

GRAHAM ODDIE

## RECOMBINANT VALUES

(Received 7 June 2000; received in revised form 27 November 2000)

**ABSTRACT.** An attractive admirer of George Bernard Shaw once wrote to him with a not-so modest proposal: “You have the greatest brain in the world, and I have the most beautiful body; so we ought to produce the most perfect child.” Shaw replied: “What if the child inherits my body and your brains?”<sup>1</sup>

What if, indeed? Shaw’s retort is interesting not because it reveals a grasp of elementary genetics, but rather because it suggests his grasp of an interesting and important principle of axiology. Since the brainy but ugly Shaw and his beautiful but apparently dim admirer both fall short of the ideal, she suggests that the best thing would be to genetically recombine his intelligence with her beauty. But what then would be the value of another genetic possibility: that of recombining his ugliness with her stupidity? Underlining the prompted inference is a fundamental principle of the theory of value which, perhaps surprisingly, has so far gone largely unnoticed in the ethical literature.<sup>2</sup> I will call it the principle of *recombinant values*.

It is the purpose of this paper to formulate the principle in a way which makes its content obvious and accessible; to motivate the principle philosophically; to both disentangle it from, and exhibit its relations to, principles of evaluative reasoning; to show how this purely qualitative principle meshes with the infamous thesis of additivity of value; and finally, to use it to ground a rather simple but quite general theory of the intrinsic value of states.

### 1. GOOD FEATURES AND BARE DIFFERENCES

Some features are good in themselves (cleverness and beauty, perhaps) and some aren’t (stupidity and ugliness, perhaps). How are we to assess what is good in itself, and what is not? The well known method of bare differences (or the contrast strategy) is one such method. It is based to the simple but appealing idea that an intrinsically good feature enhances the overall value of wholes of which it is a part, whereas an intrinsically bad feature detracts from overall value.



*Philosophical Studies* **106**: 259–292, 2001.

© 2001 Kluwer Academic Publishers. Printed in the Netherlands.

The method of value inquiry based on the idea is also straightforward. In order to determine whether or not a feature is intrinsically good, compare possibilities which differ solely in respect of the feature in question. If a possibility containing the feature is better than the corresponding possibility without the feature then, since the only difference between the two is that the better one possesses the feature in question, it must be that feature which makes the difference in value. So the feature itself must be good. Conversely, given the intrinsic value of the feature then it seems reasonable to assume that its presence anywhere must make a positive difference for the better – certainly in other barely different pairs, but presumably wherever the feature occurs. At least, that's the idea.

Evidently there are two different procedures here. Firstly there is the procedure of identifying good and bad features by comparing the value of complex wholes in which they occur. We could call this the *top-down* bare-difference procedure, since it proceeds from the comparative value of barely different wholes to the value of the parts which constitute them. Secondly once we are given an evaluative ordering of individual features we can then start ordering wholes in which those features occur. We could call this the *bottom-up* bare-difference procedure, since it enables us to evaluate wholes by considering the value of features which constitute them. Of course, the principles can also be used to refute claims of the superiority of features over their rivals.

Examples of the application of one or other of these procedures are not difficult to find, both in contemporary literature and in the history of value theory. Kant, in a well-known passage at the start of the *Metaphysical Foundations of Morals*, makes use of the bottom-up procedure in his attempt to refute the thesis of the intrinsic goodness of happiness, as well as of traditionally esteemed virtues, like moderation in the passions.<sup>3</sup> Plato (at least according to Aristotle) uses the top-down procedure to establish that pleasure is not the only intrinsic good.<sup>4</sup> More recently, both procedures have been used to undermine the moral relevance of the killing/letting-die distinction.<sup>5</sup>

This method of value inquiry and the underlying principles, have come under some sustained criticisms of late. The principles are connected to the assumption of additivity of value – that the value

of a complex state of affairs is the arithmetical sum of the value of its component parts – and such an assumption is bound to offend for two reasons: its evident simplicity as well as its presumption of numerical measurability. (Ethicists tend to have an abhorrence of numbers.)<sup>6</sup> A recent attack on the method of bare-differences (by Shelley Kagan) is called, simply, “The Additive Fallacy”.<sup>7</sup> Kagan claims that it is the assumption of additivity which makes the bare-difference approach both reasonable and attractive. He goes on to criticize the thesis of additivity by the use of various intuitive counterexamples. He thus joins a long line of critics of additivity of value and adherents of the thesis of the *organic unity* of value.<sup>8</sup>

Elsewhere I have defended additivity as a fruitful heuristic principle from the kinds of apparently intuitive counterexamples used by Kant and Moore, and more recently by Kagan, Lemos and others.<sup>9</sup> What has not been generally appreciated is that bare difference, additivity, and organic unity are notions which presuppose a *factorization* of a space of possibilities. As is almost always the case, no hypothesis stands alone in the face of a counterexample. Only in the apparently innocuous company of various auxiliaries does an hypothesis face the tribunal of intuitive counterexamples. In this case the hidden auxiliary concerns the relevant evaluative factors – of what makes *other things equal*. The apparent counterexamples often evaporate if we change to a different, often more salient, factorization.<sup>10</sup> Additivity can then be used as a heuristic device to identify salient ways of cutting up possibilities into their evaluative factors.<sup>11</sup>

In this paper I want to focus on a different issue. Additivity is, of course, a numerical concept, whereas the principles which underlie typical evaluative reasoning – like top-down and bottom-up principles of bare-difference – are purely qualitative. Two questions arise: Is there an intuitively graspable, purely qualitative condition on an evaluative ordering of possibilities that is necessary and sufficient for additivity? Further, what relations does such a condition, if it exists, bear to the intuitively compelling principles underlying the method of bare-difference?

Mathematicians have shown that principles closely related to the top-down and bottom-up principles underlying bare-difference –

which they call principles of *separability* or of *independence* – go hand in hand with additivity when it comes to magnitudes which satisfy very strong conditions of density, or continuity, or “unrestricted solvability”. (Roughly, speaking, the change in value brought about by adjustment in one magnitude can always be exactly compensated for by a suitable adjustment in any other magnitude.)<sup>12</sup> However, in normal evaluative and ethical reasoning we are usually trafficking not in continuous or dense magnitudes satisfying this rather strict requirement, but rather in a finite number of attributes, or of finite determinables, which jointly generate a finite class of complex attributes (or possible states). Thus the strong principles which are part and parcel of the standard necessary and sufficient qualitative conditions for additivity seem rather remote from our concerns.

Refining my two questions then: Is there an intuitively graspable, purely qualitative condition on an evaluative ordering of a *finite* or *discrete* set of possibilities that is necessary and sufficient for additivity. And what relations does such a condition bear to the method of bare-differences?

## 2. INDEPENDENCE AND SEPARABILITY

We can think of the complex wholes as involving a number of *factors* each of which is a determinable embracing a finite number of determinate features. Thus the intelligence factor is a determinable of which *smart*, *stupid*, or more fine-grained degrees of intelligence (*extremely smart*, *very smart*, *moderately smart*, and so on) are the determinate features. We can take the determinables associated with *beauty*, *wisdom*, *virtue*, *justice* – and anything else which comes in degrees – as such factors. Each factor is thus a collection of jointly exhaustive and mutually exclusive features. Features are *rivals* if they are incompatible determinates of such a determinable factor. Thus being smart and being stupid are rivals of the intelligence factor, being beautiful and being ugly are rivals of the beauty factor.<sup>13</sup>

A *whole* or *complex*  $V$  specifies, for each factor  $f$  a determinate feature. Each factor,  $f$ , can be associated with an  $n$ -tuple of features  $\langle f_1, f_2, f_3, \dots, f_n \rangle$  each of which is a particular determinate real-

ization of  $f$ . So with each whole  $V$  we associate an assignment of features to factors.  $V(f)$  is the feature which  $V$  assigns to  $f$ ,  $V(g)$  the feature  $V$  assigns to  $g$ , and so on. If we order the factors then each whole  $V$  is associated with sequence of features:  $V = \langle V(f), V(g), V(h), \dots \rangle$ .<sup>14</sup>

Where  $V$  is a complex which specifies some determinate feature (say, *smart*) for the intelligence factor and some determinate feature to the beauty factor, then let  $V^{stupid}$  be the complex which differs from  $V$  in *at most* that *stupid* replaces *smart*.

For example:

If  $V$  is  $\langle \text{smart}, \text{beautiful} \rangle$ ,  $V^{stupid}$  is  $\langle \text{stupid}, \text{beautiful} \rangle$ .  
If  $U$  is  $\langle \text{smart}, \text{ugly} \rangle$ ,  $U^{smart}$  is also  $\langle \text{smart}, \text{ugly} \rangle$ .

In general, let  $V^{f_k}$  be the complex which differs from  $V$  in at most that  $V$  assigns  $f_k$  to the factor  $f$ . Then  $V^{f_k}$  and  $V^{f_j}$  are a pair of wholes which differ only in that  $V^{f_k}$  assigns the feature  $f_k$  to the factor  $f$ , whereas  $V^{f_j}$  assigns the feature  $f_j$ .<sup>15</sup>

Although it would be somewhat more intuitive to work with the relation of *better than*, it will prove useful to work with the evaluative relation  $\geq$ , *as good as*, defined on a space of complex wholes. All I assume at this stage is that  $\geq$  is *connected* – for each complex whole  $V$  and  $U$ , either  $V \geq U$  or  $U \geq V$ .<sup>16</sup>

The statements of the top-down and bottom-up bare-difference principles relate the relative goodness of parts to the relative goodness of wholes, but jointly they place a single restriction on the evaluation relation  $\geq$  on complex wholes alone which I will call *independence*: if substitution of one feature for a rival renders a whole at least as good, then that substitution renders every whole at least as good.<sup>17</sup> More precisely:

#### *Independence of factor $f$*

Factor  $f$  is *independent* just in case for any  $j$  and  $k$ : if for some  $V$ ,  $V^{f_j} \geq V^{f_k}$  then for all  $V$ ,  $V^{f_j} \geq V^{f_k}$ .

We can define overall separability of the relation  $\geq$  in terms of the independence of every factor.

*Separability*

A space of complex states is *separable* just in case every factor is independent.

Consider, as an example, the Shaw anecdote. Let us suppose that both factors – intelligence and beauty – come in just two determinates. I write  $s_1$  for *smart*,  $s_0$  for *stupid*;  $b_1$  for *beautiful*,  $b_0$  for *ugly*. I will also write  $s_1 \& b_0$  for  $\langle s_1, b_0 \rangle$ . A table represents an ordering of wholes. Each row of the table represents a rank in the evaluation. Being in a higher row represents being *strictly better*.

Several orderings are compatible with the evaluative judgements implicit in the Shaw anecdote.

TABLE I

Example 1

Rank	Complex
1	$s_1 \& b_1$
2	$s_1 \& b_0$
3	$s_0 \& b_1$
4	$s_0 \& b_0$

In Example 1 we have the ordering that perhaps Shaw would have favored – being clever but ugly is more valuable than being beautiful but stupid. In Example 2 that particular judgement is reversed. In Example 3 the two complexes are ranked equal.

TABLE II

Example 2

Rank	Complex
1	$s_1 \& b_1$
2	$s_0 \& b_1$
3	$s_1 \& b_0$
4	$s_0 \& b_0$

TABLE III  
Example 3

Rank	Complex
1	$s_1 \& b_1$
2	$s_0 \& b_1$ $s_1 \& b_0$
3	$s_0 \& b_0$

Each of these three ordering satisfies separability. Take the intelligence factor, and take any complex  $V$  in any of the three tables. In each table  $s_1 \& b_1 \geq s_0 \& b_1$ , but it is not the case that  $s_0 \& b_1 \geq s_1 \& b_1$ . To satisfy independence, then, we must have that for any  $V$ ,  $V^{s_1} \geq V^{s_0}$ . But we must also have that for no  $V$ ,  $V^{s_0} \geq V^{s_1}$ . That is, we must have that for any  $V$ ,  $V^{s_1}$  is strictly better than  $V^{s_0}$  – and it is not hard to check that that is so.

What orderings are not separable? Someone might take the view that neither beauty nor intelligence add value by themselves, that only in the presence of both is there any value at all. Such an ordering has the following structure:

TABLE IV  
Example 4

Rank	Complex
1	$s_1 \& b_1$
2	$s_0 \& b_1$ $s_1 \& b_0$ $s_0 \& b_0$

$s_0 \& b_0$  is at least as good as  $s_0 \& b_1$ . For separability to hold, for any  $V$ ,  $V^{b_0} \geq V^{b_1}$ . But  $s_1 \& b_0$  is not at least as good as  $s_1 \& b_1$ . Separability fails.

## 3. INTRINSIC VALUE OF FEATURES

To see that this account of separability meshes with our top-down and bottom-up bare-difference principles we must first state the latter in our notation.

*Bottom-up bare-difference for factor f*

For all  $f_j$  and  $f_k$ , if  $f_j \geq f_k$  then, for every  $V$ ,  $V^{f_j} \geq V^{f_k}$ .

*Top-down bare-difference for factor f*

For any  $f_j$  and  $f_k$ , if for some  $V$ ,  $V^{f_j} \geq V^{f_k}$ , then,  $f_j \geq f_k$ .

To state these principles we have had to assume that not only are complexes ordered by the relation *as good as*, but so too are the individual features which constitute those complexes. So to connect independence and separability with bare difference, we have to extend the relation  $\geq$  from the complexes to the features that make up the complexes, and do so in a natural manner. I will argue that for one feature to be intrinsically as good as another the substitution of the latter by the former (holding everything else fixed, of course) must always renders things at least as good overall.

To show this, let's restrict ourselves to two degrees of value, and simply talk about intrinsic goodness and badness. The *intrinsic goodness* of a feature must in some way or other be connected with the goodness of wholes of which it is a part. The two cannot simply float freely and independently of one another. A good feature must *in some way or other* be connected with *enhancing* the goodness of the wholes of which it is a part, rather than detracting from their overall goodness. Presumably not even a Moorean advocate of organic unity would claim that something perfectly good in itself *invariably* detracts from the value of the wholes of which it is a part. What, then, is the relationship?

It is clearly not sufficient for intrinsic goodness that substituting a feature for its rival *sometimes* enhances the goodness of the whole. For that would allow a feature to be both intrinsically better than its rival, as well as intrinsically worse.

Perhaps it is sufficient that the feature enhances value under *normal conditions*. Let's discount the difficulties inherent in postu-

lating normal conditions. Still, if a feature enhances goodness under some range of conditions (but not under others) then it seems more accurate to say that what is *really* intrinsically good is not the feature itself, but rather the conjunction of the feature with those conditions.

Finally, taking a cue from decision theory, one might argue that for a feature to be intrinsically good it is necessary that *on average* it enhances overall goodness. For the goodness to be *intrinsic* the notion of average would have to be based on purely logical considerations, rather than on contingent or empirical considerations. (It cannot be just an *accident* that a feature is intrinsically good. There has to be some kind of necessity attached to it.) So we would require a robust notion of *a priori* logical probability. Then the notion could be cashed out in term of the *average* or *expected* value (utilizing logical probability) of possessing the feature in question.

There is certainly *something* important in the connection between intrinsic value of states in general, and weighted-average value. I will return to a more detailed consideration of this idea in the final section of the paper. But it is not clear that enhancing goodness on average is sufficient for the fundamental or basic notion of *intrinsic* goodness. Rather it seems that intrinsic value comes in two grades: one basic, which is compatible with, but not definable in terms of, average value; the other derived, applicable to attributes or states in general, and definable in terms of average value.

To see that there are distinct notions here note the following possibility. Suppose a feature reduces overall value in most barely-different pairs, but in conjunction with one particular constellation of other features it enormously enhances value (comparing with its barely different rival). In the calculation of average value that one case, if sufficiently valuable, might swamp the other cases. But would we want to say that it is an *intrinsically* good feature? It is arguable that the answer is no.

Note that this counterexample involves an apparent violation of separability. Granted separability it is harder to construct an intuitive counterexample involving a basic feature (one of the features which generate the set of possibilities). Given separability, a basic feature will enhance average value if and only if it enhances value everywhere. But even granted separability that is not necessarily so for non-basic features. In that case a logically complex state might

well be more of a boon on average than, say, its negation, although it might not even make sense to say that, all else being equal, the feature is better than its negation. So, for basic features, we will accept the following:

*Definitional extension of  $\geq$  to features*

For any features  $f_j$  and  $f_k$  of factor  $f$ :

$$f_j \geq f_k =_{df} \text{ for every } V, V^{f_j} \geq V^{f_k}.$$

This definitional extension of  $\geq$ , together with independence of  $f$  entails both the bottom-up and top-down principles, as promised:

Proof: The definitional extension of  $\geq$  to features yields bottom-up immediately. Now assume separability of factor  $f$ , and the antecedent of top-down: that for some  $V$ ,  $V^{f_j} \geq V^{f_k}$ . By separability of factor  $f$ , we get for all  $V$ ,  $V^{f_j} \geq V^{f_k}$ . And hence by the definitional extension of  $\geq$ ,  $f_j \geq f_k$ . This establishes top-down.

Top-down, in conjunction with the definitional extension of  $\geq$  to features, yields independence.

Note that it is by no means guaranteed that the extended relation  $\geq$  will be connected on each of the factors even though it is connected on the wholes. It turns out that the definitional extension of  $\geq$  is connected on a particular factor  $f$  just in case that factor satisfies independence.

Proof: Assume independence of  $f$ . By connectedness of the parent relation, for each  $V$  and any  $f_j$  and  $f_k$ , either  $V^{f_j} \geq V^{f_k}$  or  $V^{f_k} \geq V^{f_j}$ . If the former then, by the independence of  $f$ , we have for all  $V$ ,  $V^{f_j} \geq V^{f_k}$ , and so by the definitional extension of  $\geq$ ,  $f_j \geq f_k$ . If the latter then analogously,  $f_k \geq f_j$ . Now suppose that for some  $V$ ,  $V^{f_j} \geq V^{f_k}$ . By connectedness we know that either  $f_j \geq f_k$  or  $f_k \geq f_j$ . By the definitional extension of  $\geq$ , only the former is compatible with  $V^{f_j} \geq V^{f_k}$ . But then, also by the definitional extension of  $\geq$ , for all  $V$ ,  $V^{f_j} \geq V^{f_k}$ . That is, the factor  $f$  is independent.

#### 4. ADDITIVITY

What is it for value to be additive? The notion of additivity is numerical. It involves a numerical measure of the value of both parts and wholes. But suppose all we have is the qualitative relation – *as good as* – defined initially over complex wholes? In the absence

of numbers is there any way of cashing out talk about additivity or must it remain merely metaphorical or suggestive.

Given a non-numerical ordering,  $\geq$ , the assumption of additivity clearly involves the possibility of *representing* or *realizing* the ordering numerically – of preserving the qualitative ordering with some assignment of numbers.

### *Realization*

A real-valued function  $\Psi$  is a *realization* of a relation  $\geq$  just in case  $\Psi(V) \geq \Psi(U)$  if and only if  $V \geq U$ .

Not any old realization will do for additivity, of course. The realization has to vindicate the idea that the value of the whole is the sum of the value of its parts. So for the function  $\Psi$  to be an additive realization it must assign numbers to all the parts of wholes as well as the wholes themselves, and the value of the  $\Psi$ -value of the whole must be the sum of the  $\Psi$ -value of the parts.

Where  $P$  is the set of all constituent features (parts) which make up the different complexes, let  $\underline{V}$  be the indicator function associated with  $V$ :  $\underline{V}(p) = 1$  if  $p$  is a part of  $V$  and  $\underline{V}(p) = 0$  otherwise. We can characterize the additivity of an assignment of numerical values thus:

### *Additive functions*

A real-valued function  $\Psi$  defined on wholes and parts is *additive* just in case

$$\Psi(V) = \sum_{p \text{ a part of } V} \Psi(p) = \sum_{p \in P} \underline{V}(p) \Psi(p).$$

So value is additive if the evaluative ordering can be appropriately realized by an additive function. Otherwise it is not additive – or, as G.E. Moore would say, it exhibits *organic unity*.

### *The additivity of an evaluative ordering*

The ordering  $\geq$  is *additive* if and only if it possesses an additive realization. Otherwise it exhibits *organic unity*.

Thus the question of the additivity of an ordering boils down to this: is there a way of assigning numerical values to both wholes and component features in such a way that the relative values of

complex wholes are adequately captured by summing the values of the component features?

Consider the Shaw orderings.  $\Psi_1$ ,  $\Psi_2$ ,  $\Psi_3$  are additive realizations:

TABLE V

Example 1 :  $\Psi_1(\mathbf{s}_1) = 2$ ,  $\Psi_1(\mathbf{s}_0) = 0$ ,  $\Psi_1(\mathbf{b}_1) = 1$ ,  $\Psi_1(\mathbf{b}_0) = 0$

Rank	Complex	$\Psi_1(\text{Complex})$
1	$\mathbf{s}_1$ & $\mathbf{b}_1$	$\mathbf{3} = 2+1$
2	$\mathbf{s}_1$ & $\mathbf{b}_0$	$\mathbf{2} = 2+0$
3	$\mathbf{s}_0$ & $\mathbf{b}_1$	$\mathbf{1} = 1+0$
4	$\mathbf{s}_0$ & $\mathbf{b}_0$	$\mathbf{0} = 0+0$

TABLE VI

Example 2:  $\Psi_2(\mathbf{s}_1) = 1$ ,  $\Psi_2(\mathbf{s}_0) = 0$ ,  $\Psi_2(\mathbf{b}_1) = 2$ ,  $\Psi_2(\mathbf{b}_0) = 0$

Rank	Complex	$\Psi_2(\text{Complex})$
1	$\mathbf{s}_1$ & $\mathbf{b}_1$	$\mathbf{3} = 1+2$
2	$\mathbf{s}_0$ & $\mathbf{b}_1$	$\mathbf{2} = 0+2$
3	$\mathbf{s}_1$ & $\mathbf{b}_0$	$\mathbf{1} = 1+0$
4	$\mathbf{s}_0$ & $\mathbf{b}_0$	$\mathbf{0} = 0+0$

TABLE VII

Example 3:  $\Psi_3(\mathbf{s}_1) = 1$ ,  $\Psi_3(\mathbf{s}_0) = 0$ ,  $\Psi_3(\mathbf{b}_1) = 1$ ,  $\Psi_3(\mathbf{b}_0) = 0$

Rank	Complex	$\Psi_3(\text{Complex})$
1	$\mathbf{s}_1$ & $\mathbf{b}_1$	$\mathbf{2} = 1+1$
2	$\mathbf{s}_0$ & $\mathbf{b}_1$	$\mathbf{1} = 0+1 = 1+0$
3	$\mathbf{s}_0$ & $\mathbf{b}_0$	$\mathbf{0} = 0+0$

However, no additive function can represent the ordering in Example 4 (see Table IV).

Proof: Suppose  $\Psi$  is such an additive realization. Since  $\mathbf{s}_1 \& \mathbf{b}_1 \geq \mathbf{s}_0 \& \mathbf{b}_1$  by realization  $\Psi(\mathbf{s}_1 \& \mathbf{b}_1) \geq \Psi(\mathbf{s}_0 \& \mathbf{b}_1)$ . By additivity:  $\Psi(\mathbf{s}_1) + \Psi(\mathbf{b}_1) \geq \Psi(\mathbf{s}_0) + \Psi(\mathbf{b}_1)$ , and so  $\Psi(\mathbf{s}_1) \geq \Psi(\mathbf{s}_0)$ . Since  $\mathbf{s}_0 \& \mathbf{b}_0 \geq \mathbf{s}_1 \& \mathbf{b}_0$ , by analogous reasoning,  $\Psi(\mathbf{s}_0) \geq \Psi(\mathbf{s}_1)$ . Hence  $\Psi(\mathbf{s}_0) = \Psi(\mathbf{s}_1)$ . Similarly,  $\Psi(\mathbf{b}_0) = \Psi(\mathbf{b}_1)$ . Hence, by realization, all complexes have the same rank – contradiction.

Example 4 thus gives us an ordering exemplifying *organic unity*.

Here is an interesting objection to the sufficiency of this notion of additivity.<sup>18</sup> Imagine a theory of the good which yields the evaluative ordering in Example 1, but according to which the complex  $\mathbf{s}_1 \& \mathbf{b}_1$  is immensely more valuable than either of the features  $\mathbf{s}_1$  and  $\mathbf{b}_1$ . Intuitively, on this theory, the value of  $\mathbf{s}_1 \& \mathbf{b}_1$  is *much* greater than the sum of the values of  $\mathbf{s}_1$  and  $\mathbf{b}_1$ . Wouldn't we say that this theory is *not* additive? Rather, it seems to exhibit organic unity. Nevertheless my definition seems to judge the theory to be additive (at least restricted to this example) since its evaluative *ordering* can be realized by an additive function. If this is right, the existence of a merely order-preserving additive function is insufficient to ensure the additivity of value in the intuitive sense.

We must distinguish between the background theory of value ( $T$ ); what the theory  $T$  says about the four complexes in the example ( $T_O$ ); and the particular evaluative ordering at issue in Example 1 ( $O$ ).  $T$  implies  $T_O$  which in turn implies  $O$ . But  $O$  clearly does not entail  $T$  and neither does  $O$  entail  $T_O$  since nothing in  $O$  yields information about the assumed evaluative *distance* between the components and the whole. Any numerical representation of  $T$ 's evaluations will have to reflect its judgement that the value of  $\mathbf{s}_1 \& \mathbf{b}_1$  is immensely more valuable than either of the features  $\mathbf{s}_1$  and  $\mathbf{b}_1$ , and so any such representation will have to be non-additive. For example, suppose that we have the following assignments:

$$\begin{array}{lcl}
 T_O & \Psi(\text{smart \& beautiful}) & = 10 \\
 & \Psi(\text{smart \& ugly}) & = 3 \\
 & \Psi(\text{stupid \& beautiful}) & = 2 \\
 & \Psi(\text{stupid \& ugly}) & = 1
 \end{array}$$

There is no way we can extend this assignment to an additive realization of the ordering  $O$ , as can be easily checked. This shows that

$T$  is a non-additive theory of value, because  $T$  yields judgements which cannot be represented by an additive function. However it does nothing to undermine the claim that the ordering  $O$ , which follows from  $T_O$ , is additive.

Finally, note that *additivity entails separability*

Proof: Assume an ordering  $\geq$  of complex wholes is additive – that is, it has an additive realization,  $\Psi$ . What we have to show that for any factor  $f$ , and any features  $f_j$  and  $f_k$ : if for some  $V$ ,  $V^{f_j} \geq V^{f_k}$  then for all  $V$ ,  $V^{f_j} \geq V^{f_k}$ . So assume  $V^{f_j} \geq V^{f_k}$ . By the fact the  $\Psi$  is a realization of the ordering  $\geq$ , it follows that  $\Psi(V^{f_j}) \geq \Psi(V^{f_k})$ . Let  $w$  be the sum of the  $\Psi$ -values of the features in  $V$  other than its  $f$ -feature. Since  $\Psi$  is additive,  $\Psi(V)$  is the sum of the component features in  $V$ . Hence  $\Psi(V^{f_j}) = \Psi(f_j) + w$  and (since  $V^{f_j}$  differs from  $V^{f_k}$  only in the  $f$ -factor)  $\Psi(V^{f_k}) = \Psi(f_k) + w$ . Hence  $\Psi(f_j) \geq \Psi(f_k)$ . For any  $U$ ,  $\Psi(U^{f_j}) = \Psi(f_j) + u$ , and  $\Psi(U^{f_k}) = \Psi(f_k) + u$ , and so  $\Psi(U^{f_j}) \geq \Psi(U^{f_k})$ . Since  $\Psi$  is a realization of the ordering  $\geq$  it follows that  $U^{f_j} \geq U^{f_k}$ .

Separability is thus necessary for additivity, but is it sufficient?

#### 5. SEPARABILITY TOO WEAK

Other things being equal, smartness, courage and power all seem to be good features to have – at least, better than lacking them.<sup>19</sup> But consider, the relative merits of being courageous but not smart, on the one hand, and smart but not courageous on the other. Being courageous has a kind of nobility to it which smartness lacks. On the other hand, if you are courageous but lack the power to carry out your courageous resolves, your courage will serve you badly indeed. So it is arguable that in the presence of power it is better to be nobly courageous at the cost of sacrificing smartness; but in the absence of power better to be smart than stupidly courageous. If this is right then the relative merits of smartness and courage can depend on the presence and absence of power. At least, grant these judgements for the sake of an argument.

We have here an example of non-additivity of factors which is nevertheless compatible with the individual independence of factors. To see this explicitly the reader is invited to contemplate the extension of the ordering given in Table VIII ( $s_1$  for *smart*,  $s_0$  for *not-smart*, etc.).

Best of all is to have the three features. And it is not hard to see that any complex which embraces any one of the three features is better than its barely different counterpart. We have granted, however, that in the presence of power it is better to be courageous and lack smartness rather than vice-versa. Brute power without either smartness or courage seems rather lowly, even if better than being deprived of all three goods – plausibly the worst state of all to be in. If you are going to be without power then both courage and smartness have something going for them. Smartness is clearly a boon, and courage, while it has its downside for those without power, is at least ennobling. However, if forced to choose between smartness and courage in the absence of power, it seems wise to choose smartness. Some of the details here are not important. For example, we could exchange the relative positions of  $s_0 \& c_1 \& p_0$  and  $s_0 \& c_0 \& p_1$  without altering the conclusions.

TABLE VIII  
Example 5

Rank	Complex
1	$s_1 \& c_1 \& p_1$
2	$s_0 \& c_1 \& p_1$
3	$s_1 \& c_0 \& p_1$
4	$s_1 \& c_1 \& p_0$
5	$s_1 \& c_0 \& p_0$
6	$s_0 \& c_1 \& p_0$
7	$s_0 \& c_0 \& p_1$
8	$s_0 \& c_0 \& p_0$

Each factor in this attribute space satisfies independence. Each of the positive features – smart, courageous, powerful – satisfies the condition that, other things being equal, it is better to have it than to lack it. But the presence or absence of power makes a difference in the ordering of the complex attributes: courage without smartness and smartness without courage. Furthermore it is clear that this renders the ordering non-additive.

Proof: Suppose  $\Psi$  is an additive realization of the ordering. Since  $(s_0 \& c_1 \& p_1)$  is better than  $(s_1 \& c_0 \& p_1)$ , by additive realization:  $\Psi(s_0) + \Psi(c_1) + \Psi(p_1) > \Psi(s_1) + \Psi(c_0) + \Psi(p_1)$ . Hence (\*)  $\Psi(s_0) + \Psi(c_1) > \Psi(s_1) + \Psi(c_0)$ .  $(s_0 \& c_1 \& p_0)$  is worse than  $((s_1 \& c_0 \& p_0)$ , so by additive realization:  $\Psi(s_0) + \Psi(c_1) + \Psi(p_0) < \Psi(s_1) + \Psi(c_0) + \Psi(p_0)$ , and hence: (\*\*)  $\Psi(s_0) + \Psi(c_1) < \Psi(s_1) + \Psi(c_0)$ , contradicting (\*).

Separability, as characterized, evidently does not guarantee additivity.<sup>20</sup>

## 6. STRONG SEPARABILITY STILL TOO WEAK

It is not hard to see a way of strengthening our interpretation of the separability condition which would rule out this kind of counterexample. What we need to do is to extend the treatment applied to individual factors to complex factors. Intuitively, the following are complex features which may be ranked for goodness: *smart & courageous*; *smart & not-courageous*; *not-smart & courageous* . . . and so on. Two complex features are *rivals* of the sample complex factor if they assign features to the same factors. So *smart & courageous* is a rival of *smart & not-courageous*, but not a rival of *smart & not-powerful*. Now we can simply apply the definitions developed in section 2 to complex factors and features as well as to single factors. We can say that an ordering is *strongly separable* if every complex factor is independent in the modified sense. That is to say, suppose that  $g_j$  and  $g_k$  are complex rivals. Then if there is a pair  $V$  and  $U$  which differ only over  $g_j$  and  $g_k$  and  $V$  is better than  $U$ , then every such pair is similarly ordered.

We can now see that the ordering above does not satisfy strong separability.  $(s_0 \& c_1 \& p_1)$  and  $(s_1 \& c_0 \& p_1)$  differ only with respect to the complex factor <courage, smartness>, as do  $(s_0 \& c_1 \& p_0)$  and  $(s_1 \& c_0 \& p_0)$ , in the very same way. But the two pairs clearly violate the condition of strong separability.  $(s_0 \& c_1 \& p_1)$  is better than  $(s_1 \& c_0 \& p_1)$  but  $(s_0 \& c_1 \& p_0)$  is worse than  $(s_1 \& c_0 \& p_0)$ .

Is strong separability sufficient for additivity? Consider Example 6 below (Table IX) of a possible ordering of three kinds of goods, each of which comes in three degrees. Note that an improvement in the degree of each factor is, other things being equal, an improve-

ment overall. It is not hard to check that if you raise the degree of any one of the factors while holding the other two factors fixed, then you move up the ordering. So each degree of any one factor is better than a lower degree (i.e. we have weak separability). Further, if a certain combination of, say, knowledge and goodness of will is better than another combination, then the inequality holds regardless of degree of power, although this is not so easy to ascertain. If any one factor is held fixed (say, degree of power) then the ordering of the remaining complex factor (of knowledge-and-goodness) is the same whichever degree of power we have picked (i.e. we also have strong separability).

The ordering contains the following three inequalities:

$$\begin{aligned} (g_2 \ \& \ k_3 \ \& \ p_3) &> (g_3 \ \& \ k_2 \ \& \ p_2) \\ (g_3 \ \& \ k_1 \ \& \ p_2) &> (g_2 \ \& \ k_3 \ \& \ p_1) \\ (g_1 \ \& \ k_2 \ \& \ p_1) &> (g_1 \ \& \ k_1 \ \& \ p_3). \end{aligned}$$

By realizability:

$$\begin{aligned} \Psi(g_2 \ \& \ k_3 \ \& \ p_3) &> \Psi(g_3 \ \& \ k_2 \ \& \ p_2) \\ \Psi(g_3 \ \& \ k_1 \ \& \ p_2) &> \Psi(g_2 \ \& \ k_3 \ \& \ p_1) \\ \Psi(g_1 \ \& \ k_2 \ \& \ p_1) &> \Psi(g_1 \ \& \ k_1 \ \& \ p_3). \end{aligned}$$

The sum of the terms on the left is thus greater than the sum of the terms on the right:

$$\begin{aligned} \Psi(g_2) + \Psi(k_3) + \Psi(p_3) + \Psi(g_3) + \Psi(k_1) + \Psi(p_2) + \Psi(g_1) + \Psi(k_2) + \Psi(p_1) &> \\ \Psi(g_3) + \Psi(k_2) + \Psi(p_2) + \Psi(g_2) + \Psi(k_3) + \Psi(p_1) + \Psi(g_1) + \Psi(k_1) + \Psi(p_3). \end{aligned}$$

However, by rearranging terms we find both sides are identical to:

$$\Psi(g_3) + \Psi(g_2) + \Psi(g_1) + \Psi(k_3) + \Psi(k_2) + \Psi(k_1) + \Psi(p_3) + \Psi(p_2) + \Psi(p_1).$$

Hence strong separability is insufficient for additivity.<sup>21</sup>

## 7. RECOMBINANT VALUES

Recall that what we are after is a purely qualitative constraint on the *as good as* relation on complex wholes which is necessary and sufficient for the ordering to be additive. We begin by noting a condition which additivity clearly entails – the principle of equivalence. It is

TABLE IX

Example 6

Rank	Complex
1	$g_3 \& k_3 \& p_3$
2	$g_3 \& k_3 \& p_2$
3	$g_3 \& k_2 \& p_3$
4	$g_2 \& k_3 \& p_3$
5	$g_3 \& k_2 \& p_2$
6	$g_3 \& k_3 \& p_1$
7	$g_2 \& k_3 \& p_2$
8	$g_2 \& k_2 \& p_3$
9	$g_3 \& k_2 \& p_1$
10	$g_3 \& k_1 \& p_3$
11	$g_2 \& k_2 \& p_2$
12	$g_3 \& k_1 \& p_2$
13	$g_2 \& k_3 \& p_1$
14	$g_2 \& k_2 \& p_1$
15	$g_2 \& k_1 \& p_3$
16	$g_3 \& k_1 \& p_1$
17	$g_2 \& k_1 \& p_2$
18	$g_2 \& k_1 \& p_1$
19	$g_1 \& k_3 \& p_3$
20	$g_1 \& k_3 \& p_2$
21	$g_1 \& k_2 \& p_3$
22	$g_1 \& k_2 \& p_2$
23	$g_1 \& k_3 \& p_1$
24	$g_1 \& k_2 \& p_1$
25	$g_1 \& k_1 \& p_3$
26	$g_1 \& k_1 \& p_2$
27	$g_1 \& k_1 \& p_1$

a principle which the above counterexample to additivity violates. However the equivalence principle is not purely qualitative and so will not serve as a qualitative constraint sufficient for additivity. However it in turn entails a qualitative condition – which I will call *the principle of recombinant values* – which is thus necessary for additivity.<sup>22</sup>

In the above counterexample three inequalities gave rise to the problem. These three inequalities define two groups of complexes with an interesting feature:

superior group		inferior group
(g <sub>2</sub> & k <sub>3</sub> & p <sub>3</sub> )	>	(g <sub>3</sub> & k <sub>2</sub> & p <sub>2</sub> )
(g <sub>3</sub> & k <sub>1</sub> & p <sub>2</sub> )	>	(g <sub>2</sub> & k <sub>3</sub> & p <sub>1</sub> )
(g <sub>1</sub> & k <sub>2</sub> & p <sub>1</sub> )	>	(g <sub>1</sub> & k <sub>1</sub> & p <sub>3</sub> )

If you look down the columns you find that each feature occurs in each *group* of complexes exactly once. (That’s what gave us the contradiction.) Thus the uniformly inferior group can be constructed out of exactly the same elements from which the superior group is constructed. In such a case we say that the members of the second group are a *recombination* of the first.

This is just a special case of recombination. For recombination in general we do not require that all the features occur in each of the groups. Nor do we require that each feature occurs at most once in a group. Each complex assigns a feature to each factor. Imagine we start with a group of *n* complex wholes – U<sub>0</sub>, U<sub>1</sub>, . . . , U<sub>n-1</sub>. We note how many times each feature occurs in this group of complexes. Now consider another (possibly overlapping) group of *n* complexes – V<sub>0</sub>, V<sub>1</sub>, . . . , V<sub>n-1</sub> – and suppose that each feature occurs exactly the same number of times in this group as in the former. We can take the building blocks of which the first group of complexes are composed and recombine them into the complexes of the second group. Then we say that the latter group arises from the *recombination* of the elements of the first group.

*Recombination*

V<sub>0</sub>, . . . , V<sub>n-1</sub> are a *recombination* of U<sub>0</sub>, . . . , U<sub>n-1</sub> if and only if the total number of times any feature occurs in the first set of possibilities is the same as the total number of times it occurs in the second set of possibilities.

If additivity holds then each feature has a certain value independently of the configuration of other features in which it occurs. Thus additivity implies that the sum of the values of the first group must be the same as the sum of the values of the second, recombinant group.

### *Equivalence*

If  $V_0, V_1, \dots, V_{n-1}$  are a recombination of  $U_0, U_1, \dots, U_{n-1}$  then

$$\Psi(V_0) + \dots + \Psi(V_{n-1}) = \Psi(U_0) + \dots + \Psi(U_{n-1})$$

If we had a straightforward qualitative equivalent of the numerical sum of the values of a group of complexes, then this principle would be tantamount to a qualitative condition. Unfortunately there is no such straightforward equivalent. Nevertheless it is not hard to see that this principle in turn entails a condition that can be expressed qualitatively. Call  $V_i$  the *counterpart* of  $U_i$ . Then suppose that for all the  $V_i$  except  $V_0$ ,  $V_i$  is at least as good as its counterpart  $U_i$  (that is:  $\Psi(V_i) \geq \Psi(U_i)$ ). Then since the sum of the the values of both groups must be the same it follows that  $U_0$  cannot be worse than its counterpart  $V_0$  – that is,  $\Psi(U_0) \geq \Psi(V_0)$  – for if  $U_0$  were strictly worse than  $V_0$  the left hand side of the above equality would be strictly less than the right hand side.  $\Psi(U_0)$  has to be at least as large as  $\Psi(V_0)$  to *compensate* for all the other inequalities of the form:  $\Psi(V_i) \geq \Psi(U_i)$ . So the equivalence principle entails what might be called the *compensation* principle:

### *Compensation*

If  $V_0, V_1, \dots, V_{n-1}$  are a recombination of  $U_0, U_1, \dots, U_{n-1}$  then

if  $\Psi(V_i) \geq \Psi(U_i)$  for all  $i > 0$ , then  $\Psi(U_0) \geq \Psi(V_0)$ .

Now by simply translating these numerical inequalities into the corresponding qualitative conditions the Compensation Principle yields:

### *Recombinant Values*

**If  $V_0, V_1, \dots, V_{n-1}$  are a recombination of  $U_0, U_1, \dots, U_{n-1}$   
then if  $V_i \geq U_i$  for all  $i > 0$  then  $U_0 \geq V_0$ .**

## 8. CONSEQUENCES OF RECOMBINANT VALUES

The content of the principle of recombinant values is thus not hard to grasp. Before stating the main result of this section I want to work through a few simple applications of the principle.

*Shaw's admirer*

Return to the little story with which I began. Shaw and his admirer each exemplified two ways of distributing the four features: beautiful, ugly; smart, stupid. Shaw was ugly & smart, his admirer beautiful & stupid. Shaw's admirer implicitly endorsed two plausible judgments:

- i beautiful & smart is better than ugly & smart;
- ii beautiful & smart is better than beautiful & stupid.

Shaw's query concerned the relative merits of being ugly & stupid. Note that the first pair of complexes are a recombination of the second:

$V_0$ ugly & stupid,	$V_1$ beautiful & smart
$U_0$ beautiful & stupid,	$U_1$ ugly & smart

So given the principle of recombinant values, judgement i ( $V_1 > U_1$ ) yields  $U_0 > V_0$ : that is,

- iii ugly & stupid is worse than beautiful & stupid;

Similarly, judgement ii yields:

- iv ugly & stupid is worse than ugly & smart.

Thus we do not need to add further judgements to draw Shaw's undesirable conclusion. Nor do we need to make a complex excursion through a valuation of the individual component features. All one needs are two evaluative judgements ranking the complex wholes, together with recombinant values, and the rest follows inexorably.

*Transitivity of  $\geq$* 

You will recall that I placed only one constraint on the evaluative relation *as good as* – that it is connected – and left other desirable properties unspecified. It turns out that we can recover the usual properties from connectedness and recombinant values. We get reflexivity (for all  $V$ ,  $V \geq V$ ) immediately from connectedness. And, somewhat surprisingly we get transitivity itself immediately from recombinant values.

Proof suppose  $V \geq U$  and  $U \geq W$ .  $W, V, U$  are a recombination of  $V, U, W$  – trivially, since we have here the very same set of possibilities reordered. Hence by recombinant values, since  $V \geq U$  and  $U \geq W$ , we must have  $V \geq W$ .

The transitivity of *better-than* and *same-value* follows from this together with the usual definitions of  $\geq$  and  $\approx$ . That a fundamental feature of value should flow so readily from recombinant values is a pleasant bonus.

*Recombinant values and the bare-difference principles*

If recombinant values turns out to be tantamount to additivity (as is the case) then of course it entails separability and the associated bare-difference principles. However, it is interesting to see more directly how recombinant values immediately subsumes both weak and strong separability. What we need to show is that recombinant values delivers the independence of each factor. Top-down and bottom-up then follow immediately via the definitional extension of  $\geq$  to features.

Proof: Suppose that for some possibility  $V$ ,  $V^{f_j} \geq V^{f_k}$ . Consider  $U^{f_j}$  and  $U^{f_k}$ , for any other possibility  $U$ .  $U^{f_k}$  contains the same elements as  $U^{f_j}$  except that  $f_j$  occurs in the latter where  $f_k$  occurs in the former. Similarly,  $V^{f_j}$  contains the same elements as  $V^{f_k}$  except that  $f_k$  occurs in the latter where  $f_j$  occurs in the former. Hence the pair  $U^{f_k}, V^{f_j}$  are a recombination of the pair  $U^{f_j}, V^{f_k}$ . Since  $V^{f_j} \geq V^{f_k}$  by recombinant values we must have  $U^{f_j} \geq U^{f_k}$ . Thus for all  $V$ ,  $V^{f_j} \geq V^{f_k}$ . *Strong separability* is essentially the same provided  $f_j$  and  $f_k$  are interpreted as complex rivals.

We can see how the smartness-courage-power ordering violates recombinant values. Recall that the ordering included the following:

- (i)  $(s_0 \& c_1 \& p_1) > (s_1 \& c_0 \& p_1)$
- (ii)  $(s_1 \& c_0 \& p_0) > (s_0 \& c_1 \& p_0)$

$(s_0 \& c_1 \& p_1)$ ,  $(s_1 \& c_0 \& p_0)$  are a recombination of  $(s_1 \& c_0 \& p_1)$ ,  $(s_0 \& c_1 \& p_0)$ . From (i)  $(s_0 \& c_1 \& p_1) \geq (s_1 \& c_0 \& p_1)$  so by recombinant values we must have  $(s_0 \& c_1 \& p_0) \geq (s_1 \& c_0 \& p_0)$ , contradicting (ii).

It should be noted that recombinant values is a much stronger constraint on an evaluative ordering than that imposed by the method of bare difference alone – certainly bare-difference applied to single features (weak separability), but also bare difference applied to complexes of features (strong separability). Minimally, for the application of purely negative bare difference arguments, we require the bottom-up bare difference principle: that if substituting a feature for a rival does not always make things better overall, then the feature in question is not bearer in itself than the rival feature. This thesis is just one half of the definitional extension of  $\geq$  from wholes to parts. We have persuasive arguments for this account of intrinsic value for features *regardless* of the structure of the relation of *as good as* on wholes. So bottom-up bare-difference is actually weaker than the definitional extension of  $\geq$  from wholes to parts; considerably weaker than full separability (both weak and, of course, strong); and weaker still than recombinant values.

*Recombination and the additivity of value*

We can now state the main result of this paper:

***The principle of recombinant values is necessary and sufficient for the additivity of value.***

*PROOF:* Suppose the ordering  $\geq$  on  $\Omega$  is additive, and let  $\Psi$  be an additive realization of  $\geq$ . Suppose that  $V_0, \dots, V_{n-1}$  is a recombination  $U_0, \dots, U_{n-1}$  and that  $V_i \geq U_i$  for  $0 < i < n$ . Given that  $V_0, \dots, V_{n-1}$  are a recombination of  $U_0, \dots, U_{n-1}$  the sum of the values of the  $V_i$  must be the same as the sum of the values of the  $U_i$ :

$$(*) \Psi(V_0) + \dots + \Psi(V_{n-1}) = \Psi(U_0) + \dots + \Psi(U_{n-1}).$$

Since  $V_i \geq U_i$  for all  $i > 0$ , by realization,  $\Psi(V_i) \geq \Psi(U_i)$  for all  $i > 0$ , and so to preserve the equality (\*),  $\Psi(U_0) \geq \Psi(V_0)$ . Hence, by realization, we have  $U_0 \geq V_0$ , thereby establishing recombinant values.

Now suppose we have a space of complexes  $\Omega$  with a connected evaluative ordering  $\geq$  satisfying the principle of recombinant values. First, a result in linear algebra due to Dana Scott.<sup>23</sup> Let  $L$  be a finite dimensional real linear vector space. A vector is *rational* if it has rational coordinates. A subset  $S$  of  $L$  is rational if all its members are rational. Let  $\geq$  be a binary relation on  $S$ . We say  $\geq$  is *realizable*

if there exists a linear functional  $\varphi$  on  $L$  such that for all  $x, y \in Y$ :  $x \geq y$  if and only if  $\varphi(x) \geq \varphi(y)$ .

*LEMMA*(Scott) Let  $S$  be a finite rational subset of  $L$ . For a binary relation  $\geq$  on  $S$  to be realizable it is necessary and sufficient that the following two conditions hold:

- (i) for all  $x, y \in S$ ,  $x \geq y$  or  $y \geq x$ ,
  - (ii) for all sequences  $x_0, \dots, x_{n-1}, y_0, \dots, y_{n-1} \in S$ ,
- if  $x_i \geq y_i$  for  $0 < i < n$  and  $n > 0$ ,

and

$$\sum_{i < n} x_i = \sum_{i < n} y_i \quad \text{then } y_0 \geq x_0.$$

Each member of  $\Omega$  is an  $n$ -tuple of features, one feature from each factor ( $f, g, h, \dots$ ). So each complex  $V$  is associated with a function from the class of all features  $P$  into the set  $\{0, 1\}$ . Let  $\underline{\Omega}$  be the set of all the indicator functions of members of  $\underline{\Omega}$ :  $\underline{V}$  is in  $\underline{\Omega}$  just in case  $V$  is in  $\underline{\Omega}$  (where for each  $p \in P$ ,  $\underline{V}(p)=1$  if  $p$  is a part of  $V$  ( $V$  assigns  $p$  to the factor of which  $p$  is a determinate feature) and  $\underline{V}(p)=0$  otherwise).  $\underline{\Omega}$  is a finite rational subset of the standard  $|P|$ -dimensional linear space  $L(P)$  of functions from  $P$  into the reals. We will call the vectors in  $\underline{\Omega}$  *possibility-vectors*. Let  $\underline{V} \geq \underline{U}$  just in case  $V \geq U$ . Since  $\geq$  is connected on  $\underline{\Omega}$  we have condition (i) of Scott's lemma: for all  $\underline{V}, \underline{U} \in \underline{\Omega}$  either  $\underline{V} \geq \underline{U}$  or  $\underline{U} \geq \underline{V}$ . The vector sum  $\underline{V}_0 + \dots + \underline{V}_{n-1}$  of the possibility-vectors  $\underline{V}_0, \dots, \underline{V}_{n-1}$ , assigns to each  $p \in P$  the number of times  $p$  occurs as a part of the complexes  $\underline{V}_0, \dots, \underline{V}_{n-1}$ . Hence  $\underline{V}_0, \dots, \underline{V}_{n-1}$  are a recombination of  $\underline{U}_0, \dots, \underline{U}_{n-1}$  if and only if:  $\underline{V}_0 + \dots + \underline{V}_{n-1} = \underline{U}_0 + \dots + \underline{U}_{n-1}$ . Condition (ii) of Scott's lemma then follows immediately from the principle of recombinant values. So there exists a realization of  $\geq$ : a linear functional  $\varphi$  on  $L(P)$  such that for all  $\underline{V}, \underline{U} \in \underline{\Omega}$ :  $\underline{V} \geq \underline{U}$  if and only if  $\varphi(\underline{V}) \geq \varphi(\underline{U})$ . To show that  $\varphi$  yields an additive realization  $\Psi$  of  $\geq$ , note that each feature  $p$  in  $P$  can be associated with a base vector  $\underline{b}_p$  of  $L(P)$ :  $\underline{b}_p$  assigns 1 to  $p$  and 0 to all the other members of  $P$ . Since

$$\underline{V} = \sum_{p \in P} \underline{V}(p) \underline{b}_p$$

by the linearity of  $\varphi$

$$\varphi(\underline{V}) = \sum_{p \in P} \underline{V}(p) \varphi(\underline{b}_p).$$

For each  $V$ , let  $\Psi(V)=\varphi(\underline{V})$ , and for each feature  $p$ , let  $\Psi(p)$  be  $\varphi(\underline{b}_p)$ . Then:

$$\Psi(V) = \sum_{p \in P} \underline{V}(p)\Psi(p). = \sum_{p \text{ a part of } V} \Psi(p)$$

$\Psi$  is clearly the required additive realization of  $\geq$ .

Recall that for simplicity we originally assumed logical independence of all factors. This is a restriction which we can now lift. Logical dependence means that the space  $\Omega$  of complexes is “missing” certain specifications of features because they are not genuinely possible. Let  $\Omega$  be such a space, and let  $\geq_{\Omega}$  be the associated relation. Let  $\Omega^+$  be the set of *all* (combinatorial) specifications over the same set of factors, whether or not the specification is genuinely possible. Let us say that a relation  $\geq_{\Omega^+}$  on  $\Omega^+$  is a *full extension* of  $\geq_{\Omega}$  if  $\geq_{\Omega}$  is the restriction of  $\geq_{\Omega^+}$  to  $\Omega$ . Finally,  $\geq_{\Omega^+}$  satisfies the principle of recombinant values (RV) iff there is a full extension of  $\geq_{\Omega}$ , satisfying RV. Now it is not difficult to show that RV is necessary and sufficient for additivity. For if  $\geq_{\Omega}$  satisfies RV then a full extension  $\geq_{\Omega^+}$ , satisfies RV and there will be an additive realization  $\Psi^+$  of the latter. The restriction of  $\Psi^+$  to  $\Omega$  will be an additive realization of  $\geq_{\Omega}$ . Now suppose  $\geq_{\Omega}$  is additive, and that  $\Psi$  is an additive realization of it. Since  $\Psi$  assigns values to all the features in all factors, we can extend  $\Psi$  to an additive assignment  $\Psi^+$  to all complexes in  $\Omega^+$  by making the value of each of the complexes in  $(\Omega^+ - \Omega)$  the sum of the  $\Psi$ -values of their component features. This induces an additive relation  $\geq_{\Omega^+}$  on  $\Omega^+$  which, by the main theorem, satisfies RV. Since  $\geq_{\Omega^+}$  is clearly a full extension of  $\geq_{\Omega}$  this latter also satisfies RV.

9. NON BASIC FEATURES: INTRINSIC VALUE AND EXPECTED VALUE

Given the principle of recombinant values we are assured a way of additively representing the relation between simple parts and complex wholes. But there are states other than simples, and complexes concatenated of simples. There are also the states which can be specified by the Boolean operations of disjunction and negation of both simples and complexes. A complete account of value

should encompass such states, and it should mesh appropriately with the account given so far. To put this another way, suppose we start with an ordering on complexes which satisfies recombinant values. We can then extend the ordering to the simples which make up the complexes. Can we further extend the ordering in a natural and principled way to all the relevant logical sums and products?

Consider Shaw's little example again, and in particular Example 1. *Being clever* and *being beautiful* are both good, and *being ugly* and *being stupid* are both bad. But what of the value of features like: *being beautiful or smart*, *being beautiful or stupid*, and *being ugly or stupid*. Presumably these should diminish in value, in that order. However nothing in what we have said so far enables us to capture those judgements.

A powerful principle, suggested by decision theory, is that the value of any state is the expected value (the weighted-average value) of the complex states which realize it. That value just is expected value of realization can be derived from some really rather weak and plausible principles.<sup>24</sup> In section 3 I presented this expectation thesis as a possible alternative to the account of intrinsic value of features motivated by considerations of independence and separability (that for a feature to be more valuable it must make a positive difference to value everywhere).

The expectation thesis presupposes that we have already been given the values of complete, complex states of affairs. Incomplete states are those which can be realized by more than one complete state. According to the expectation thesis, the value of an incomplete state S is just the probabilistically weighted average of the value of the complete states which realize it, conditional upon the truth of S. For the value so defined to be *intrinsic* it would have to be a matter of necessity. For example, we would not want the intrinsic value of a state to change on incoming information. Nor would we want it to be contingent upon, say, the structure of propensities in quantum mechanics. This suggests that the relevant probability measure would have to be a matter of necessity – that is, something *like* logical probability.

Suppose that we begin with a space of possibilities which satisfies recombinant values, and for which there is an obvious realization: e.g. Example 1 (Table V).

If we had some sort of measure of logical probability  $\mathbf{P}$  then it would be natural to extend the valuation of complete states to incomplete states by the usual formula for expected value:

*Expectation thesis*

Where  $\mathbf{P}$  is a measure of logical probability and  $S$  an arbitrary state, the intrinsic  $\Psi$ -value of  $S$  is just the expected intrinsic value of the complete states which realize  $S$ :

$$\text{Value}_\Psi(S) = \text{df } \sum_V \mathbf{P}(V|S)\Psi(V).$$

Note that whatever  $\mathbf{P}$  is, so long as  $\mathbf{P}(V|V)=1$ , the intrinsic  $\Psi$ -value of  $V$  turns out to be identical to its given  $\Psi$ -value.<sup>25</sup> That's obviously right. But we have a problem. If we apply the expectation thesis to the basic features themselves the value of the feature is *not* necessarily its  $\Psi$ -value. In Example 1, assuming that  $\mathbf{P}$  assigns equal weight to each complex,

$$\begin{aligned} \text{Value}_\Psi(s_1) &= 2.5 \neq \Psi_1(s_1) = 2; \\ \text{Value}_\Psi(s_0) &= 0.5 \neq \Psi_1(s_0) = 0, \text{ and so on.} \end{aligned}$$

The question then arises of whether the account of intrinsic value of basic states derived from recombinant values can be reconciled with this more general thesis of intrinsic value derived from expectations. It would be if each additive realization  $\Psi$  could be rescaled in a natural and uniform way to another additive realization  $\Psi^*$ , such that for any state, basic or non-basic, the intrinsic value of that state is its expected  $\Psi^*$ -value, and for all basic states, intrinsic value is identical to  $\Psi^*$ -value.

*Reconciliation*

$$\text{For any } f \text{ and } i: \Psi^*(f_i) = \sum_V \Psi^*(V)\mathbf{P}(V|f_i).$$

Take our additive realization  $\Psi_1$ . This might already strike you as unsatisfactory as a numerical measure of intrinsic value for two connected reasons.

Firstly, if *being smart* or *being beautiful* adds to overall value, then *being stupid*, or *being ugly*, must *detract* from overall value. Those features which enhance overall value should be assigned a positive real value, and those which detract, should be assigned a

negative value. But how could such an assignment, reflecting absolute levels of good and evil, be arrived at non-arbitrarily simply on the basis of the initial ordering?

Secondly, a factor – the whole range of possible realizations – should be of neither positive nor negative value. For each complex necessarily realizes the factor in some way or other, and necessity should favor neither good nor evil. This is closely connected to the first point. For if all the determinates of a determinable have non-negative value, then necessity must favor the good.

Call an additive realization which satisfies these two related strictures, as well as reconciliation, an *absolute* realization – the idea being that it captures the structure of absolute good and evil, and not just of the relation of better-than. It turns out that we can naturally transform any additive realization  $\Psi$  of a relation into an absolute realization  $\Psi^*$  of the same relation with generally pleasing properties. There is just one proviso, a *factor-neutrality* constraint on probability: information about one factor yields no information about any other factor. All the determinables are, from the logical point of view, probabilistically independent of one another. That seems a reasonable constraint to place on *logical* probability.

#### *Factor neutrality*

For all  $f_i$  and  $g_j$  ( $f \neq g$ ):  $\mathbf{P}(f_i|g_j) = \mathbf{P}(f_i)$ .

We begin with the idea that from the purely logical point of view neither good nor evil starts with an advantage. The space as a whole is axiologically neutral – it has zero expected value. It follows that each determinable should be, from the purely logical point of view, axiologically neutral.

Let  $f^\Psi$  be the average  $\Psi$ -value of the determinates of  $f$ .  
 $f^\Psi = \text{df } \sum_i \Psi(f_i)\mathbf{P}(f_i)$

To satisfy neutrality we need to rescale  $\Psi$  so that this expectation is always 0 for all factors. So let's do just that:

For each  $f$  and  $i$ :  $\Psi^*(f_i) = \text{df } \Psi(f_i) - f^\Psi$ .

Further, let us define an additive assignment of values to complexes using the new scale  $\Psi^*$ :

For each complex  $V$ :  $\Psi^*(V) = \text{df} \sum_f \Psi^*(V(f))$

Thus  $\Psi^*(V)$  uniformly scales  $\Psi(V)$  down by the sum of the average  $\Psi$ -values of all the factors:

(#)  $\Psi^*(V) = \Psi(V) - \sum_f f^\Psi$ .

It is obvious that if  $\Psi$  is a realization of  $\geq$  so too is  $\Psi^*$ , since, by (#),  $\Psi^*$  must preserve the ordering of complexes induced by  $\Psi$ . So, in Example 1, assuming that our measure of logical probability assigns equal weight to each complex, we have that  $\mathbf{s}^{\Psi_1}$  is 1, and  $\mathbf{b}^{\Psi_1}$  is 0.5. This yields the following  $\Psi_1^*$ -values of the complexes (Table X):

TABLE X

Example 1:  $\Psi_1^*(\mathbf{s}_1) = 1$ ,  $\Psi_1^*(\mathbf{s}_0) = -1$ ,  $\Psi_1^*(\mathbf{b}_1) = 0.5$ ,  
 $\Psi_1^*(\mathbf{b}_0) = -0.5$

Rank	Complex	$\Psi_1(\text{Complex})$
1	$\mathbf{s}_1$ & $\mathbf{b}_1$	<b>1.5</b> = 1+0.5
2	$\mathbf{s}_1$ & $\mathbf{b}_0$	<b>0.5</b> = 1-0.5
3	$\mathbf{s}_0$ & $\mathbf{b}_1$	<b>-0.5</b> = -1+0.5
4	$\mathbf{s}_0$ & $\mathbf{b}_0$	<b>-1.5</b> = -1-0.5

We can now strengthen the main result of this paper.

***The principle of recombinant values holds just in case there is an absolute additive realization of value.***

Proof: Clearly the right-hand side entails the left. Assume recombinant values. Then there is an additive realization  $\Psi$  of  $\geq$  and so  $\Psi^*$  is an additive realization of  $\geq$ .  $\Psi^*$  Recall the conditions for being an absolute realization.

(i) Features are assigned positive and negative values in a systematic and non-arbitrary way.

This is clearly satisfied.

(ii) Necessity favors neither good nor evil.

To see this, note that:

$$\begin{aligned} f^{\Psi^*} &= \sum_i \Psi^*(f_i) \mathbf{P}(f_i) = \sum_i (\Psi(f_i) - f^\Psi) \mathbf{P}(f_i) \\ &= \sum_i (\Psi(f_i) \mathbf{P}(f_i) - f^\Psi \mathbf{P}(f_i)) \\ &= f^{\Psi^*} - f^\Psi \\ &= 0 \end{aligned}$$

Hence each factor has zero expected  $\Psi^*$ -value, as desired.

(iii) Reconciliation:

$$\begin{aligned} \sum_V \Psi^*(V) \mathbf{P}(V|f_i) &= \sum_V [\Psi^*(V(f)) + \sum_{g \neq f} \Psi^*(V(g))] \mathbf{P}(V|f_i) \\ &= \sum_V \Psi^*(V(f)) \mathbf{P}(V|f_i) + \sum_V \sum_{g \neq f} \Psi^*(V(g)) \mathbf{P}(V|f_i) \\ &= \Psi^*(f_i) + \sum_{g \neq f} \sum_V \Psi^*(V(g)) \mathbf{P}(V|f_i) \end{aligned}$$

Now:

$$\begin{aligned} &\sum_V \Psi^*(V(g)) \mathbf{P}(V|f_i) \\ &= \sum_V \sum_j \Psi^*(g_j) \mathbf{P}(V(g) = g_j|f_i) \\ &= \sum_j \sum_V \Psi^*(g_j) \mathbf{P}(V(g) = g_j|f_i) \\ &= \sum_j \Psi^*(g_j) \sum_V \mathbf{P}(V(g) = g_j|f_i) \\ &= \sum_j \Psi^*(g_j) \mathbf{P}(g_j|f_i) \\ &= \sum_j \Psi^*(g_j) \mathbf{P}(g_j) \quad (\text{by Probabilistic Independence}) \\ &= g^{\Psi^*} \\ &= 0. \end{aligned}$$

$$\begin{aligned} \text{So } \sum_V \Psi^*(V) \mathbf{P}(V|f_i) &= \Psi^*(f_i) + \sum_{g \neq f} g^{\Psi^*} \\ &= \Psi^*(f_i) + 0 \\ &= \Psi^*(f_i) \end{aligned}$$

thereby establishing Reconciliation.

The resulting measure depends on both the realization  $\Psi$  and the logical probability function  $\mathbf{P}$ . These functions can vary, of course, and such variations carry with them some variations in the overall orderings of non-basic states. To derive the set of orderings mandated by the initial ordering of complexes alone, we need to filter out these variations. It is thus natural to say that for any states A and B, state A is at least as good as state B, if A's intrinsic value is at least as great as that of B under all realizations  $\Psi$  and all admissible probability functions. We have already mentioned factor neutrality as a condition of admissibility. But there might be others. One which seems fairly attractive pertains to the *width* of determinates of a

determinable. We can think of logical probability as measuring the logical width of a feature, and it seems as though the factors should be divided up into features which all have the same logical width as each other:

*Equal logical width*

For all  $f_i$  and  $f_j$  :  $P(f_i) = P(f_j)$ .

Equal logical width and factor neutrality jointly yield a unique measure – the one which (naturally enough) assigns equal weight to all the complexes. Given this we can easily show that the orderings of non-basic states in Example 1 are as one would intuitively expect. For example:

*beautiful or smart > beautiful or stupid > ugly or stupid.*

Thus given the principle of recombinant values we can derive a complete account of the intrinsic value of both wholes and parts. Whether or not we are given recombinant values depends, of course, on the actual structure of the good.<sup>26</sup>

NOTES

<sup>1</sup> Pearson (1942), p. 310.

<sup>2</sup> Surprising, perhaps, because the principle has been lurking in the mathematics of measurement theory for over thirty years, most elegantly and explicitly in Scott (1964).

<sup>3</sup> Kant (1959), pp. 9–10.

<sup>4</sup> Aristotle (1962), Book X, Chapter 2, 1172b, 27–33.

<sup>5</sup> The term “bare difference” was coined by James Rachels. See his (1986), pp. 111–4. For another explicit application of the method, see Oddie (1993).

<sup>6</sup> There are notable exceptions of course. Harman argued for an additive account of intrinsic value in his (1967). Quinn criticized and modified Harman’s admittedly simple account in his (1974). Rival accounts along somewhat similar lines have been advocated by Oldfield (1977), Carlson (1997) and Danielsson (1997). The positive account I offer here delivers in a very straightforward and non-ad hoc manner the fundamental intuitions which motivate those accounts.

<sup>7</sup> Kagan (1988).

<sup>8</sup> Moore (1960), p. 28. W.D. Ross rejected Moore’s examples of organic unity, but did endorse the Kantian example as a “genuine illustration of the doctrine”. See his (1963), pp. 70–72, and (1939), pp. 185–6. Others have detected in Ross’s own treatment of duty what is effectively a thesis of inseparability. See Slote (1992),

p. 33. See, also, Lemos (1994), especially chapter 3. Although employing rather more obscure argumentation F.H. Bradley also appears to be arguing against additivity in his (1893), p. 405.

<sup>9</sup> See Oddie (2001).

<sup>10</sup> Often, but not always. Some spaces are intrinsically inseparable.

<sup>11</sup> Two further heuristic principles constrain the selection of a factorization. First, there is a way of measuring the degree of logical dependence of factors, and if logical dependence is maximal then additivity is totally trivial and uninteresting. So we should be guided by the heuristic of maximizing logical independence of factors. Further, some factorizations may be too distant from natural categorizations. If possible we should select as the basic axiological factors those that satisfy additivity, maximize logical independence, and are as “natural” as possible (Oddie (2001)).

<sup>12</sup> See Krantz (1971), p. 256, and also p. 423 ff.

<sup>13</sup> I am using *determinate* here in the relative sense. Thus “red may be determinate with respect to color, but determinable with respect to specific shades of red” Swoyer (2000), section entitled “Determinables”.

<sup>14</sup> Thus I also take the notion of a *whole* to be a relative one: relative to a specification of factors. See the previous footnote. Throughout I will simplify by not distinguishing complex features from complex states. A state of affairs can be thought of as an assignment of a feature from each factor to each one of a range of individuals. The various principles can easily be extended to states.

<sup>15</sup> I assume here, for the sake of simplicity, that the factors are logically independent – features can vary within one factor without logical consequences for features in other factors. As we will see this simplification is innocuous. The main theorem of the paper generalizes smoothly to the case of spaces of complexes with logically dependent factors.

<sup>16</sup> We can define being *the same value*,  $\approx$ , in the obvious way:  $V \approx U$  just in case each is at least as good as the other (both  $V \geq U$  and  $U \geq V$ ). Finally,  $V$  is *better than*  $U$  just in case  $V$  is as good as  $U$  but  $U$  is not as good as  $V$ :  $V > U =_{df} V \geq U$  and not  $V \approx U$ . Of course, without further constraints on  $\geq$  we cannot yet establish what most take to be desirable features of  $\approx$  and  $>$ , like transitivity.

<sup>17</sup> See Kraft (1971) pp. 248ff. Kagan (1988) calls this *ubiquity*. p. 12: “If a variation in a factor makes a difference *anywhere*, it makes a difference *everywhere*.”

<sup>18</sup> The objection was suggested by an anonymous referee, and the rest of this paragraph is a direct quotation with only inconsequential stylistic modifications.

<sup>19</sup> If you feel, with Kant, that these judgements may be true only conditional upon a good will, but false conditional upon a bad will, then it will suffice, for my purpose, that we assume a good will, and hold that fixed.

<sup>20</sup> This has been known, of course, for a long time: Krantz (1971), p. 259 “Historical Note”. Kagan (1988) concedes as much, but his argument for the lack of entailment is obscure (see his footnote 7, p. 16).

<sup>21</sup> The question posed and answered here is closely related to a question first posed by de Finetti in 1937: what does it take for a qualitative probability relation to be realized by a finitely additive probability measure. The analogue of separ-

ability in this domain is the following: if  $A \cap C = \emptyset = B \cap C$  then  $A \geq B$  if and only if  $AUC \geq BUC$ . de Finetti's other conditions were: transitivity, connectedness, non-triviality ( $S > \Delta$  where  $S$  is the whole space), and non-negativity (for all  $A$ ,  $A \geq \Delta$ ). In 1959 a counterexample to de Finetti's conjecture was constructed by Kraft et al. using a thirty two element Boolean algebra generated by five atoms. See Kraft (1959).

<sup>22</sup> This condition is an adaptation and generalization of the answer to de Finetti's query concerning the additive realization of qualitative probability measures. I arrived at it through Dana Scott's application of a theorem in linear algebra to a range of problems in the theory of measurement. See Scott (1964). For closely related methods and results, and a brief history, see Krantz (1971) pp. 423–427.

<sup>23</sup> Scott (1964), Theorem 1.3, p. 236.

<sup>24</sup> See Oddie and Milne (1991).

<sup>25</sup> On the standard account of conditional probability  $P(A|B) = P(A \& B)/P(B)$ , so if  $P(B)=0$ ,  $P(A|B)$  is undefined.  $P(W|W)$  and  $\text{Value}(W)$  would thus also be undefined if  $P(W)=0$ . Both results seem wrong. Intuitively  $P(W|W)$  should be 1 come what may. This argues for taking conditional probability as the basic notion, with the standard definition being a constraint in the case where  $P(B)$  is non-zero.

<sup>26</sup> I would like to thank Erik Carlson for numerous helpful suggestions.

## REFERENCES

- Aristotle (1962): *Nicomachean Ethics*, translated by Martin Ostwald, Indianapolis: Bobbs-Merrill Co.
- Bradley, F.H. (1893): *Appearance and Reality*, New York: Macmillan.
- Broome, J. (1990): *Weighing Goods*, Oxford: Blackwell.
- Carlson, E. (1997): 'The intrinsic value of non-basic states of affairs', *Philosophical Studies* 85, 95–107.
- Danielsson, S. (1997): 'Harman's equation and the additivity of intrinsic value' in Lars Lindahl et al. (eds.), *For Good Measure*, Uppsala.
- Harman, G. (1967): 'Towards a theory of intrinsic value', *Journal of Philosophy* 64, 792–804.
- Kagan, S. (1988): 'The additive fallacy', *Ethics* 99, 5–31.
- Kant, I. (1959): *Foundations of the Metaphysics of Morals* trans. Lewis White Beck, Indianapolis: Bobbs-Merrill.
- Kraft, C.H., Pratt, J.W. and Seidenberg, A. (1959): 'Intuitive probability on finite sets', *Annals of Mathematical Statistics* 30, 408–419.
- Lemos, N.M. (1994): *Intrinsic Value: Concept and Warrant*, New York: Cambridge University Press.
- Moore, G.E. (1960): *Principia Ethica*, London: Cambridge University Press.
- Oddie, G. and Milne, P. (1991): 'Act and value: expectations and the representability of moral theories', *Theoria* 48, 165–182.
- Oddie, G. (1993): 'Addiction and the value of freedom', *Bioethics* 7(5), 373–401.

- Oddie, G. (2001): 'Axiological atomism', forthcoming in *Australasian Journal of Philosophy*.
- Oldfield, E. (1977): 'An approach to a theory of intrinsic value', *Philosophical Studies* 32, 233–249.
- Quinn, W.S. (1974): 'Theories of intrinsic value', *American Philosophical Quarterly* 11, 123–132.
- Pearson, H. (1942): *Bernard Shaw*, New York: Harper.
- Rachels, J. (1986): *The End of Life*, Oxford: Oxford University Press.
- Ross, W.D. (1939): *The Foundations of Ethics*, London: OUP.
- Ross, W.D. (1963): *The Right and the Good*, London: OUP.
- Scott, D. (1964): 'Measurement structures and linear inequalities', *Journal of Mathematical Psychology* 1, 233–247.
- Slote, M. (1992): *From Morality to Virtue*, New York: Oxford University Press.
- Swoyer, C. (2000): 'Properties', *Stanford Encyclopedia of Philosophy*, (<http://plato.stanford.edu/>) <http://plato.stanford.edu/entries/properties/>.

*University of Colorado*  
*Boulder CO*  
*USA*  
(*E-mail: graham.oddie@colorado.edu*)