



This is an author produced version of a paper published in *Philosophical Studies*. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the published paper:

Brännmark, Johan. (2017). Good-making and organic unity. *Philosophical Studies*, vol. 174, issue 6, p. null

URL: <https://doi.org/10.1007/s11098-016-0769-1>

Publisher: Springer

This document has been downloaded from MUEP (<http://muep.mah.se>).

Good-Making and Organic Unity

Johan Brännmark
[johan.branmark@mah.se]

ABSTRACT. Since G. E. Moore introduced his concept of organic unity there has been some discussion of how one should best understand this notion and whether there actually are any organic unities in the Moorean sense. Such discussions do however often put general questions about part-whole relations to the side and tend to focus on interpreting our intuitive responses to possible cases of organic unity. In this paper the focus lies on the part-whole relation in valuable wholes and it is suggested that we should distinguish between two kinds of wholes, collections and complex unities, where the latter can involve values that do not pass on their value to the greater whole in which they are included. Given this distinction we are then able to distinguish between two kinds of organic unity phenomena, the first involving a form of goodness that emerges on the level of the whole, the second involving a form of goodness that is embedded on the level of parts. In order to properly understand the latter form of goodness, there is also a need to distinguish final value from inherent value.

KEY WORDS: Goodness; G. E. Moore; Organic unity; Mereology; Inherent value

G. E. Moore's principle of organic unity, *i.e.*, the idea that there are unities in which the value of such a whole "bears no regular proportion to the sum of the values of its parts" (1903, p. 27), mainly makes a negative point. If correct it has implications both for the mathematics and the metaphysics of values, but Moore himself does not really elaborate on either. For Moore the principle can be said to be the consequence of two views. The first is invariabilism about final value (that the final value of any part in a valuable whole is the same as the value that it has by itself or in any other whole, *i.e.*, all final value is intrinsic value¹). The second is his observation that there are wholes where the values involved cannot all be understood in terms of the invariable values of these parts. In recent years, many writers have rejected invariabilism² and some³ have therefore also thought that there is no need to accept the principle of organic unities in order to account for the kind of examples to which Moore and others tend to appeal.

¹ Moore himself only operates with the notion of intrinsic value, but many contemporary value theorists have insisted on separating the notion of something having its value based only in intrinsic properties from the notion of something having its value as an end, *i.e.*, as something for the ultimate sake of which we act. For an example, see Christine Korsgaard (1983).

² Thomas Hurka (1998), Jonathan Dancy (2003; 2007), David McNaughton and Piers Rawling (2008).

³ Hurka is open to the principle being true, while Dancy as well as McNaughton & Rawling reject it.

Ultimately, whether we accept something like Moore's principle should not just turn on our intuitions about certain key examples, but on our wider understanding of values, what they are, and how they relate to each other. Questions concerning organic unities are on a philosophical level primarily questions within the metaphysics of value, albeit ones that are strongly related to questions within the mathematics of values. In what follows here it will be argued that there can be differences in how the values of parts and the values of wholes are related to each other that depend on what kind of fusions that are involved in making certain parts form certain wholes. A general framework for considering possible organic unity phenomena will be outlined, and it will be suggested that two distinct such phenomena, emergence and embedding, can be identified. Before getting to this, however, there is a need for some stage-setting.

1. Parts, Wholes, and Bearers of Value

The question of how wholes are composed from parts is not an easy one. And while the value theoretician might have a need for some metaphysics, there is a need to keep things somewhat simpler than they actually are. The overarching question of how parts come to constitute wholes is however clearly relevant to how the goodness of those parts come to compose a totality goodness of the object in question. The most straightforward way of understanding wholes, at least in a philosophical context, is mereologically, where parthood obeys at least the following three principles: (i) *Reflexivity* (everything is part of itself), (ii) *Transitivity* (if x is a part of y and z a part of x , z is a part of y), and (iii) *Antisymmetry* (if x is part of y and y part of x , x is identical to y). This means that wholes are also parts and that we have proper parthood when x is part of y , but x is not identical with y . These principles are then usually supplemented with additional principles, so that, we also get *Extensionality* (if every proper part of x is part of y , and every proper part of y is part of x , then $x = y$).

In his discussion of organic unities, Moore himself did not distinguish between mereological and non-mereological wholes. He can certainly be interpreted along mereological lines and when he talks about "parts" he should then be understood to be talking about "proper parts",⁴ but for Moore the most important notion of parthood would presumably be *whatever can bear intrinsic value*. In the Hegelian tradition, from which Moore took the notion of organic unity, the relevant parts and wholes should perhaps not be understood mereologically, but this is not the place to delve deeper into Hegelianism.⁵ It is however important to keep in mind that depending on how you conceive of a whole, different parts will come out as proper parts of that whole. A classic example in this context (Rescher 1955) is about how an individual can be part of a squad that is part of a platoon that is part of a company.

⁴ This interpretation is, for instance, suggested by Erik Carlson (2001, p. 336).

⁵ Although see Schaffer (2010) for a reading of Hegelian-style monism which makes it less radical than that kind of position is often understood to be.

But while the parts of the squad are individuals, the parts of the company are military units, so the individual is not a proper part of the company even if he is a proper part of the squad and the squad is a proper part of the company. A similar example is about how the house has a door, and the door has a handle, but the house does not have a handle (Lyons 1977).

In a taxonomy of different senses of *part of* in English, Winston, Chaffin & Herrmann (1987) identify six meronymic relations, among them component-integral object (*e.g.*, handle-door), member-collection (*e.g.*, tree-forest), and portion-mass (*e.g.*, slice-pie).⁶ Apart from these senses of parthood, there are also different forms of inclusion, such as topological relations (*e.g.*, the wine is in the cooler) and member-class relations (*e.g.*, roses are flowers), which might colloquially sometimes be expressed in terms of parthood, but which are distinct from it. While it is in principle possible that some meronymic relations just are intransitive, the suggestion they make is that intransitivity occurs when two or more senses of parthood are involved. Take the example with the handle: it is a component of a door, part of what makes the door function as a door; the door is however not a component of the house (since the house is not an integral or functional object in that kind of way, even though it obviously can have its uses). The door is rather a piece of the house, which is arguably about the portion-mass relation instead. The handle is certainly still included in the house in a topological sense, and surely there is some loose sense of being a part in which it is a part of the house, but is not in any robust sense a part of the house, either in its own right or because it is a part of something that is a part of the house. On this kind of picture, transitivity is only guaranteed as long as we are dealing with one specific kind of parthood. A key issue for whether we have transitivity of parthood when it comes to valuable wholes is then about whether a full understanding of how the relevant parts involved in forming valuable wholes will involve only one or (at least) two kinds of parthood and about whether the greater wholes that are ultimately of interest in deciding what to do will always only involve lesser wholes of the same kind (at least with respect to understanding them as valuable wholes) or if they will at times involve lesser wholes that are wholes of a different kind, the parts of which will then not be proper parts of the greater whole, even though they might still be contained by the whole or be part of the whole in some loose sense.

In what follows here, a key distinction will be between collections and complex unities,⁷ where the former are created simply by demarcation, and the latter essentially have a deeper level of integration. Mereological composition allows for

⁶ Winston, Chaffin & Herrmann are approaching this issue from a linguistic perspective (and the three other meronymic relations they identify in English are stuff-object, feature-activity, and place-area). From a philosophical perspective it is an open question how many such meronymic relations that we need in constructing an ontology for some specific domain and whether we might even need to stipulate additional more abstract sense of parthood.

⁷ These terms are borrowed from Ivar Segelberg (1999[1945], p. 49).

really cheap wholes. In the context of ethics, this is perfectly fine, by and large, since we are often dealing with scattered wholes, such as *all the consequences of action x*, and presumably these are always collections – wholes defined simply by a certain causal demarcation. The interesting question is whether, in such collections, there might also be some complex unities, where the defining feature of such complex unities will be that they are more strongly integrated, that the parts define or shape each other in ways that go beyond what is required in order to be a mereological sum. It is quite possible that such complex unities should not be understood as mereological sums at all, but as having a non-mereological mode of composition instead, yet nothing here will hinge on that being the case. It is enough that there is a difference in senses of parthood, similar to the example with the handle, so that there is a break in the transitivity of parthood. This would mean that it cannot be taken for granted that all value-bearing parts scooped up into a collection will ultimately be proper parts of that collection, at least not if we are to make sense of them as bearers of value. They might instead be parts of complex unities, which are in turn proper parts of the relevant collections, but for the sake of calculating the overall value of the whole, these complex unities will behave like basic bearers of value, even though there are values borne by their sub-parts.

If we are to consider the parts that make up valuable wholes, there is first a question as to which ontological categories they belong to. What kind of things are the makers and the bearers of value? There is no consensus here in the literature, although two main camps might be discerned on the issue of value-bearers. We have the propositionalists, who maintain that propositionally structured entities like states of affairs or facts are the fundamental value-bearers, and then we have the non-propositionalists who maintain that entities like concrete objects or properties, are the fundamental value-bearers. There are several prominent value theoreticians in the propositionalist camp, W. D. Ross (1930) perhaps being the clearest early example, and among later philosophers there is what might be called the Chisholm school of value theory: Roderick Chisholm himself (1986), Fred Feldman (2000), Noah Lemos (1994), and Michael Zimmerman (2001). Among contemporary philosophers in the non-propositionalist camp a Kantian like Christine Korsgaard might belong here, as does Elizabeth Anderson (1993), and Wlodek Rabinowicz & Toni Rønnow-Rasmussen (2003) have at least a foot in this camp.

We should however not make too much of the difference between these two camps. To begin with, any complete ontology should ultimately allow us to make sense of talk framed in terms of non-fundamental ontological categories and this task of translation is hardly a burden that rests solely on the shoulders of the value theoretician. While there might be a need for some paraphrasing, talk in terms of concrete objects and talk in terms of states of affairs should accordingly at the end of the day be interchangeable, albeit at times perhaps clumsily so. Certain examples will probably be easier to explicate in one idiom, others in another. There is also good reason for the value theoretician *qua* value theoretician not to be strongly

committed to a particular ontology of bearers of value since this a stand that ultimately depends on, or is even determined by, our overall ontological framework. If the world is a world of states of affairs, then of course, the bearers of value must be states of affairs, and if the world is a world of concrete objects, then the bearers of value are presumably concrete objects too.⁸ The value theoretician is not faced with an ontological smorgasbord from which one can pick and choose the ontological categories of one's fancy, at least not without implicitly committing to an overall ontology.

While examples might sometimes have to be translated into different idioms, the crucial presupposition here is that the distinction between collections and complex unities can be maintained irrespective of which idiom we use. Some complex unities might have a non-mereological mode of composition, but the idea here is that it can be enough that there are different parthood relations involved, for certain parts of a larger whole to be considered as complex unities. Even if they will have identifiable parts themselves, these parts will not be proper parts of the larger whole, the one that we are ultimately considering when deliberating about what to do. Other parts that themselves have proper parts will be mere collections and those parts will be proper parts of the larger whole. It should be noted that even if we do not think that there are any organic unity phenomena with respect to values, we might very well still accept that there are complex unities, and even that there is non-mereological composition. Take the state *Harry's being happy*, clearly we can have both *Harry's existing* and *happiness being instantiated* without having this state. There is a fusion here that one might very well see as going beyond mere mereological fusion, which would mean that we are dealing with a complex unity.⁹ Presumably, we are dealing with some form of structuring uniting the two, where this structuring cannot be seen as yet another part of the whole, since in that case there would simply be an additional question about how that part is fused with the first two parts (and so on).¹⁰ The easiest way out is perhaps to understand states of affairs as simple, but then again, they really do seem to be complex entities. The main lesson of this, in the present context, is simply that there might be axiologically unimportant complex unities, because while they might bear values, they do not contain parts that bear values. They might contain good-makers under the hood, so to speak, but again, this does

⁸ On one standard understanding (Armstrong 1997) of propositionally structured entities like states of affairs, they ultimately involve particulars and properties, related in certain ways, so even if you have states of affairs as a kind of atomic building blocks, they would still have sub-atomic parts that are not states of affairs.

⁹ Armstrong (1978, p. 118) accepts that states of affairs have a non-mereological mode of composition with respect to their constituents.

¹⁰ This is an example of the Bradley Regress, which is one horn of a dilemma presented by F. H. Bradley (1908, pp. 19-25). The other horn of the dilemma is that we do not get fusion. It should be made clear that this problem does not automatically go away just by accepting non-mereological fusion, we just increase our options that way.

not commit us to the existence of organic unity phenomena. It does however point to why, in order to look more closely at how to understand possible organic unity phenomena, there is a need to first look at good-making more generally.

2. The Ways of Good-Making

In what follows here I will here speak of *constitutive parts* as a way of talking about proper parts (in the mereological sense) as well the constituents of non-mereological wholes. Some of these constitutive parts can be subatomic in the sense that we find them on a lower level than that of the fundamental value bearers, but they might be atomic and supra-atomic as well. It is an inclusive notion, but in a mereological whole, the whole is neither a proper part nor a constitutive part of itself.

Now, there would seem to be at least the following four basic possible modes of good-making (and, *mutatis mutandis*, bad-making) with respect to constitutive parts:

- (i) *Primitive*: where constitutive parts of the object make it good, but where these parts themselves are not bearers of goodness.
- (ii) *Aggregative*: where constitutive parts of the object, which are themselves bearers of goodness, make the object correspondingly better by passing on this goodness to the object.
- (iii) *Amplificatory*: where some constitutive parts of the object, which are themselves not bearers of goodness, makes the contribution of some other good-making part to the goodness of the object greater or smaller.
- (iv) *Formative*: where constitutive parts of the object, which are themselves bearers of goodness, by virtue of their goodness, make the object good in a way distinct from passing on their goodness to the object.

Whether we ultimately need to appeal to all four of these in order to make sense of how the value of wholes is made up is an open question, although it is difficult to see how we could make do without at least (i) and (ii). The former would typically be a matter of certain abstract constitutive parts, such as properties, making the thing in question good; *e.g.*, an instance of pleasure is constituted by certain properties and at least some of these (like intensity, duration, and the bare quality of pleasantness, if there is such a thing) make it good without themselves being bearers of goodness. This is similar to how a certain set of physical properties might be required for a mental quality to supervene on them without any mental quality supervening on any single one of them or any of them being mental. Given that it is extended in time, this instance of pleasure will however also have temporal parts, which are proper parts, that are themselves bearers of a goodness which is then aggregated in the pleasure as whole; this is something that corresponds to (ii). If you are an invariabilist who thinks that value aggregation is summative, you have no need for anything beyond (i) and (ii), and would probably be wary of going beyond them. One is then only dealing with two ways of being an axiologically relevant part. The first, involved in primitive good-making, can largely be ignored, since it will be sub-atomic

(although in principle the primitive good-makers might be embedded in complex unities). The second, aggregative good-making, is ideally suited for understanding relevant wholes as mere collections. There will be no problems with the transitivity of parthood and values will simply be passed on from the lower parts to the larger whole.¹¹

If we turn to (iii), it opens up for a part making a different contribution to the whole than the value it carries by itself and one notion that can be used to make sense of this is *contributive goodness* (Moore 1966, pp. 130-31, and Ross 1930, pp. 72-73): the net difference something makes for the value of the whole. This is a conceptual possibility that can be represented by an example of the pleasant contemplation of beauty, where the pleasure might already be good, but where the beauty of the object in which the pleasure is taken amplifies the goodness of the pleasure. However, it should be said that the fact that a certain part makes a certain marginal contribution of goodness says nothing about whether any extra value is located in this part, or where it is located at all. The notion of contributive goodness does not by itself do much explanatory work.¹² In fact, a variabilist could very well accept that there are parts that are amplificatory good-makers without seeing a need to locate this value anywhere else than in the proper parts (and hence be able to reject Moore's principle of organic unity). It is when we turn to (iv), where we have something that bears value but still acts similar to a primitive good-maker, that we might need to accept the principle of organic unity in order to make sense of the values of the relevant parts and wholes. Already (iii) is problematic on the simple summative view, but (iv) even more clearly introduces the possibility that there might be different kinds of parthood involved in some value-bearing wholes and that some value-bearing parts might not be proper parts of the larger whole.

There seems to be at least two possible phenomena that can be understood as exemplifying formative good-making as well as the principle of organic unity. To

¹¹ There is still a complication in that we might decompose a whole in many different ways and in terms of different levels. Five minutes of pleasure might be divided into five minute-long parts but also into 300 second-long parts, and while all these parts might be considered as bearing value we cannot simultaneously consider them as contributing this value to the whole since that would involve double-counting. We might address this problem by identifying some atomic states (Feldman 2000), ones that cannot be further divided without ceasing to be bearers of value, but there might be cases where it is unclear what would constitute such atoms or cases where the choice between some different ways of cutting the thing up into parts just seems arbitrary. The problem of double-counting is however primarily a problem in the mathematics of value, *i.e.*, in formulating a general method for the computation of value, we must do so in a way that precludes double-counting. When the focus lies on the metaphysics of value this problem can largely be set aside. It will not be important for any of the examples considered here exactly how to determine the quantities of values involved, rather it is a matter of understanding the structures of these wholes.

¹² Since contributive value is a form of marginal value, it is possible that there are several valuable parts of a given whole that are such that if the object loses any one of them, it will lose all of the value that there is in addition to the sum of the values of the parts by themselves; so all of these parts will have a contributive value corresponding to that additional value.

begin with, there is what one might call *emergence*, where a value emerges on the level of the whole as a whole, partly depending on the values that belong to the constitutive parts of the object in question. Then we have what might be called *embedding*, where the value borne by some constitutive part(s) will not be passed on to the whole. What this means is that there are two possible steps that one can take away from the simple picture that there are no organic unities. The first step need not take us beyond the realm of mere collections, but involves accepting that there are structural features that supervene on certain sets of atomic value-bearers and that these structural features can be good- or bad-making. The second step takes us into the realm of complex unities. Presumably, some people would prefer if they did not have to take any of these steps, but would rather just take the first one, if they have to move at all. It will be argued here that we need to go all the way. Of these two, the latter one is the more radical possibility, since it will potentially complicate the relation between values and reasons for making good things persist or come to exist. It is often assumed that there is a strong link between values and reasons, that to have value is to be something for the sake of which we ought to act in certain ways and to be a fitting object of pro-attitudes. There are many who would even say that there is a necessary connection between having value and being a provider of such reasons.¹³ On a simple summative picture of good-making, the parts bearing final values, whether negative or positive, can be seen as providing reasons for or against making the whole in question exist or persist, and when you add them up, you get a net balance of reasons. If there is such a phenomenon as embedding, however, certain parts might have value without providing reasons for making wholes that include them exist or persist.

Borrowing a notion from C. I. Lewis (1946, Chapter 14), one can say that some of the things that we tend to call intrinsically or finally good are to be considered in the first instance as being inherently good.¹⁴ Something which is inherently good is good when considered by itself.¹⁵ There is a default reason for promoting it.¹⁶ But in

¹³ According to the buck-passing or fitting-attitudes analysis of value, this is simply what being valuable is about. Objections to this idea tend to focus mainly on whether being such a provider of reasons is sufficient for being valuable; for a seminal discussion see Wlodek Rabinowicz & Toni Rønnow-Rasmussen (2004).

¹⁴ It should be said that Lewis' own use of this notion is tied to his idea that the only things that have intrinsic/final value are experiences and that the things that have inherent value are certain possible objects of experience.

¹⁵ Moore famously proposed his method of absolute isolation (1903, p. 187) for determining whether something has intrinsic value. This is a poor method for determining whether something always makes the world a better place (as far as it goes); instead, we would presumably need to consider all possible situations where we might find that something. Still, when something appears valuable when considered by itself, surely that says something important about that thing. The method of absolute isolation can accordingly be a test for whether something has inherent value. Note that for the argument here it need not be presumed that if something has inherent value, it always has inherent value, although that might ultimately be the case. In that case, one could simply use the notion of

actuality, goods are always parts of larger wholes and the reasons we have for promoting these goods will merely typically be reasons for promoting these larger wholes. If the value borne by a part is not passed on to the whole, it seems reasonable that the part is not reason-providing either, at least not in the typical way. The suggestion here, then, is that if we allow for formative good-making, we should see inherent value as the fundamental form of value. Things that have inherent value are inherently attractive. But they do not always have final value; they do not always provide reasons for promoting the whole of which they are a part. Take pleasure. When you consider pleasure as such, it is hard to find fault with it; there is something positive to be said about it without having to think of it in terms of pleasure taken in innocent matters or anything like that. Pleasure has inherent value. But does this mean that pleasure always has final value? There might be values in the world, the presence of which would typically make the world a better place, but which in their present incarnation do not make the world a better place.

Alternatively, if we want to pass the buck entirely to reasons, what we will say is that as long as we have a “common currency” in terms of reasons for and against promoting, these reasons are simply aggregated and the value borne by the greater whole is to be understood in terms of the net balance of aggregated reasons for and against promoting. But if, with respect to certain parts, the buck is not to be passed to reasons for and against promoting, but rather to other reasons instead, *e.g.*, reasons for certain emotional responses, then we no longer have that common currency and *Transitivity* no longer holds. The idea that the value of the greater whole is a sum of the values of its parts only makes sense if the parts bear value in the same manner, and if we pass the buck this would seem to require that they at the very least share a primary way of being reason-providing that can be used in order to achieve a net balance of reasons. When determining what to do, we will be looking at reasons for and against promoting, and all those value-bearing parts in the relevant whole that can be understood in terms of such reasons can be understood as forming a collection. In fact, irrespective of whether we are buck-passers or not, membership in the kind of collections that primarily will interest us in an ethical context will be about whether something provides reasons for or against promoting. But certain parts of such collections might still have to be understood as complex unities, containing values that provide other types of reasons and which play a formative role with respect to final values that can, in turn, be understood in terms of reasons for and against promoting.

intrinsic value instead, only with the caveat that intrinsic values need not always be final values, or that intrinsic value is not always passed on to the larger whole in which the relevant bearer of value is located.

¹⁶ In the literature on particularism, there is an analogue to inherent value in the notion of default or presumptive reasons; see Garrett Cullity (2002). There is a further question about whether there are any final values that are not inherent values and it will not be entered into here; suffice to say that the answer to it does not make any difference to the argument here.

The argument here will be that we do find axiologically relevant forms of complex unity, where there are constitutive parts of those complex unities that are not just primitive good-makers, but value-bearing in their own right. First, we will however look at a phenomenon that, while requiring us to accept the principle of organic unity, does not require us to model the valuable wholes in question as complex unities. This discussion will be kept relatively brief in order to leave more room for discussing what will be called *embedded values*. In both cases, some alternative interpretations will also be considered, but these discussions will be fairly brief (although hopefully still indicative of the kinds of problems that are faced by these alternative interpretations).

3. The Phenomenon of Emergence

Say that a certain part of one's life might contain prudential goods in either of two ways, $A+A+A+A$ or $A+B+C+D$, and that all of these instances of goodness would, taken by themselves, be equally good; some would then still say that the second option is better. There are two basic interpretations of this. One is that there is a diminishing marginal value of single types of final goods (which means that in the context of $A+A+A+A$ an instance of A does not have the goodness that it would have if you considered merely a single instance of A). The other is that there is some holistic positive value in variety or, alternatively, negative value in monotony. If we put the first interpretation (which would amount to a variabilist account of final value) to the side, then it would seem to be the case that if variety among the parts makes the whole better, then the point is really that we cannot have just $A+B+C+D$, we will always have $A+B+C+D+E$, where E is a value that supervenes on a structural feature of a whole that has among its proper parts the ones on which values A, B, C, D supervenes.

Are wholes like these mere collections or complex unities? Variety is clearly an emergent feature of the relevant whole, but it is at the same time simply a feature of the whole, it does not alter the way in which the parts of the whole are parts. That a whole is a mereological sum does not preclude the possibility of it having properties not had by any of its proper parts. The French army might have the property of being stronger than the Prussian army without its units having this property, but it could still be reasonable to think that the French army is nothing but a mereological sum of its units. Even when there is emergence, something can accordingly still be just a collection, since the unity involved here is simply the unity involved in proper parts being grouped together, which need not be understood as a matter of complex unity. Emergence is certainly a phenomenon that might put into question whether a whole is merely a mereological sum, but at least when it comes to properties the emergence of which are simply due to the mere grouping together of certain parts, emergence does not seem to rule out a strictly mereological mode of composition. The difference between a collection and a complex unity is that *qua* object the former is

created by the mere demarcation of it. With this demarcation in place, however, certain features might emerge on the level of the whole that cannot be found on the level of proper parts. If the variety in the example above is just a product of certain distinct goods happening to fall within a certain segment of a life, then that portion is nothing more than a collection. In a complex unity there is something more than mere demarcation which ties the object together. Of course, human lives certainly seem to be candidates for being understood as complex unities, and there are writers who take positions seemingly involving seeing them as such when making sense of how good lives are made good (*e.g.*, Brännmark 2003, Levinson 2004). The argument here is just that with respect to emergence, there are no good parts that stand in relations to each other that necessitate understanding them as complex unities. The value-bearing parts of the wholes $A+A+A+A$ and $A+B+C+D$ can function both as aggregative and formative good-makers, but on the above interpretation they are still fully final values. They are formative good-makers in a one-sided way, and what happens is simply that some additional value, providing reasons for or against promoting, supervenes on them. There is accordingly a common currency of value throughout, and this is why these wholes can ultimately be understood simply as collections rather than involving any complex unities that contain embedded values.

There are also other examples of possibly emergent values. One is inequality. For some people this is a disvalue that resides on the level of wholes, in this case collections of people and their levels of welfare.¹⁷ This is a kind of example which is often not described as a case of organic unity, but if we regard the welfare levels of the individuals involved as the atomic value-bearers, it certainly seems to exemplify Moore's principle of organic unity. The contributive goodness of the individual instances of welfare will often not equal the final value borne by these instances; adding an individual with a very high level of welfare to a whole will, for instance, not just make it better by the amount of final value belonging to that welfare, the net contribution will be less than that, because there will also be an increase of inequality. There is a difference between inequality and variety in that while the latter could possibly be interpreted in terms of the emergent value supervening on a structure that is created by the arrangement of the atomic value-bearers rather than their values, the disvalue of inequality seem to be more strongly linked to patterns of distribution of value, *i.e.*, it is even more clearly an example of formative good-making.

Another example, which is reminiscent of inequality, since it also about distributive patterns, has to do with how goods are distributed across a single life (*e.g.*, Nozick 1989, p. 100). Even if we think that pleasure is the only basic building

¹⁷ It should be said that many philosophers reject such an interpretation, since they find that the badness of equality must be a badness that concerns the individuals involved; they believe in what Larry Temkin has called the Slogan (1993, p. 256): "one situation *cannot* be worse (or better) than another *in any respect* if there is *no one* for whom it *is* worse (or better) *in any respect*."

block of a good life, we might still find that a life where this pleasure is predominantly located in the later parts of that life is superior to one where it is predominantly located in the earlier parts of that life. A progressive structure is better than a regressive. Again, the contributive value of goods might then vary depending on where they are located temporally. This type of example is however more difficult to interpret. In the case of inequality the distribution of goods is clearly directly relevant, but in this case our intuitive responses to the relevant distributive structures might very well have to do with implicitly assuming that the lives in question are characterized by something more than mere distributive patterns, that they also have some narrative features characterizing the lives in question, that the life with more pleasure at the end is a life building towards success while the other life is one of ultimate failure, but then the difference between the two would be deeper than just a matter of pure distribution. We will not delve deeper into this particular example here.

4. The Phenomenon of Embedding

If the interpretation of emergent values provided here is correct, we do not need to appeal to complex unities in order to account for such values. There is however another important class of examples in the literature, ones involving what might be called *embedding*. Two standard such examples are deserved punishment and malicious pleasures. We will go through these in turn.

Deserved punishment is Moore's primary example (1903, p. 216). A simple collection of a crime and a punishment (or perhaps better: the act that constitutes the punishment) would just be two bad things, but an emergent retributive value can potentially make sense of why adding punishment is actually an improvement of the situation. Some might perhaps dispute the evaluative views involved, thinking that retribution has no positive value, but here the question is rather one of how to make sense of the values involved in Moore's example *if* there were to be such values. There is however a problem with Moore's interpretation, namely that if whatever constitutes the punishment has negative value, then if there is a connection between values and reasons, this would mean that there is a reason against inflicting the punishment. While there might certainly be reasons for inflicting the punishment in a somber mood, regretting its necessity, this is not the same thing as there being a positive reason against inflicting it. This can be contrasted with how yet another instance of a particular good might decrease variety in life – here it seems reasonable to say that there are reasons pulling in two directions. Punishment does not seem to be like that.

Part of the explanation for this is that while variety or monotony is an extra feature “on top” of a line of other goods, the case of crime and punishment involves two constitutive parts permeating each other, transforming each other: the crime is no longer just a crime, it is a punished crime, and if the punishment is completed it is

also a crime atoned for. And if we look at the punishment, then as F. H. Bradley points out, “[p]unishment is punishment, only when it is deserved” (1927, p. 26). The suggestion here, then, is that arguably there is a stronger form of fusion, compared to mere demarcation, involved in binding together the constitutive parts of the whole that we create by meting out a punitive measure to a deserving person. Of course, there are other consequences as well that might bear on what we do, such as the importance of deterring future possible crimes, but if we look at the final values involved in a whole like *all the value-bearing consequences of punitive action x*, then they will include a complex unity where the badness of the punitive action is embedded. The complex unity of a crime having been punished bears a certain final value, which is still negative, but which is better than the negative value that would have been borne by the same crime having gone unpunished.¹⁸ In the case of emergence, we had an external relation of formative good-making, but here we have an internal relatedness. The disvalue of the crime plays a part in making the punitive act into an act of punishment, and the punishment plays a role in transforming the criminal act into an atoned crime.

In the whole that comes into being, the exact natures of both of its main parts are dependent on the values of one another. This means that there is a meshing here that we do not find in the case of emergence. It also means that there is a deeper level of integration and that the crime and the punishment are parts of the whole which they compose in a different way than when values are just scooped up in a collection. The parthood relations involved are more like functional parthood relations. This means that when we consider the larger whole *all the value-bearing consequences of punitive action x*, there are differences in how value-bearing parts are parts of that larger whole, and *Transitivity* accordingly no longer holds for all value-bearing parts. The complex unity composed by the crime and its punishment has value-bearing parts, but in contrast to the case of emergence, these do not function as aggregative good-makers in relation to the larger whole, only as formative good-makers in relation to the complex unity that they form. That complex unity does however have a final value, which can be aggregated to form part of the value of a larger whole. This kind of picture seems to get the reasons right: the inherent values involved can ground responses such as regret or somberness, but it is not the case that there is actually a positive reason against inflicting punishment.

¹⁸ When the crime goes unpunished it seems reasonable to say that we do not have a complex unity, we just have the badness of the crime, but the crime is on the other hand never a direct part of the consequences of our actions – it was a consequence of the actions of the criminal in question. The crime can however still play the role of a formative good-maker in relation to the complex unity that is created when we mete out punishment. Without the crime, the act that constitutes punishment would in fact be even worse than that act carried out by itself (because it would involve a miscarriage of justice). By understanding the badness of the crime as embedded in a complex unity we can have it playing a good-making role, while still being completely clear about its inherent badness.

A variabilist account could get the sums right, in the sense that our all-things-considered reason for inflicting punishment would have the strength that it should have, but there would be problems in getting the full picture right. Either we would have to say that the act that constitutes the punishment is good, and that therefore there is a reason for inflicting it, or that the infliction of punishment is a bad thing, but that it changes the value of the crime so that the contributive value of the act of punishment is still positive. Neither alternative is satisfactory. Presumably, the punishment would not really be punishment if it was not a bad thing (*cf.* Lemos 1994, pp. 43-4), so the first option is not acceptable. And even if one might, perhaps even should, think that there is some positive value in a crime atoned, this would presumably still only be a rather small silver lining on a rather large dark cloud – not something so good that it has more positive value than the badness of the act of punishment. In contrast, treating the relevant wholes as involving complex unity, and hence certain values as inherent rather than final, allows the values involved to simply not be passed on to the larger whole of which they are a part.

A second example in this context is that of malicious pleasures. When considered by itself, pleasure certainly does seem good, just as pain seems bad, but if you put them together in a certain way, things are no longer that straightforward. This particular example has a complication that needs to be considered eventually, but let us start with a simple case where someone is taking pleasure in someone else's pain. The pain is there and the pleasure is caused by the pain in a way that involves sadistic inclinations. Let us say that the pain and the pleasure are quantitatively equivalent (whatever that means); evaluatively, they would then presumably cancel each other out. But this does not seem right. Adding the pleasure seems to make the situation worse, not better. The problem for a Moorean approach here is similar to the case of punishment. If we are to make sense of the total value of the situation by adding a negative emergent value, that would mean that the pleasure involved would still make a difference for the better, *i.e.*, there would be something to be said for bringing the malicious pleasure about. This approach would not really capture how the pleasure in question is spoiled or perverted by the malice. By stipulating an emergent value we can get the net balance right, but we would not get a satisfactory analysis of how that net balance comes about. Cannot the emergentist say that the emergent bad brought about by something like a malicious pleasure does not just give rise to a reason that outweighs the reason based in the pleasure *qua* pleasure, but defeats it in a way that leaves nothing to be said for bringing this particular pleasure about? The argument of this paper certainly implies that the connection between reasons and values is not such that the goodness or badness of something will always give rise to a reason for and against realizing that thing, but the point is that there should be a principled difference between cases where this is the case and where it is not. The mere fact that an emergent value is involved does not seem to provide such an explanation, since basically it is just yet another value, whereas an explication in

terms of different kinds of parthood being involved can provide us with such an explanation.

A variabilist interpretation is considerably more plausible; the pleasure would then, under these circumstances, have no value or even negative value. At the same time, one might however wonder whether this gets the entire picture right. After all, is not the value of pleasure part of what makes the malicious pleasure so objectionable? Take a parallel case: sympathetic pain. If I see someone in pain and is then pained by this, this is not just adding one bad thing to another. One might even say that the situation as a whole is somewhat improved by the sympathetic pain. But at the same time it does not seem right to say that in this case my pain has no negative value or even that it has positive value. An alternative explication would be in terms of complex unities. The value of the pleasure would then act as a formative bad-maker and the positive value it still has is not passed on to the larger whole into which it is embedded. The same thing goes for sympathetic pain; pain is inherently bad, but that badness might at times function as a formative good-maker with respect to the larger wholes of which it is a part. Ultimately, if we look at the world as a whole, it can from an evaluative point of view arguably be seen as collection having a value that is aggregated from its proper parts (as well, perhaps, as the value borne by the whole as a whole). This means that if we, so to speak, want to block certain values from being part of this larger sum, the simplest way of doing so is to deny the transitivity of parthood at some point in the hierarchy of parts; and in order to do so we need complex unities that are not just atomic value-bearers, like *Harry's being happy* (where *happiness being instantiated* is a primitive good-maker making *Harry's being happy* bear value), but where certain inherent values are embedded in those complex unities, not having their value passed on upwards, but still playing a role as formative good-makers with respect to the values of the complex unities of which they are a part.

As already mentioned, there is a complication with respect to this particular example. Someone can feel malicious pleasure (or sympathetic pain) without there actually being any pleasure (or pain). He or she might just have misread the situation. What should we say about this type of situation? Pleasure is still inherently good and pain inherently bad.¹⁹ In these cases, however, we have more complex states of consciousness, involving not just the pleasure or pain felt, but beliefs about what happens in the world as well. If we regard these complex states as complex unities rather than mere collections, the relevant instances of pleasure and pain can still be understood as embedded, in the form of intentional objects, and acting as formative

¹⁹ As already pointed out, the use of the notion *inherent value* in this paper has to do with leaving it open that there might not be any values that keep their value in all contexts. The reasoning here could however allow adherents of there being intrinsic values to handle some counterexamples that tend to be made to certain plausible candidates, like pleasure. If intrinsic value can be embedded, it need not provide reasons to promote whatever it is embedded in. For instance, malicious pleasures need not be a counterexample to pleasure having intrinsic value.

rather than aggregative good/bad-makers with respect to these larger complexes. Is it reasonable to regard them as complex unities rather than mere collections of mental states? From an evaluative perspective, yes, since at least on a reasonable interpretation of the pleasures and pains in question, these are not free-standing pleasures and pain, they have as their objects pains or sufferings that are inherently bad, and where this inherent disvalue is what grounds the values of the attitudes in question. But while this inherent disvalue is involved in these attitudes, as an intentional object, it is not actually instantiated in the world so the way in which it is a part of these attitudes is quite distinct from the way in which a number of mental states make up a collection of mental states. Still, the badness of a malicious appreciation of an episode of pain, whether a real pain or not, surely depends on the badness of pain, and this badness is not part of the whole in any sense created by mere demarcation.

Note that the point here is not that pleasure is always pleasure in something or that it is always pleasure caused by something so that pleasures simply considered as pleasures are never bearers of value, and that to describe a pleasure simply as a pleasure would therefore be too unspecific, or evaluatively incomplete.²⁰ Instead, the object of one's pleasure is often just irrelevant to that instance of pleasure having value, and then the intentional object is not a constitutive part of that object *qua* value-bearer. Here is an analogue: drinking poisoned water is not good for you, but that does not mean that what is good for you is drinking non-poisoned water: it is water that is good for you. In a situation where pleasure is taken in some innocent matter, the value-bearing object is just that instance of pleasure *qua* pleasure, not pleasure-taken-in-an-innocent-matter. However, while pleasures are not evaluatively incomplete, they are evaluatively porous. What this means is that they are open to meshing, similar to what we saw in the case of crime and punishment. We then have a deep integration between the component parts, where inherent values are involved as formative good-makers, but where we do not have any final value until we reach the level of the complex unity as a whole. That complex unity, and its final value, can in turn be part of a collection, but that is another matter.

There might, of course, be cases where it is not clear whether we should understand a certain whole as a mere collection or as a complex unity; indeed, there might often be several ways of describing a given complex and no clear choice between them. Many of the key examples here are, as anyone familiar with the literature will know, possible to interpret in different ways.²¹ This means that the

²⁰ Though not framed in these terms exactly, Michael Zimmerman's rejection of organic unities (2001, Chapter 5) builds on the idea that when the parts of valuable wholes are completely specified, there is no need to appeal to organic unities in order to explicate the values of any wholes (for instance, we should never just analyze pleasures merely as pleasures but always in terms of pleasures in certain things, and so on).

²¹ Unfortunately, it is difficult to avoid having to rely on examples and appeals to intuitions here. We are dealing with different ways of modelling value-bearing parts and wholes and their relations, so the

distinction between collections and complex unities can hardly be mechanically applied to sort cases of embedding or meshing from non-cases. At the end of the day, then, not much more can be concluded here beyond noting that this distinction provides us with a way of approaching, in a principled manner, the issue of why some good parts do not provide us with reasons, as would be expected were they simply finally valuable, and why some things seem to have a certain value or disvalue, without just passing it on to the wholes of which they are parts. On the analysis provided here, the underlying reason for this is that there is a stronger fusion of the relevant constitutive parts, and through this fusion some values borne by constitutive parts are embedded or encapsulated. They then act as formative good/bad-makers, rather than aggregative ones.

5. Concluding Remarks

The argument here has been that in order to satisfactorily account for all organic unities, we need two additions to our framework compared to a Moorean one. We need a notion of inherent value and we need a distinction between collections and complex unities. These two ideas are not unconnected. Rather, the idea here has been that the stronger fusion involved in complex unities can allow us to make sense of how, for example, something of negative value can exist in the world, without making the world as a whole a worse place. The reason is that, apart from the weak form of fusion involved in composing collections, there are also stronger forms of fusion and that, because of this, a value-bearing part need not be a proper part of the greater whole that we can affect through action. And because of this, *Transitivity* does not always hold and there are value-bearing parts that no longer function as aggregative good-makers, but formative good-makers instead. Looking at certain examples, there are values that do not seem to be final values, but neither can they always be explicated as being merely instrumental values. This is why we need the notion of *inherent value*. Inherent values have a connection to reasons, but what sets them apart from final values is that they do not essentially provide reasons to bring them about or enable them to persist, which is presumably something which characterizes final values. They can however still give us cause for certain emotional responses and they can make a life go more or less well (*e.g.*, being punished presumably makes the criminal's life go less well).

The conceptual apparatus developed here is more complex than either the traditional Moorean one or more recent variabilist approaches, so there is potentially an Occamite reason against it. That Occamite reason does however usually come with a *ceteris paribus* clause, and the argument here has been that all else is not equal, but that with respect to cases of what has here been called embedding, the present

question of which such way is preferable will always turn on how we feel about different applications, *i.e.*, the extent to which they enable us to anatomically dissect the values and reasons involved in different wholes. And we only have our intuitions to judge the accuracy of those results by.

account allows us to not just get the net sums right, but also the anatomical details of all the reasons involved. In the end, the phenomena of emergence and embedding have been understood as distinct phenomena, but they both involve what has here been called formative good-making, and they are both explicated in terms of values being borne by parts that are not proper parts of the whole, in one case the whole as a whole, in the other case parts of complex unities.

Acknowledgments

This paper was originally presented at a Conference on Organic Unities organized by the Thumos group at the University of Geneva and for the discussion there I'm grateful to all the participants, although especially Noah Lemos, Kevin Mulligan, and Wlodek Rabinowicz. For helpful comments on earlier versions of the text I would also like to thank Anna-Sofia Maurin, as well as two anonymous referees. My work on it has been enabled by a position as Research Fellow granted by the Royal Swedish Academy of Letters, History and Antiquities.

Bibliography

- Anderson, Elizabeth. 1993. *Value in Ethics and Economics*. Cambridge, Mass.: Harvard University Press.
- Armstrong, David. 1997. *A World of State of Affairs*. Cambridge: Cambridge University Press.
- Armstrong, David. 1978. *Universals and Scientific Realism, Vol. I*. Cambridge: Cambridge University Press.
- Bradley, F. H. 1908. *Appearance and Reality, 2nd ed.* Oxford: Oxford University Press.
- Bradley, F.H. 1927. *Ethical Studies, 2nd Ed.* Oxford: Clarendon Press.
- Brännmark, Johan. 2003. Leading Lives, *Philosophical Papers* 32: 321-43.
- Carlson, Erik. 2001. Organic Unities, Non-Trade-Off, and the Additivity of Intrinsic Value. *The Journal of Ethics* 5: 335-360.
- Chisholm, Roderick. 1986. *Brentano and Intrinsic Value*. Cambridge: Cambridge University Press.
- Cullity, Garrett. 2002. Particularism and Presumptive Reasons. *Proceedings of the Aristotelian Society*, Suppl. Vol. 76: 169-190.
- Dancy, Jonathan. 2003. Are There Organic Unities? *Ethics* 113: 629-650.
- Dancy, Jonathan. 2007. Moore's Account of Vindictive Punishment: A Test Case for Theories of Organic Unities," in S. Nuccetelli and G. Seay (eds.), *Themes from G. E. Moore: New Essays in Epistemology and Ethics*. Oxford: Oxford University Press.
- Feldman, Fred. 2000. Basic Intrinsic Value. *Philosophical Studies* 99: 319-46.
- Hurka, Thomas. 1998. Two Kinds of Organic Unity. *Journal of Ethics* 2: 299-320.
- Korsgaard, Christine. 1983. Two Distinctions in Goodness. *The Philosophical Review* 92: 169-195.
- Lemos, Noah. 1994. *Intrinsic Value: Concept and Warrant*. Cambridge: Cambridge University Press.

- Levinson, Jerold. 2004. Intrinsic Value and the Notion of a Life, *The Journal of Aesthetics and Art Criticism* 62: 319-29.
- Lewis, C.I. 1946. *An Analysis of Knowledge and Valuation*. La Salle, Illinois: Open Court.
- Lyons, John. 1977. *Semantics I*. Cambridge: Cambridge University Press.
- Moore, G.E. 1903. *Principia Ethica*. Cambridge: Cambridge University Press.
- Moore, G. E. 1966[1912]. *Ethics*. Oxford: Oxford University Press.
- McNaughton, David and Rawling, Piers. 2008. Benefits, Holism, and the Aggregation of Value. *Social Philosophy & Policy* 26: 354-374.
- Nozick, Robert. 1989. *The Examined Life*. New York: Simon and Schuster.
- Rabinowicz, Wlodek & Rønnow-Rasmussen, Toni. 2003. Tropic of Value. *Philosophy and Phenomenological Research* 66: 389-403.
- Rabinowicz, Wlodek & Rønnow-Rasmussen, Toni. 2004. The Strike of the Demon: On Fitting Pro-Attitudes and Value. *Ethics* 114: 391-423.
- Rescher, Nicholas. 1955. Axioms for the Part Relation. *Philosophical Studies* 6: 8-11.
- Ross, W.D. 1930. *The Right and the Good*. Oxford: Clarendon Press.
- Schaffer, Jonathan. 2010. The Internal Relatedness of All Things. *Mind* 119: 341-376.
- Segelberg, Ivar. 1999[1945]. Zeno's Paradoxes" in *Three Essays in Phenomenology and Ontology*, trans. Herbert Hochberg & Susanne Ringström Hochberg. Stockholm: Thales.
- Temkin, Larry. 1993. *Inequality*. New York: Oxford University Press.
- Winston, Morton; Chaffin, Roger & Herrmann, Douglas. 1987. A Taxonomy of Part-Whole Relations. *Cognitive Science* 11: 417-444.
- Zimmerman, Michael. 2001. *The Nature of Intrinsic Value*. Littleham, Md.: Rowman & Littlefield.