure or pain is 'deserved' if it roughly corresponds to her desert level, that is, if she receives exactly what she deserves or fairly close to what she deserves. If a person's pleasure or pain does not roughly correspond with her desert level and thus is clearly more (less) than she deserves, then this pleasure or pain is 'under-deserved' ('over-deserved'). Roughly, my idea was that desert enhances the goodness of deserved pleasure and mitigates the badness of deserved pain, whereas positive desert mitigates the goodness of under-deserved and over-deserved pleasure, and negative desert enhances the badness of over- and under-deserved pain.

Now, although these principles capture the fit-idea quite well, I think they do not properly bring out the central idea and structure of the fit-idea, and the relationship between its different parts. Firstly, there is an ambiguity that the above formulation of the fit-idea shares with Feldman's original formulation of Justicism. As formulated, the intrinsic value of an episode of pleasure or pain depends on the amount of pleasure or pain the recipient deserves in that episode. Strictly speaking, this is not compatible with Feldman's own idea of intrinsic value. As he writes in another context, '[s]urely, if something is intrinsically good, it must be good in virtue of the way it is in itself, not merely because of some extrinsic relation it happens to bear to some other thing' (Feldman, 1997, p. 138). But according to Justicism, as stated above and by Feldman, it seems that the intrinsic value of an episode of pleasure or pain depends on contingent facts regarding the desert level of the recipient. Nevertheless, this should only be taken metaphorically. What Feldman actually means is that the only carriers of basic intrinsic value are compound states of affairs consisting of a person's experience of pleasure or pain and her desert level.[11]

An alternative to Feldman's formulation would be to let the fit between desert and receipt be another carrier of intrinsic value in addition to pleasure and pain. Let us call the value of the fit between desert and receipt in a life its *desert value* and the value of the pleasure or pain in a life its *hedonic value*. The intrinsic value of a life would then be the sum of its hedonic value and its desert value. For example, Feldman says that the value of a life enjoying a deserved one unit of pleasure is two.[12] On the suggested revision of Justicism, this means that the intrinsic value of the pleasure in this life is one unit, and the intrinsic value of the fit between desert and receipt in this life, its desert value, is also one unit. These two values taken together yield that the intrinsic value of this life is two units.

One putative advantage with this approach as compared to Feldman's has to do with the proper attitude towards good and bad states of affairs. As Thomas Hurka writes:

> Feldman treats desert not as a separate value additional to those of pleasure and pain but as a factor that adjusts those states' values up and down, so undeserved pleasure is purely and simply evil, and deserved pain purely and simply good… Feldman's view implies that our response to the infliction of deserved punishment should be pure and simple pleasure, since all that is being created is a good. The view I have described

> implies, more plausibly, that our response should combine satisfaction that justice is being done with pain at the infliction of pain, with the latter emotion limiting and qualifying the first (Hurka, 2001a, p. 11, fn. 8). [13]

Consider again the case of a life led by person P who deserves 100 units of pleasure but receives only one unit. According to Feldman's view, there is only one basic carrier of value here, consisting of P's experience of one unit of pleasure and P's desert level of a 100 units of pleasure, which has an intrinsic value of -49. Accordingly, one might think, we should only have a negative attitude towards this state of affairs since it is only bad. On the view suggested above, there are two carriers of intrinsic value in this case, one consisting of P's experience of one unit of pleasure, and another one consisting of the fit between P's pleasure and her desert level. Consequently, there is one state of affairs towards which we should have a positive attitude – the experience of pleasure – and another state of affairs towards which we should have a negative attitude – the mismatch between desert and receipt.

Nonetheless, I do not find this objection against Feldman's formulation of Justicism decisive. Firstly, one could respond that the proper attitudes to states of affairs do not only concern the intrinsic value of the states but also what is good or bad *for* people, that is, their well-being. In the case described above, there is one thing we should have a positive attitude towards, namely, that one person is experiencing one unit of pleasure, which is good for this person. Secondly, one could deny the intuition that one should have a positive attitude towards P's experience of one unit of pleasure, given that one believes in an axiology where the basic carriers of value are complex states of affairs consisting of a person's experience of pleasure or pain and her desert level. Rather, one could claim, one should have no attitude toward P's experience of pleasure since, given the axiology in question, the value of that state is evaluatively underdetermined.[14]

At any rate, I do not think the important issue for the present paper is whether we have one or two carriers of intrinsic value. Rather, the question is whether the best way of spelling out Justicism is in terms of an additively separable value function with the pleasure or pain and the fit between desert and receipt of pleasure or pain as the two arguments in the value function. Whether or not this is true will not depend on metaphysical considerations regarding the number of carriers of intrinsic value but rather on whether such a formulation best fits our intuitions concerning pleasure and desert. If we would like to retain the idea that there is just one carrier of intrinsic value, we just need to slightly adjust our definition of desert value in terms of the difference between the intrinsic value of a life and its hedonic value, that is, in terms of how much the fit between desert and receipt contributes or detracts from the intrinsic value of a life. We are, so to say, factoring out the desert component of the intrinsic value of a life. Thus, if the intrinsic value of a life enjoying a deserved one unit of pleasure is two, then the contributive value of the pleasure in this life is one unit, and the contributive value of the fit between desert and receipt, the desert value, is also one unit, and these two values taken together

yield that the intrinsic value of this life is two units, in accordance with Feldman's view.

Another problem for my reformulation of Feldman's principle is how we should understand the talk of enhancement and mitigation. Compared to what value is there an enhancement or mitigation? It is quite easy to make sense of this idea in connection with the merit-idea since, according to that idea, an increase in the positive desert level in a life with positive hedonic level increases the intrinsic value of such a life, and a decrease in the negative desert level (e.g., from slightly negative to very negative) in a life with negative hedonic level increases the intrinsic value of such a life. This might be clearer if we formulate M1 more exactly. Let (p,d) be a life with hedonic level p and desert level d and let IV be a function that returns the intrinsic value of a life. M1 can now be formulated as follows:

M1′: If $p > 0$ and $d_1 > d_2 \geq 0$, then $\mathrm{IV}(p,d_1) > \mathrm{IV}(p,d_2)$.

In other words, if two lives have the same positive hedonic level but one has higher positive desert level, then the latter life has higher intrinsic value.

This does not work with the fit-idea, however. An increase in positive desert might sometimes increase the intrinsic value of a life and sometimes decrease the intrinsic value of a life. For example, recall that the value of a life enjoying a deserved one unit of pleasure is two and that the value of a life led by a person who deserves 100 units of pleasure but receives only one unit is -49. Thus, in this case an increase in the desert level from 1 unit to 100 units yields a decrease in the intrinsic value of a life.

There might be a way around this problem but I think there is a way of formulating the fit-idea and Justicism without any talk about enhancement and mitigation. This is what I shall now turn to.

3. Justicism and The Fit-Idea Reformulated

For the purpose of stating Justicism and the fit-idea more precisely, it will be useful to state some definitions and assumptions, and introduce some notational conventions. Let p_1, p_2,… and so on be the numerical representation of a certain hedonic level of a life, that is, the pleasure and pain in that life taken as a whole, and let d_1, d_2,… and so on be the numerical representation of a certain desert level of a life, that is, the desert of that life taken as a whole.[15] For the sake of simplicity, I shall assume that the hedonic and desert level can be measured on a ratio scale. Let (p,d) be a life with hedonic level p and desert level d. Thus, (1,2) represents a life in which a person receives one unit of pleasure but deserves two units of pleasure. Let A, B, C and so forth be populations of lives, represented by vectors $((p_1,d_1), (p_2,d_2),\ldots, (p_n,d_n))$. For example, if A=((1,5), (3,4)), then 3 is the numerical representation of the hedonic level of the second person in population A. Let IV be a function that returns the numerical representation of the intrinsic value of a life or

a population, let HV be a function that returns the numerical representation of the hedonic value of a life or a population, and let DV be a function that returns the desert value of a life or a population.

Any version of Justicism can be divided into three parts. One part is the value function that tells us how to aggregate hedonic and desert value into a measure of the intrinsic value of lives and populations. The other two parts tell us about how the hedonic value of a life depends on the pleasure or pain in the life, and how the desert value of a life depends on the fit between the pleasure or pain and the desert in a life. As we said above, we are going to investigate a version of Justicism according to which the intrinsic value of a life equals the sum of its hedonic and desert values, and where the desert value is understood along the lines of the fit-idea. Let us call this theory Additively Separable Fit Justicism or ASFJ for short. The value function of ASFJ is defined by the following principles:

V1: $IV(p,d) = HV(p,d)+DV(p,d)$.

It should be noticed that it is not self-evident that we should formulate Justicism as an additively separable value function. It seems a quite natural move if we consider the fit between desert and receipt another intrinsic value in addition to pleasure. As we said above, we have not committed ourselves to any specific view on this matter. If one holds the view that there is only one type of basic carrier of intrinsic value, that is, compound states of affairs consisting of a person's experience of pleasure or pain and her desert level, then one might think that it does not follow in a natural way that the intrinsic value of a life is a sum of its hedonic and desert values. It could instead be, say, the product of its hedonic and desert values, or some more complicated function. Making it an additively separable function has the advantage of simplicity, however, and the purpose of the present paper is to investigate how far we can get with such a value function in combination with the fit-idea.

In line with Feldman's theory, we are going to make the intrinsic value of a population A equal to the sum of the intrinsic value of the lives in A:[16]

V2: $IV(A) = IV((p_1,d_1), (p_2,d_2),\ldots, (p_n,d_n)) =$
$IV(p_1,d_1)+IV(p_2,d_2)+\ldots+IV(p_n,d_n)$.

It follows from the two definitions above that the hedonic value of a population equals the sum of the hedonic values of the lives in the population, and that the desert value of a population equals the sum of the desert values of the lives in the population.

Since Justicism is a desert-adjusted version of utilitarianism, we are also going to assume that the hedonic value of a life equals its hedonic level:

HV1: $HV(p,d) = p$.

It follows from the above definitions that the hedonic value of a population equals the sum of the hedonic levels of the lives in the population. There are, of course, alternatives to the above formulations that we could have considered. For example, we could have made the hedonic value of a life a marginally decreasing function of the pleasure or pain in it and in such a way captured the intuition that we should give priority to the worst off. Likewise, instead of summing each life's hedonic value to get a measure of a population's hedonic value, we could instead have used, say, averaging to avoid Derek Parfit's repugnant conclusion.[17] However, the attractive feature of Feldman's original suggestions is that the desert part of the equation might be able to handle all of these intuitions, that is, introducing a desert component in the axiology might make tinkering with the aggregation of pleasure and pain superfluous. That is clearly worth investigating; hence the definitions above.

In the following, I shall take the above definitions for granted and when I sometimes claim, without any reference to the above definitions, that a certain version of the fit-idea has a certain implication with respect to the intrinsic value of a life or a population, then that is just short for saying that it has this implication in conjunction with the above definitions.

4. The First Central Fit-Idea

As I said above, I think there are two central parts to the fit-idea. Here is the first one:

> *The First Central Fit-Idea*: The better the fit between receipt and desert, the higher the desert value.

How should we formulate this idea more exactly? Actually, there are several options available here, and it will be worthwhile to consider some different alternatives. Here is a first try:

F1-1: If $d \geq p_1 > p_2$, then $DV(p_1,d) > DV(p_2,d)$;
if $p_1 > p_2 \geq d$, then $DV(p_1,d) < DV(p_2,d)$.

In words: If two lives both have less pleasure or more pain than they deserve, or both have more pleasure or less pain than they deserve, then the life with the lesser difference between desert and receipt of pleasure or pain has the higher desert value.

This principle partly captures the first central idea and in a rather weak way. It is compatible with there being an asymmetry between the value of over- and under-deserved pleasure such that, for example, DV(4,3) is greater than DV(2,3). One might find this an attractive feature since one might reasonably think that it is better that a person gets one unit of pleasure more than she deserves rather than one unit less. F1-1 is, on the other hand, also compatible with DV(4,3) being

smaller than DV(2,3) which seems hard to defend. We could avoid this implication by adding a further requirement to F1-1. The question is, however, whether we should capture the intuition that it is better that one get one unit more rather than one unit less than one deserves by introducing an asymmetry between the value of over- and under-deserved pleasure or pain, or by assigning increasing value to increases in the hedonic level, other things being equal. Given our definitions of the intrinsic value of a life and the value of pleasure and pain (V1 and HV1), it follows that if DV(4,3) = DV(2,3), then IV(4,3) is greater than IV(2,3) since HV(4,3) = 4 is greater than HV(2,3) = 2. I think this sufficiently accounts for the intuition that it is better that a person gets more than they deserve rather than less than they deserve.[18] Thus, I think we should go for an explication of the central fit-idea that implies symmetry between the value of under- and over-deserved pleasure and pain:

F1-2: If $|p_1-d| < |p_2-d|$, then $DV(p_1,d) > DV(p_2,d)$; and
if $|p_1-d| = |p_2-d|$, then $DV(p_1,d) = DV(p_2,d)$.[19]

In words: If two persons deserve the same amount of pleasure or pain, and the difference between the first person's receipt and desert is less than (equal to) the difference between the second person's receipt and desert, then the desert value of the first person's life is greater than (equal to) the desert value of the second person's life. F1-2 implies a symmetry such that DV(4,3) = DV(2,3), or in general that $DV(y+x,y) = DV(y-x,y)$.

One might consider that apart from F1-2, there should also be a symmetry such that DV(4,5) = DV(4,3), that is, if $|p-d_1| \leq |p-d_2|$, then $DV(p,d_1) \geq DV(p,d_2)$. The following principle is a combination of this idea and F1-2:

F1-3: If $|p_1-d_1| < |p_2-d_2|$, then $DV(p_1,d_1) > DV(p_2,d_2)$; and
if $|p_1-d_1| = |p_2-d_2|$, then $DV(p_1,d_1) = DV(p_2,d_2)$

In words: If the difference between receipt and desert in a certain life is smaller than (equal to) the difference between receipt and desert in another life, then the desert value of the former life is greater than (equal to) the desert value of the latter life.

This principle implies a further and quite interesting symmetry. It implies that the desert value only depends on the absolute difference between receipt and desert and not on the magnitude of the receipt or the desert. For example, it implies that DV(2,2) = DV(100,100), DV(1,2) = DV(99,100), and that DV(1,2) > DV(98,100). An intuitive support in favor of this implication of F1-3 could be that when a person receives exactly what she deserves, then perfect justice is done, and perfect justice has one and the same value in all situations. Likewise for discrepancies between desert and receipt; that is, the imperfect justice of someone receiving, say, two units less than she deserves has the same value in all situations.[20]

F1-3 in conjunction with V1 also implies two principles that seem to fit our intuitions about desert and receipt. Since the intrinsic value of a life equals the sum

of its hedonic and desert values, F1-3 implies that for a given hedonic level, the better the fit between receipt and desert, the higher the intrinsic value of a life:

D1: If $|p\text{-}d_1| < |p\text{-}d_2|$, then $IV(p,d_1) > IV(p,d_2)$.

Moreover, F1-3 implies that, for any given hedonic level, the life with the highest intrinsic value is the life with a perfect match between the hedonic level and the desert level. Any deviation between the receipt and the desert decreases the intrinsic value:

D2: If $d_1 = p$ and $d_2 \neq p$, then $IV(p,d_1) > IV(p,d_2)$.

F1-3 has an implication, however, that gives us reason to reject it, given our definition of the intrinsic value of a life. Firstly, F1-3 implies that DV(1,2) = DV(99,100) and that DV(2,2) = DV(100,100). Assume that we can give one unit of pleasure either to (1,2) or to (99,100). It follows from HV1 above that the increase in hedonic value will be the same in both cases, namely one unit. Since F1-3 implies that DV(1,2) = DV(99,100) and that DV(2,2) = DV(100,100), the increase in desert value will be the same in both cases (that is, DV(2,2)-DV(1,2) = DV(100,100)-DV(99,100)). Since, from V1 above, $IV(p,d) = HV(p,d) + DV(p,d)$, F1-3 implies that giving one unit of pleasure to (1,2) increases intrinsic value equally much as giving one unit of pleasure to (99,100). Yet, it seems reasonable to claim that from the perspective of proportional justice, (1,2) is much worse off than (99,100) since she only has half of the pleasure that she deserves, whereas (99,100) has almost all she deserves. Hence, we should give the one unit of pleasure to (1,2) rather than (99,100). Thus, the desert value does depend on the magnitude of the receipt and the desert.

Erik Carlson has suggested a theory that implies this kind of symmetry. The intrinsic value of a life is determined by the following formulas in Carlson's theory:[21]

$IV(p,d)=d+(p-d)^k$ if $d \leq p$, $0 < k < 1$;
$IV(p,d)=d-(d-p)^m$ if $d > p$, $m > 1$.

It follows from above that $IV(1,2) = 2-(2-1)^m = 1$, $IV(2,2) = 2+(2-2)^k = 2$, $IV(99,100) = 100-(100-99)^m = 99$, and that $IV(100,100) = 100+(100-100)^k = 100$. Again, assume that we can give one unit of pleasure either to (1,2) or to (99,100). Since IV(2,2)-IV(1,2) = 2-1 and IV(100,100)-IV(99,100) = 100-99, the increase in intrinsic value will be the same in both cases according to Carlson's theory, namely one unit. Consequently, his theory implies that giving one unit of pleasure to (1,2) increases intrinsic value equally as much as giving one unit of pleasure to (99,100).

In light of the above discussion, I think we should jettison F1-3 and with it Carlson's theory. What about F1-2? It does not have the implications of F1-3 just discussed, but one might suspect that it has analogous consequences, for example

that we should be indifferent between giving ten units of pleasure to (100,0) and (-100,0) since, according to F1-2, DV(100,0) = DV(-100,0). However, since F1-2 does not imply that DV(110,0) = DV(-90,0), it does not follow that we should be indifferent between giving ten units of pleasure to (100,0) and (-100,0).

F1-2 has implications that are desirable from the perspective of distributive justice. It sometimes implies that we should redistribute pleasure and pain from people that have more than they deserve to people that have less than they deserve. For example, a population A = ((100,0),(-100,0)) is worse than a population B = ((90,0),(-90,0)) since HV(A) = HV(B) = 0, DV(90,0) > DV(100,0), and DV(-90,0) > DV(-100,0).

More importantly, F1-2, in conjunction with V1, V2, and HV1, satisfies a plausible adequacy condition suggested by Carlson. He has proposed that '[if] a consequentialist theory is to alleviate the objection from justice' then it should satisfy the following adequacy condition:

> J1: If *n* units of pleasure or pain are to be distributed among a certain number of people with a total desert level of *n*, then the distribution where each person gets exactly what she deserves is better than any alternative distribution.[22]

As Carlson puts it, '[a] theory which does not satisfy J1 sometimes allows that some person get more that she deserves, and another person gets less, even though there is no reason in terms of maximizing net pleasure to allow this. Such a theory, it seems, permits us to depart from the requirements of justice for no good reason' (Carlson, 1997, p. 311). As Carlson has shown, Feldman's version of Justicism does not satisfy this condition.[23] Now, since the hedonic value is going to be the same for any distribution of a fixed amount of pleasure or pain, the population with the highest intrinsic value will be the population with the highest aggregate desert value according to V2. It follows from F1-2 that the maximal desert value of a life with a given desert level is the life where the person gets exactly what she deserves. Consequently, the aggregate desert value is maximized when everybody gets exactly what they deserve. Hence, F1-2, V1, V2, and HV1 together imply J1.

5. The Second Central Fit-Idea

The next central principle of the fit-idea concerns the relative importance of increases in fit:

> *The Second Central Fit-Idea*: The contributive value of a given increase in fit decreases the closer to the desert level one gets.

The intuitive idea behind the second fit-idea is that the greater the mismatch between receipt and desert, the greater the urgency to increase the fit between receipt and pleasure. Here is an example: Assume that d=10, p_1=5, p_2=4, p_3=3. Then the increase in desert value from a change in pleasure from 4 to 5 is less than the increase in desert value from a change in pleasure from 3 to 4.

Here is how we can formulate the second fit-idea more exactly:

F2: If $|e_1|=|e_2|$, $|p_1-d|>|p_1+e_1-d|$, $|p_2-d|>|p_2+e_2-d|$, and $|p_1-d|>|p_2-d|$, then $DV(p_2+e_2,d)-DV(p_2,d) < DV(p_1+e_1,d)-DV(p_1,d)$.

In words: If we can increase the fit between desert and receipt in two lives by adjusting their hedonic level up or down by a fixed amount, then the life with the greater difference between receipt and desert will get the greater increase in desert value from the adjustment of its hedonic level.

F2, in conjunction with F1-2, V1, V2, and HV1, has some attractive implications in regard to distribution of pleasure and pain. For a start, it implies that the best distribution of a given amount of pleasure or pain between two persons with the same desert level is an equal distribution:

D3: For any d, if $p_1>p_2>p_3$ and $p_1+p_3=2p_2$, then, $IV((p_2,d),(p_2,d))>IV((p_1,d),(p_3,d))$.

Here is a proof. We shall divide the proof into three different cases depending on the desert level: $d > p_1-e/2$, $d < p_3-e/2$, and $p_1-e/2 \geq d \geq p_3+e/2$, where e is the difference in pleasure between p_1 and p_2. Let us start with $d > p_1-e/2$. Let

(1) p_1, p_2, p_3 be three pleasure levels, d be a desert level, and e a difference in pleasure levels such that $p_1>p_2>p_3$, $p_1+p_3=2p_2$, $e=p_1-p_2$, and $d > p_1-e/2$;
(2) A and B be two populations such that A=((p_2,d), (p_2,d)) and B=((p_1,d), (p_3,d)).

It follows from (1), (2), and the definition of the hedonic value of a population that

(3) HV(A)=HV(B).

From (2), (3), and the definition of the intrinsic value of a population, it follows that

(4) IV(A)>IV(B) iff DV(A)>DV(B) iff $2DV(p_2,d)>DV(p_1,d)+DV(p_3,d)$.

That is, the ranking of A and B will be decided by their respective desert values. Since $d > p_1-e/2$ and $e = p_1-p_2$ (from (1)), we get

(5) $|p_3-d|>|p_2-d|$;

(6) $|p_2\text{-}d|>|p_2+e\text{-}d|$;
(7) $|p_3\text{-}d|>|p_3+e\text{-}d|$.

F2 and ((5)-(7)) imply

(8) $DV(p_2+e,d)\text{-}DV(p_2,d) < DV(p_3+e,d)\text{-}DV(p_3,d)$.

Since $p_1+p_3 = 2p_2$ and $p_1\text{-}p_2 = e$ (from (1)), it follows that

(9) $p_1\text{-}p_2=p_2\text{-}p_3=e$,

which implies

(10) $p_1=p_2+e$ and $p_2=p_3+e$,

which in turn yields that (8) is equivalent to

(11) $DV(p_1,d)\text{-}DV(p_2,d) < DV(p_2,d)\text{-}DV(p_3,d)$.

By rearranging the terms in (11) we get

(12) $DV(p_1,d) + DV(p_3,d) < DV(p_2,d) + DV(p_2,d)$,

which together with (4) implies that

(13) IV(A) > IV(B). Q.E.D.

The proof for $p_3 > d+e/2$ follows the same pattern as the above proof, so I shall not spell it out here. The proof for $p_1\text{-}e/2 \geq d \geq p_3+e/2$ is quite simple. Assume first that $p_1\text{-}e/2 \geq d > p_3+e/2$. It follows that $|p_2\text{-}d| \leq |p_1\text{-}d|$ and $|p_2\text{-}d| < |p_3\text{-}d|$ (just look at the maximal and minimal values of d). F1-2 thus implies that $DV(p_2,d) \geq DV(p_1,d)$ and $DV(p_2,d) > DV(p_3,d)$. It follows that $DV(p_2,d)+DV(p_2,d) > DV(p_1,d)+DV(p_3,d)$ which together with (4) above implies that IV(A)>IV(B). Assume secondly that $p_1\text{-}e/2>d \geq p_3+e/2$. It follows that $|p_2\text{-}d| < |p_1\text{-}d|$ and $|p_2\text{-}d| \leq |p_3\text{-}d|$. F1-2 thus implies that $DV(p_2,d) > DV(p_1,d)$ and $DV(p_2,d) \geq DV(p_3,d)$. It follows that $DV(p_2,d) + DV(p_2,d) > DV(p_1,d) + DV(p_3,d)$, which together with (4) above implies that IV(A) > IV(B). Thus, for any d, if $p_1>p_2>p_3$ and $p_1+p_3=2p_2$, then F2 implies that $IV((p_2,d),(p_2,d)) > IV((p_1,d),(p_3,d))$, which is exactly what D3 states. Q.E.D.

In general, F2, F1-2, V1, V2, and HV1 together imply that an equal distribution is the best distribution of any given amount of pleasure or pain to any given number of people with the same desert level, that is, a generalization of D3:

D4: For any d, and any populations A = $((p_1,d), (p_2,d),\ldots,(p_n,d))$ and B = $((q_1,d), (q_2,d),\ldots,(q_n,d))$, $n \geq 2$, if $p_i > p_j$ for some i, $j \leq n$, and $q_i=(p_1+p_2+\ldots+p_n)/n$ for all $i \leq n$, then IV(B) > IV(A).

Since the intrinsic value of a population is the sum of the intrinsic values of the lives in the population, that is, the value-function is additively separable, D4 follows from D3. For any population, we can by repeated application of D3 to different pairs of lives in the original population generate successively better populations until we reach a population with an equal distribution of pleasure or pain. The final population will be better than the original population by virtue of the transitivity of intrinsic value. Here is an example. Let us say that population A consists of four people: (10,d), (8,d), (6,d), and (4,d). The average pleasure in A is 7 units (28/4). D3 implies that IV((10,d), (8,d)) < IV((9,d), (9,d)) and IV((6,d), (4,d)) < IV((5,d), (5,d)). It follows from this and V2 that IV(A) < IV((9,d), (9,d), (5,d), (5,d)). Applying D3 again, we get that IV((9,d), (5,d)) < IV((7,d), (7,d)). Again, it follows from this and V2 that IV((9,d), (9,d), (5,d), (5,d)) < IV((7,d), (7,d), (7,d), (7,d)) which, by transitivity, is better than A.

In light of the above discussion, I suggest that we define Additively Separable Fit Justicism as a conjunction of V1, V2, HV1, F1-2 and F2. We should appreciate that ASFJ implies D4. Carlson suggests that D4 is a crucial requirement for a theory to meet the objection from justice. As he puts it, '[i]f everybody has equal desert, it is a breach of justice to give more to one person than to another' (Carlson 1997, p. 311). As Carlson has pointed out, Feldman's theory violates D4 and directs us to distribute pleasure or pain unequally although everyone deserves the same level of pleasure or pain.[24] In other words, ASFJ avoids another criticism directed against Feldman's version of Justicism.

6. Conclusion

The fit-idea is a promising concept for consequentialists seeking a theory that alleviates the objection from justice. Additively Separable Fit Justicism is superior to Feldman's original version of Justicism since the former but not the latter satisfies two crucial adequacy conditions for such a theory. More work is clearly needed to develop a full-fledged version of Justicism along the lines of the fit-idea. The framework presented here may provide a fruitful avenue for such further work.[25]

Notes

[1] Feldman, 1997. Feldman also suggests that Justicism generates a plausible answer to the paradoxes of population axiology and our duties to future generations.

[2] See Carlson, 1997, Persson, 1997, and Vallentyne, 1995. In Arrhenius, 2000 and 2003, I show that Justicism is unsatisfactory as a population axiology.

[3] See Persson, 1997, Arrhenius, 2003.

[4] See Arrhenius, 2000 and 2003.

[5] Feldman formulates Justicism as a version of classical hedonism mainly for pedagogical reasons. It could equally well have been stated in terms of Feldman's propositional theory of pleasure or in terms of some other theory of welfare. See ibid., p. 152.

[6] Feldman, 1997, pp. 161f., 202 f.

[7] Feldman, 1997, pp. 163-9.

[8] Feldman, 1997, pp. 206, 163.

[9] Carlson, 1997, p. 315, makes the same point.

[10] Persson, 1997.

[11] Feldman, 2000, and personal communication with Feldman.

[12] More exactly, Feldman, 1997, p. 212, considers a world of a billion billion people where each person deserves and receives one unit of pleasure. Then he concludes that 'the … value of this world is two billion billion.'

[13] See also Hurka, 1998, pp. 309-311; 2001, pp. 193-7.

[14] Michael Zimmerman uses the term 'evaluatively inadequate' for this phenomenon. See Zimmerman, 2001.

[15] Notice that we are not taking a stand on how to aggregate episodes of pleasure and pain into a measure of the pleasure and pain in a life as a whole. It could be done by just summing the episodes, but there are other approaches that we might find more in line with our intuitions. Likewise for desert. Cf. Arrhenius, 2005.

[16] Feldman, 1997, p. 169, writes: 'The intrinsic value of a whole consequence is the sum of the justice-adjusted intrinsic values of the episodes of pleasure and pain that occur in that consequence.' On p. 208 he says that '…the relevant…value of a world…is the sum of the values of the lives lived there, adjusted for desert….'

[17] See Parfit, 1984, and Arrhenius, 2000.

[18] Those who believe that from a *pure* desert perspective it is better that one get one unit more rather than one unit less than one deserves would not be satisfied by this account, however, and should stick with F1-1 as the best explication of the first fit-idea.

[19] '$|a-b|$' represents the absolute difference between the numerical values a and b. If $a \geq 0$, then $|a|=a$; if $a<0$, then $|a| = -a$. For example, $|5| = 5$ and $|-5| = 5$.

[20] If Feldman were to agree with our definition of the intrinsic value of a life, then he would disagree with this symmetry. Recall that Feldman holds that the value of a life enjoying a deserved one unit of pleasure is two. However, in his discussion of the repugnant conclusion he says that the life of a person who deserves 100 units of pleasure and receives exactly that amount of pleasure has an intrinsic value of 200 (Feldman, 1997, p. 206). Now, since IV(1,1) = 2 and HV(1,1) = 1, it follows that DV(1,1) = 1, and since IV(100,100) = 200 and HV(100,100) = 100, it follows that DV(100,100) = 100. According to F1-3, however, DV(1,1) = DV(100,100). On the other hand, Feldman might very well reject V1 above and instead opt for a definition of the intrinsic value of a life according to which $IV(p,d) = HV(p,d)\times DV(p,d)$. Given this definition, the above evaluations are compatible with F1-3. It has, however, in combination with HV1 above, the odd feature that the intrinsic value of a life with zero pleasure is always zero, irrespective of the desert value. I shall not pursue this matter further here.

[21] Carlson, 1997, p. 312. I have reformulated Carlson's theory in terms of the notation used here.

[22] Carlson, 1997, p. 311. I have rephrased Carlson's condition.
[23] Carlson, 1997, p. 311.
[24] Carlson, 1997, pp. 309-311.
[25]I would like to thank John Broome, Erik Carlson, and especially Krister Bykvist and Michael Zimmerman, for fruitful discussions and comments. Financial support from the Bank of Sweden Tercentenary Foundation, STINT, and the Swedish Collegium for Advanced Study is gratefully acknowledged.

References

Arrhenius, Gustaf (2000), *Future Generations: A Challenge for Moral Theory*, F.D.-Diss., University Printers, Uppsala.
Arrhenius, Gustaf (2003), 'Feldman's Desert-Adjusted Utilitarianism and Population Ethics,' *Utilitas*, Vol. 15, pp. 225-36.
Arrhenius, Gustaf (2005) 'Superiority in Value,' *Philosophical Studies*, forthcoming.
Carlson, Erik (1997), 'Consequentialism, Distribution and Desert,' *Utilitas*, Vol. 9, pp. 307-18.
Feldman, Fred (1997), *Utilitarianism, Hedonism, and Desert: Essays in Moral Philosophy,* Cambridge University Press, Cambridge.
Feldman, Fred (2000), 'Basic Intrinsic Value,'*Philosophical Studies*, Vol. 99, pp. 319-46.
Hurka, Thomas (1998), 'Two Kinds of Organic Unity,' *The Journal of Ethics*, Vol. 2, pp. 299-320.
Hurka, Thomas (2001a), 'The Common Structure of Virtue and Desert,' *Ethics*, Vol. 112, pp. 6-31.
Hurka, Thomas (2001b), *Virtue, Vice, and Value,* Oxford University Press, New York.
Parfit, Derek (1991), *Reasons and Persons* [1984], Clarendon Press, Oxford.
Persson, Ingmar (1997), 'Ambiguities in Feldman's Desert-adjusted Values,' *Utilitas*, Vol. 9, No. 3, pp. 319-28.
Vallentyne, Peter (1995), 'Taking Justice Too Seriously,' *Utilitas*, Vol. 7, pp. 207-16.
Zimmerman, Michael J. (2001), *The Nature of Intrinsic Value*, Rowman & Littlefield, Lanham, Md.