

Article

"Fitting-Attitude Analyses and the Relation Between Final and Intrinsic Value"

Antoine C. Dussault

Les ateliers de l'éthique / The Ethics Forum, vol. 9, n° 2, 2014, p. 166-189.

Pour citer cet article, utiliser l'information suivante :

URI: <http://id.erudit.org/iderudit/1026685ar>

DOI: 10.7202/1026685ar

Note : les règles d'écriture des références bibliographiques peuvent varier selon les différents domaines du savoir.

Ce document est protégé par la loi sur le droit d'auteur. L'utilisation des services d'Érudit (y compris la reproduction) est assujettie à sa politique d'utilisation que vous pouvez consulter à l'URI <http://www.erudit.org/apropos/utilisation.html>

Érudit est un consortium interuniversitaire sans but lucratif composé de l'Université de Montréal, l'Université Laval et l'Université du Québec à Montréal. Il a pour mission la promotion et la valorisation de la recherche. Érudit offre des services d'édition numérique de documents scientifiques depuis 1998.

Pour communiquer avec les responsables d'Érudit : erudit@umontreal.ca

FITTING-ATTITUDE ANALYSES AND THE RELATION BETWEEN FINAL AND INTRINSIC VALUE*

ANTOINE C. DUSSAULT

DOCTORANT EN PHILOSOPHIE, UNIVERSITÉ DE MONTRÉAL

ABSTRACT:

This paper examines the debate as to whether something can have final value in virtue of its relational (i.e., non-intrinsic) properties, or, more briefly put, whether final value must be intrinsic. The paper adopts the perspective of the fitting-attitude analysis (FA analysis) of value, and argues that from this perspective, there is no ground for the requirement that things may have final value only in virtue of their intrinsic properties, but that there might be some grounds for the alternate requirement that final value be grounded only in the essential properties of their bearers. First, the paper introduces the key elements of the FA analysis, and sets aside an obvious but unimportant way in which this analysis makes all final values relational. Second, it discusses some classical counterexamples to the view that final value must be intrinsic. Third, it discusses the relation between final, contributive, and signatory value. Fourth, it examines Zimmerman's defense of the requirement that final value must be intrinsic on the grounds that final value cannot be derivative. And finally, it explores the alternative requirement that something may have final value in virtue of its essential properties.

RÉSUMÉ :

Cet article examine, selon la perspective de l'analyse de la valeur en termes d'attitudes appropriées (la AAA), le débat concernant la possibilité qu'une chose ait de la valeur finale en vertu de ses propriétés relationnelles, ou en d'autres termes, la question de savoir si toute valeur finale doit être intrinsèque. La thèse défendue par l'article est que, selon la perspective de la AAA, il n'y a aucune raison convaincante d'adopter l'exigence selon laquelle une chose ne pourrait avoir de la valeur finale qu'en vertu de ses propriétés intrinsèques, mais il semble y avoir des bases intuitives pour adopter l'exigence alternative selon laquelle la valeur finale devrait être fondée sur des propriétés essentielles de ses porteurs. L'article présente d'abord les éléments clés de la AAA et met à l'écart une manière non pertinente selon laquelle celle-ci rend toute valeur finale relationnelle. Ensuite, l'article passe en revue quelques contre-exemples classiques à la thèse selon laquelle toute valeur finale serait nécessairement intrinsèque. Troisièmement, l'article discute de la relation entre valeurs finales, contributives et signatives. Quatrièmement, il examine la défense de l'exigence selon laquelle toute valeur finale devrait être aussi intrinsèque élaborée par Zimmerman. Et finalement, l'article explore l'exigence alternative selon laquelle une chose ne pourrait avoir de valeur finale qu'en vertu de ses propriétés essentielles.

INTRODUCTION

This paper examines the resources that fitting-attitude analyses of value (hereinafter ‘FA analyses of value’) offer for dealing with a key issue in formal axiology: whether something may have *final* value in virtue of its *relational* or *extrinsic* properties¹. In less technical terms, this issue concerns whether a thing can be valuable *for its own sake*, or *as an end*, by virtue of the relations it entertains with other things. This issue has its classical origin in G. E. Moore’s interpretation of his isolation test, according to which “[i]n order to arrive at a correct decision on the first part of this question [i.e., the question “What things have intrinsic value?”], it is necessary to consider what things are such that, if they existed by themselves, in absolute isolation, we should yet judge their existence to be good” (Moore 1988, p. 187). Moore’s idea is that, in such an isolation situation, a thing would retain all and only its *intrinsic* properties—that is, its *non-relational* properties, those that are *internal* to the thing². Moore then goes on to present the isolation test as a way of demarcating what is valuable *as an end* from what is a mere *means* to value:

By employing this method, we shall guard against two errors, which seem to have been the chief causes which have vitiated previous conclusions on the subject. The first of these is (1) that which consists in supposing that what seems absolutely necessary here and now, for the existence of anything good—what we cannot do without—is therefore good in itself. If we isolate such things, which are mere *means* to good, and suppose a world in which they alone, and nothing but they, existed, their *intrinsic worthlessness* becomes apparent (Moore 1988, p. 187, italics mine)³.

Moore’s rationale seems to be that since the relations that a thing entertains with other things would be absent under the isolation test, the value that a thing possesses *for its own sake* (as an end) must therefore be one that it has in virtue of its intrinsic properties. Hence, Moore seems to take the concepts of *final* and *intrinsic* value to be equivalent; consequently, in his view, a thing may only be valuable as an end in virtue of properties that are intrinsic to it. I will call this thesis *Moore’s intrinsicity principle*.

Moore’s intrinsicity principle has recently become the focus of a debate opposing, on the one hand, Korsgaard (1983), O’Neill (1992), Kagan (1998), and Rabinowicz and Rønnow-Rasmussen (2003; 1999), all of whom have raised serious objections against the principle, and on the other, Zimmerman (2001a; 2001b, Chap. 3), who has attempted an indirect defense of the principle by means of an argument about bearers of value⁴. This paper examines this debate from the perspective of the fitting-attitude analysis of value, which analyses the notion of value in terms of attitudes that it would be “fitting” for agents to adopt towards objects and situations of the world⁵. The thesis of the paper is that, from the perspective of the FA analysis of value, Moore’s intrinsicity principle is implausible, but that “the *essentiality* principle,” a distinct but not unrelated principle, may nevertheless be valid. Section 2 introduces the key elements of FA analy-

ses and sets aside an obvious but unimportant way in which these analyses make all final values relational. Section 3 presents classical cases put forward by Korsgaard, Kagan, O'Neill, and Rabinowicz and Rønnow-Rasmussen as counterexamples to Moore's intrinsicity principle. Section 4 discusses the relation between final value and some alleged kinds of non-instrumental extrinsic values (mainly contributive and signatory values). Section 5 examines Zimmerman's defense of Moore's intrinsicity principle through a reductionist manoeuvre he attempts along with Olson (2003). Lastly, section 6 proposes and illustrates the *prima facie* plausibility of the *essentiality* principle as an alternative to Moore's principle.

FA ANALYSES AND THE CONSTITUTION RELATION

FA analyses of value propose that to be valuable is to be a fitting object of an approving attitude (a *pro-attitude*, for short). According to such analyses, to say that something is valuable amounts to saying that it is such that it would be fitting to favour it, the notion of fittingness being generally understood in deontic terms, and the notion of favouring being usually interpreted in terms of attitudes, whence the name "fitting-attitude analyses." As its advocates typically emphasize, FA analyses thus include two main components: an *attitudinal* component (usually sentiments or emotions felt towards the object), and a *normative* component (the *fittingness* of the attitude)⁶. The attitudinal component in FA analyses portrays value concepts as concepts that, like colors and sounds, cannot be characterized independently of humans' subjective responses. This makes value concepts response-dependent, and so gives FA analyses some affinity with emotivist and other non-cognitivist metaethical theories which spell out evaluative judgments in terms of the sentiments and attitudes that evaluators *actually* experience. FA analyses dissociate themselves from these views by also including the idea of fittingness, thereby putting some distance between what evaluators *actually* feel and what is fitting to feel, that is, what they *ought* to feel towards the objects they contemplate⁷. Though FA analyses can be compatible with strong forms of non-cognitivism (depending on how the concept of fittingness is spelled out), most current versions of FA analyses conceive the fittingness of attitudes as grounded in the natural properties of the objects to which they respond. Their normative component thus grants values some independence from the evaluators' responses, and so gives FA analyses some affinity with realist metaethical theories⁸.

Three main advantages have made FA analyses appealing to many metaethicists. The first one is the simple fact that this approach actually provides an analysis of value concepts. This is not negligible: after Moore's well-known verdict that goodness is an unanalyzable concept, many value theorists simply gave up on the project of formulating an analysis of evaluative concepts. Moreover, FA analyses give such an account in a way that may relieve value properties of the air of ontological *mysteriousness* often attached to them (cf. Mackie 1977). Our sentiments and attitudes are familiar features of the world, and insofar as their fittingness can be grounded in the natural properties of the objects to which they respond, FA analyses may succeed providing an account of value that fits into a

naturalistic understanding of the world⁹. A second advantage of FA analyses is their relative metaethical neutrality. Although FA analyses exclude metaethical views that reject one or the other of their two defining components, the variety of ways in which these components can be construed leaves room for a plurality of metaethical interpretations, ranging from more realist-leaning to more anti-realist-bound ones. Such flexibility is welcome, given that realist and antirealist metaethical theories both have their strengths and shortcomings, while the debate between them is unlikely to be settled soon. And a last advantage of FA analyses is that they are also flexible at the level of axiology. FA analyses can easily accommodate the various and manifold nuances of our evaluative practices, as it allows that there may be as many distinctive ways to value something as there are shades of attitudes that an evaluator could adopt. For instance, to admire something is not equivalent to respecting it, to desire something is not the same as to be awe-inspired by it. And same goes for disvalues: for instance, to be angered by something is not equivalent to feeling contempt for it. Moreover, FA analyses are also axiologically flexible as to which kinds of thing can be proper bearers of values. Some attitudes, like desiring and hoping, seem more paradigmatically directed towards states of affairs, while others, like respecting and caring, seem more appropriate towards objects like persons or other kinds of entities. Thus, FA analyses have strong *prima facie* plausibility¹⁰. This *prima facie* plausibility suggests, in turn, that FA analyses provide a suitable perspective from which to examine the plausibility of Moore's intrinsicity principle.

It may be judicious to begin this examination by considering a very broad and general way in which FA analyses may seem to contradict Moore's principle. Given the significant role that FA analyses assign to the attitudes of evaluators in the constitution of values, it may indeed seem that such analyses unavoidably make *all* properties that have some relevance to values relational *from the outset*. This would have the consequence that *all* values are relational, running radically counter to Moore's intrinsicity principle, but in a way that seems suspiciously beside the point. As Rabinowicz and Rønnow-Rasmussen note:

One might wonder, of course, whether the claim that some values are intrinsic is compatible with the FA style of value analysis. If an object is valuable only insofar as there are reasons to have pro-attitudes toward it, then, one might argue, all value is relational by its very nature: the value of an object consists in the existence of a deontic relation between that object and the potential attitude holders. This would suggest that no value is intrinsic, since all value is relational. [...] In one sense, any value on the FA view seems to be a relational property, since it requires external attitudes toward the value bearer. But at the same time, if a given value-property of an object is grounded in its internal features, then that property even on the FA view appears to be context independent: the fitting pro-attitudes toward the object remain fitting as long as its internal features remain the same, however the external context might change (Rabinowicz and Rønnow-Rasmussen 2004, p. 409).

Rabinowicz and Rønnow-Rasmussen's suggestion that there is nevertheless a sense in which values may be non-relational according to FA analyses can be elaborated by appealing to a distinction that they make in their 1999 paper. There, they distinguish between what they call the *supervenience* base of values and the *constitutive* base of values:

The idea of a source of value is ambiguous between at least two interpretations: on the one hand, one may be thinking of the features of an object on which its final value *supervenes* (its "good-making" properties). If these features are internal to the object, i.e., non-relational, then the relevant value is intrinsic, as we are using this term. On the other hand, one may have in mind the *constitutive grounds* of an object's final value. The latter may well lie outside the object itself even though the former are internal to the object (Rabinowicz and Rønnow-Rasmussen 1999, pp. 117-118).

Thus, although FA analyses involve some relation between values and evaluators, this relation may not be the one at stake in Moore's intrinsicity principle. The former relation, it would seem, pertains to the *constitutive* base of values, while the relationality of properties at stake in Moore's principle pertains to their *supervenience* base.

This can be intuitively grasped by noting that although an FA analyst would consider that the existence of the final value of, say, a beautiful sunset partly depends upon the responses of evaluators, she would not be ready prepared to say that the sunset has its value *in virtue of* this response. The FA analyst would rather say that the beautiful sunset has its final value in virtue of the natural properties that a beautiful sunset has independently of the evaluator's evaluation. The "in virtue of" relation here points to what Rabinowicz and Rønnow-Rasmussen call the supervenience base of values. They specify that the two bases concern two distinct subfields of ethics which must not be conflated: "Claims about the supervenience bases of value belong to axiology, while claims concerning the constitutive grounds are perhaps best seen as belonging to metaethics, even though the boundary between these two disciplines is not as clear-cut as one might wish" (Rabinowicz and Rønnow-Rasmussen 1999, p. 118). Claims about the supervenience base concern axiology because that subfield develops theories about the features in virtue of which things can count as valuable, whereas claims about the constitutive base concern metaethics because that subfield constructs theories about the epistemological and ontological source of values. It is in this latter sense that FA analyses make values relational to evaluators insofar as they make values partly dependent upon evaluators' attitudinal responses. Hence, the relationality of values involved from the outset by the response-dependence of values in FA analyses does not in fact contradict Moore's intrinsicity principle. In order to assess the validity of this principle, one must examine cases where the relations upon which values depend are located in their supervenience rather than in their constitutive base. The following sections examine cases of this kind.

ARE THERE NON-INSTRUMENTAL EXTRINSIC VALUES?

Korsgaard (1983), O'Neill (1992), Kagan (1998), and Rabinowicz and Rønnow-Rasmussen (1999) discuss a series of cases where seemingly final values supervene on relational properties. A first case concerns the value that some things seem to have in virtue of their rarity. As Kagan remarks: "Many people, I think, are attracted to a view according to which the intrinsic value of an object depends in part on how rare that object is, or (in the limiting case) on its being completely unique. Obviously enough, however, uniqueness is not a property that an object has independently of whatever else may exist in the world" (Kagan 1998, p. 102). And as O'Neill notes, rarity as a ground for final value seems to have central importance for the issue of endangered species in environmental ethics:

If any property is irreducibly relational, then rarity is. The rarity of an object depends on the non-existence of other objects, and the property cannot be characterized without reference to other objects. [...] The preservation of endangered species of flora and fauna and of unusual habitats and ecological systems is a major practical environmental problem. Rarity appears to confer a special value to an object (O'Neill 1992, p. 124).

As O'Neill remarks, the value based on rarity is directly tied to another major value in environmental ethics, namely diversity, which also involves relations. O'Neill mentions yet another relational property that many environmental ethicists take as a ground for final value: "one might value wilderness in virtue of its not bearing the imprint of human activity [...] To say 'x has value because it is untouched by humans' is to say that it has value in virtue of a relation it has to humans and their activities. Wilderness has such value in virtue of our absence" (O'Neill 1992, p. 125). In a much less serious vein, Rabinowicz and Rønnow-Rasmussen (1999, p. 121) introduce the hypothetical case of the final value that a dress may have in virtue of its relational property of having belonged to Princess Diana. All these examples present cases where some objects appear to be valuable for their own sakes in virtue of relational properties. In so doing, they constitute potential counterexamples to Moore's intrinsicity principle. But are these genuine cases of relational final value?

A first indication that they are indeed is the phenomenological evidence that the evaluative attitudes experienced by the evaluators in these cases appear to be authentic attitudes of final valuation. Rabinowicz and Rønnow-Rasmussen argue on these grounds that it would be implausible to interpret such examples as cases of instrumental evaluations in disguise:

Diana's dress is perhaps valuable merely as a means: merely because it allows us to establish an indirect connection to a person we admire or find important in one way or another. Having such a connection may be something that we set a final value on. Couldn't this be what is going on here? Not necessarily. Even if the desire to establish such an "affiliation" with Diana may well be a part of the causal explanation of our

evaluative attitude towards the dress, this does not imply that the evaluative attitude itself is of the instrumental kind: if we idolise Diana, we do not simply find the dress useful for some purpose; we ascribe an independent value to it. Compare this with O'Neill's example: the wilderness is not simply instrumental in allowing us to come into contact with something (otherwise) untouched by humans. Even if we could never visit the wild area, it would still keep its value from our point of view (Rabinowicz and Rønnow-Rasmussen 1999, p. 121).

To be sure, the evaluative attitudes in these cases seem to be the very opposite of instrumentalizing ones. From the perspective of FA analyses, such phenomenological evidence should be assigned much weight, I think, given the close psychological and conceptual connection that these analyses establish between values and attitudes. Unless Moore's intrinsicity principle can be shown to have at least as much *prima facie* plausibility as the FA style of analysis, and given the striking abundance of cases similar to those presented by Korsgaard, Kagan, O'Neill and Rabinowicz and Rønnow-Rasmussen, it seems legitimate to shift the burden of proof onto the supporters of Moore's principle. The next section will show that Moore's intrinsicity principle may in fact have something closer to *prima facie* implausibility, for, as we will see, this principle requires, to earn some minimal plausibility, that its supporters accept the problematic epicycle of Moore's theory of organic unities.

CONTRIBUTIVE, SIGNATORY AND INSTRUMENTAL VALUES

Besides the phenomenological evidence yielded by paradigm examples, some additional evidence against Moore's principle may be found when considering that, as O'Neill (1992, pp. 124-125) argues, the impression that all relational values must be instrumental may just stem from an equivocal habit of language which designates two different concepts—final value and non-relational value—by the same term, “intrinsic value.” Bradley (2001, pp. 49-50; 1998, pp. 109-110) offers some resources for this purpose when he proposes that instrumental value may just be one special case among others of extrinsic value. On Bradley's proposal, besides *instrumental* value, which he defines as a value that results from a thing's “*causal* relationships with other things” (Bradley 2001, p. 49, italics mine), there exist at least two other kinds of extrinsic value: *contributive* (or *contributory*) value, i.e., “the value something has in virtue of being a part of a valuable whole,” and *signatory* value, i.e., the value something has “because of what it signifies” (Bradley 1998, pp. 110-111). On this view, the value that things have in virtue of their rarity may be labeled as *contributive* insofar as this value results from these things' contribution to the intrinsic value of a larger whole (by making it more diverse, as Bradley suggests). The value of a wilderness area's not bearing the imprint of human activity may also be viewed as *contributive* in that it results from preventing the intrinsic disvalue which would result from some human presence in it¹¹. And the value of Diana's dress may be depicted as *signatory* in that it results from what the dress represents for a Diana idolizer. Thus, on this picture, the value of rare things, of wilderness areas, and of Diana's dress would be *relational* yet without thereby being *instrumental*.

Would this confer *final* value on rare things, wilderness areas, and Diana's dress? In other words, what happens to final value according to Bradley's "de-binarization" of the relation between intrinsic and instrumental value? Is some value final once it is not instrumental, including contributive and signatory values as final ones? Or does it remain the case that only intrinsic (non-extrinsic) values can be final? This question seems undecidable on the sole ground of Bradley's taxonomy of relations. Regarding the *contributive* relation, this raises an issue brought up by Moore's well-discussed principle of organic unities. Moore developed this principle in part to explain how the same thing—e.g., beauty, which, he thought, has some final value in isolation—could generate more final value in contexts where it is contemplated and appreciated. Such an increase in final value seems to violate his intrinsicity principle, as beauty then seems capable of generating some amount of final value in virtue of a relational property, namely its property of being contemplated and appreciated. To remain consistent with his intrinsicity principle, then, Moore developed his theory of organic unities, which explains that when contemplated and appreciated, the beauty involved retains the exact same amount of final value that it has in isolation, and that the increase in final value generated by its being contemplated and appreciated belongs to the whole formed by the complex "beauty + contemplation," rather than to beauty itself. Thus, in agreement with the intrinsicity principle, the contemplation and appreciation of beauty may increase the total final value of the world, yet the final value of beauty itself remains unchanged by such a relation.

However, not everyone agrees with this interpretation of the increase in final value generated by organic unities. Hurka (1998) delineates two possible interpretations of this value, which he calls the *holistic* and *conditional* interpretations. Under the holistic interpretation, which is the one to which Moore subscribes, the final value added by the combination of parts into a whole is, as we've seen, located strictly in the whole, and so the final value of the associated parts themselves remains the same. Under the conditional interpretation, in contrast, "the intrinsic value of a state can change when it enters into a larger whole, so its value or degree of value is altered by its relations to other states" (Hurka 1998, p. 303)¹². In the case of the contemplation of beauty, this would imply that the final value of a beautiful thing is itself increased when it is contemplated. In sum, on the holistic view, there is only *upward* determination of the final value of a whole by its parts in association, while on the conditional view, there is *downward* determination of the final value of the parts by the whole formed through their association.

The question now is: Which interpretation is the correct one? In fact, both views seem plausible depending on which examples they are applied to. Bradley (2002) makes a convincing case, I think, that the final value generated by the addition of rare parts into a whole (by making it more diverse) must be interpreted holistically, that is, as belonging to the whole without affecting the parts:

Suppose, for example, that A is a beautiful painting, that B is a painting exactly like A, and that C is a beautiful piece of music. The aesthetic contemplation of A may have the same [intrinsic] value as that of B

and also the same [intrinsic] value as that of C. But the whole that is the aesthetic contemplation of A followed by that of C is intrinsically better than that whole that is the aesthetic contemplation of A followed by that of B. Hence one could say that the value of a *bonum variationis* is greater than the sum of the values of its constituent parts (Chisholm 1986, p. 71; cited in Bradley 2002, p. 34).

By Bradley's (2002, pp. 39-40) own admission, however, this accounts only for some organic unities and is therefore no proof of the *general* validity of the holistic view over against the conditionalist one. Counter to Bradley, Olson (2004) formulates a convincing defense of the conditionalist view's superiority for some other cases like that of *wicked pleasure* and *compassionate pain*. As Olson notes, the difference between the final value of "John's being pleased" and the final disvalue of "John's being pleased at Mary's pain" is better accounted for by the conditionalist interpretation, for it seems implausible to say that the final value of "John's being pleased" is the same in both cases. The total final value of "John's being pleased at Mary's pain" is lower than that of "John's being pleased" in isolation because "John's being pleased" itself is made disvaluable in a context where it is a case of wicked pleasure. And similarly for the difference between the final disvalue of "John's being pained" and the final value of "John's being pained at Mary's pain," where "John's being pained" becomes valuable (or at least less disvaluable) in a context where it is a case of compassionate pain (cf. Olson 2004, pp. 37-41). Thus the right solution to the holist/conditionalist debate about Moorean organic unities seems to be: *It depends!*

To see on what it may depend, it can be informative to look at Korsgaard's discussion of the contributive value of beauty. Criticizing Moore's own treatment of this case and defending a conditionalist interpretation of it, Korsgaard complains that

Moore's view, and the intuitionistic method of isolation, veil or obscure the internal relations within the organic unity in virtue of which the organic unity has its value. Whereas the Kantian account, which focuses on rather than ignoring the internal relations of the valuable whole, allows us to see why happiness is valuable in just this case and not in another case. Moore can only say that the combination of happiness and good will works (is a good recipe, so to speak) while happiness plus the bad will does not. Kant can say that happiness in the one case is good because the condition under which it is fully justified has been met (roughly, because its having been decently pursued makes it deserved). Those internal relations reveal the *reasons* for our views about what is valuable, while Moore's view tends to cover up these reasons (Korsgaard 1983, p. 95).

The last sentence of this passage is the most important one, for it states the rationale behind Korsgaard's preference for the conditionalist view. The conditionalist interpretation is better in this case, according to Korsgaard, because it makes the supervenience base of values more transparent—it shows the reasons

which may ground the evaluators' appropriate evaluations. This rationale resonates well with the picture given by FA analyses of the relation between values and their supervenience base. Thus it seems that Korsgaard indicates a response available to FA analysts for addressing the holist/conditionalist issue. Fitting-attitude analysts may tackle this issue in the very spirit of their approach simply by stating that whether holism or conditionalism should be the privileged interpretation should be determined on a case-by-case basis according to what our fitting evaluative attitudes respond to. Thus, FA analysts can simply state that holism should apply just when the added value grounds an evaluative response that is directed towards the whole, whereas conditionalism should apply just when the added value grounds an evaluative response that is directed towards the part. On this picture, the contributive relation between rare things and the whole to which they belong would increase the final value of the whole rather than of the part by making that whole, but not the parts, more fitting for an attitude of, say, admiration (responding to the increased variety of the whole). For instance, the contributive character of the absence of humans' imprint in a wilderness area would increase the final value of the area as a whole, but not that of the human absence itself in this context, by making the wilderness area as a whole fitting for an attitude of, say, respect. Returning to the signatory value of Diana's dress, the corresponding implication would be that the dress may acquire final value from its signatory relation insofar as this signatory relation makes it a fitting object of, say, a treasuring attitude.

Such a result, however, runs partly counter to Bradley's endeavour when he delineates kinds of non-instrumental extrinsic values. The FA analytic picture proposed here has the corollary that a thing's being in a *causal* relation with something else is not a necessary condition for this thing's having *instrumental* value. When our fitting attitudes determine the location of the final value generated in organic unities, they also thereby determine the location of some instrumental value. In the case of the value that a rare thing adds to a whole, for instance, our fitting attitudes assign some additional final value to the whole, thereby assigning *instrumental* value to the parts (regardless of the fact that this value does not supervene on causal relations). Similarly, in the case of the value that the absence of humans adds to a wilderness area, our fitting attitudes assign some final value to the area, thereby assigning *instrumental* value to the absence of humans (regardless of the fact that humans' absence *constitutes* rather than *causes* the existence of the wilderness area). This picture thus retains—at the *attitudinal level*—the dualistic opposition between final and instrumental values that Bradley sought to dissolve (though this account also acknowledges his observation that relational properties need not be causal).

But there is more besides contributive and signatory final values. Kagan and Korsgaard present a plethora of allegedly paradoxical cases where some things acquire final value in virtue of their *usefulness*, suggesting that final value may even supervene on *instrumental* relations! Kagan (1998, pp. 102-104) mentions “an elegantly designed racing car,” “excellence in various practical arts,” and “the pen used by Abraham Lincoln to sign the Emancipation Proclamation, free-

ing the slaves”; and Korsgaard (1983, p. 89) enumerates “[m]ink coats and handsome china and gorgeously enameled frying pans¹³.” In these cases as much as in those of rare things, wilderness areas and Diana’s dress, the phenomenological evidence indicates that the evaluators’ attitudes at least sometimes consist in genuine final evaluations. And the existence of such cases pleads even more strongly in favour of my proposal (*pace* Bradley) to define instrumental value not on the basis of the kinds of relations (causal, contributive, etc.) occurring in the supervenience base, but rather on the basis of the attitude befitting the thing in its relational context. In fact, the term “instrumental value” seems to have two meanings: first, a *relational* meaning (Bradley’s sense), which signifies roughly a thing’s *causal* properties, that is, its ability to cause something else; and second, an *attitudinal* meaning, which signifies a person’s fitting stance towards a thing. At the attitudinal level, final value occurs when a thing is the ultimate target of a person’s fitting evaluative attitude, and instrumental value occurs when the thing is valuable *for the sake of* something else¹⁴. With this distinction in mind, the cases of final value grounded in instrumental value can easily be explicated if one pays attention to the difference between the relations “in virtue of” and “for the sake of.” While the “in virtue of” relation, as we have seen, indicates the supervenience base of evaluative attitudes, the “for the sake of” relation indicates what one may call the *target* of evaluative attitudes. And what has been shown by cases of contributive and signatory values applies in exactly the same way to instrumental values in the relational sense: our evaluative attitudes are not dictated univocally by the kind of relations characterizing the properties *in virtue of* which they are grounded. Just as nothing precludes John’s pain from being valuable for its own sake in virtue of its contribution to an occurrence of compassionate pain, nothing precludes, say, handsome china from being valuable for its own sake partly in virtue of its usefulness. We have seen in the last paragraph that attitudinal instrumental value can supervene on types of extrinsic values that are not relationally instrumental. The independence also goes the other way around: relational instrumental value does not dictate attitudinal instrumental value.

This FA analysis-inspired treatment of contributive, signatory, and relational instrumental values thus pleads strongly in favour of the genuineness of the examples of relational final value presented by Korsgaard, O’Neill, Kagan, and Rabinowicz and Rønnow-Rasmussen. The next section examines a sophisticated attempt to defend Moore’s intrinsicity principle against these examples.

ZIMMERMAN AND OLSON’S REDUCTIONIST STRATEGY

One attempt to defend Moore’s intrinsicity principle is the reductionist strategy employed by Zimmerman (2001a; 2001b, Chap. 3) and Olson (2003). This strategy consists in arguing that the value of *objects*—such as rare things, wilderness areas and Diana’s dress—can be reduced to the value of *states of affairs* or *tropes*. On this conception, the value of Diana’s dress (an object) would be reducible to that of the state of affairs “This dress belonged to Diana” or that of the trope “This dress’s having belonged to Diana.” As states of affairs and tropes have the same implications for my purposes, I will speak uniformly of “states¹⁵.”

Zimmerman and Olson's reductionist attempt formulates a defense for Moore's intrinsicity principle by building on the observation that the *relational* properties of objects are, at the same time, *intrinsic* properties of the states in which they are involved. This observation paves the way for an indirect defense of Moore's principle: if it can be shown that what looks like the relational final value of an object can always be reduced to the state in which this object occurs, then all of the (seemingly) *relational* final values of objects will ultimately be *intrinsic* final values of states. This would confirm Moore's intrinsicity principle by making it *necessarily* true, regardless of whether the value supervenes on some properties that are relational to the objects involved in the finally valuable states.

Before examining Zimmerman and Olson's proposal, it may be relevant to consider an objection that can be addressed to it from the outset, because doing so will help us to grasp what exactly is at stake. This objection has been very well formulated by Kagan, who anticipated a reductionist attempt equivalent to the one later proposed by Zimmerman and Olson:

It might be suggested, however, that although it is a common enough practice to view objects as the bearers of intrinsic value, it is nonetheless preferable to hold that facts (or, perhaps, states of affairs) are the only genuine bearers of intrinsic value. [...] [I]t is not implausible to suggest that it is an intrinsic property of a given fact that it concerns the specific objects and properties that it does. That is, it would not be implausible to claim that it is an intrinsic property of the *fact* that there exists a pen which was used to sign the Emancipation Proclamation [...] If we accept this claim about the intrinsic properties of facts [...] and we then combine it with the earlier claim that strictly speaking only facts are the bearers of intrinsic value, then the following result emerges: one can accept the thrust of all of my examples, while still accepting the dominant philosophical tradition that intrinsic value turns solely upon intrinsic properties (Kagan 1998, p. 111).

As Kagan points out, a proposal like Zimmerman and Olson's seems to save Moore's intrinsicity principle at quite a high price: on such a proposal, Moore's principle loses its regulatory role for evaluations involving objects. So what is at stake with Zimmerman and Olson's reductionist attempt is not so much Moore's principle itself as a debate over the very possibility that objects can be genuine bearers of values. As will be seen shortly, Zimmerman and Olson's attempt leans towards the result that anytime one thinks one is valuing an object for its own sake, one is in fact valuing it somehow instrumentally. This would make many evaluations *delusively* final.

For the sake of simplicity, I will begin by focusing on Zimmerman's version of the reductionist argument and later turn to Olson's only where his version usefully supplements Zimmerman's. Zimmerman's reductionist attempt starts by making the claim that final value must be *nonderivative*:

Consider the claim that pleasure is good for its own sake. I take this to mean that every state of pleasure, every state consisting of someone's being pleased, is good simply in virtue of being such a state. There is no helpful explanation why the state is good; it just is good "as such," that is, good in virtue of its own nature. But though unhelpful, in that this account of the goodness of pleasure does not cite something else in terms of which the goodness of pleasure may be understood, it does, contrary to the first objection, provide some insight into the nature of final value. Such value is *nonderivative*; it is the ground or source of nonfinal values (such as those of charity and of hitting someone on the head with a hammer), values that may thus be declared derivative. All explanation must come to an end somewhere; the explanation of values stops with the citing of final values (Zimmerman 2001a, p. 193; see also 2001b, p. 37).

So Zimmerman's test is simple: any value that can be explained in terms of another one is *derivative*. Zimmerman's purported evidence for the claim that only those things that are nonderivatively valuable can be finally valuable is that the inexplicability of their value indicates that they are valuable "as such," or "in virtue of their own natures." Zimmerman reasons as follows: "where no helpful explanation as to why something is good is available, this is because the thing in question just is good 'as such,' that is, good in virtue of its own nature. Given that something's nature is intrinsic to it, my contention and suggestion jointly imply that *nonderivative value is intrinsic to its bearer*" (Zimmerman 2001a, p. 194; see also 2001b, p. 38). Zimmerman then mobilizes this observation against the examples presented by Korsgaard, Kagan, and Rabinowicz and Rønnow-Rasmussen:

If we were to ask these authors why the objects in question are good, we might expect, given what was said in the last paragraph, that they would simply answer, "They just are good 'as such.' They're good in virtue of their natures." But this is not what they say. Korsgaard attributes the value of the objects she mentions to their "instrumentality," that is, to their helpfulness in allowing us to accomplish certain tasks. Kagan attributes the value of Lincoln's pen to the unique historical role it played. Rabinowicz and Rønnow-Rasmussen attribute the value of Diana's dress to the fact that it belonged to Diana (Zimmerman 2001a, p. 194; see also 2001b, pp. 37-38).

I suspect that I am not the only one to remain unconvinced by Zimmerman's nonderivativity criterion. But as his whole argument hinges on it, let's grant it for now and see how Zimmerman deals with an objection raised by Rabinowicz and Rønnow-Rasmussen from the perspective of FA analyses of values:

[I]f we accept the general idea that value is what calls for an appropriate response, the response in question need not consist in just preferring or promoting. There may well be other alternatives: preference is not the only attitude to be considered, nor is promoting the only behav-

our that may be relevant in this context. [...] Alternatively, one might relate value to thing-oriented behaviours, or perhaps better, to thing-oriented attitudes-cum-behaviours: value is what we should cherish, protect or care for. [...] In fact, when one realises how many various types of responses could be relevant in this context, it becomes tempting to draw the conclusion that any monistic analysis of value in terms of one particular type of response would be inadequate. [...] Given such a pluralist approach to value analysis, the main motivation to reduce thing values to state values disappears. Valuable things may be objects that call for specific thing-oriented attitudes or behaviors: a wilderness untouched by human hands calls for protection, Diana's dress is an object to be cherished and preserved, and so on (Rabinowicz and Rønnow-Rasmussen 1999, pp. 124-125).

Zimmerman explores two possible responses to this objection¹⁶. The first consists in denying that Rabinowicz and Rønnow-Rasmussen's thing-oriented attitudes cannot also apply to states:

Notice that in many cases the attitudes that are directed toward individual objects may also be directed toward other things as well. I may love, admire, and respect someone, but I may also love, admire, and respect what he does. The fact that I have these pro-attitudes toward him doesn't itself show that he has final value, for these attitudes may derive from the attitudes I have toward what he does. I may admire what John does for its own sake (for example, I may admire his display of courage), and I may thus admire him for what he does (I may admire him for his display of courage), but this doesn't show that I admire John for his own sake (Zimmerman 2001a, p. 197; see also 2001b, p. 41).

Here, I think, Zimmerman remains unconvincing. First, it would be fairly plausible, I think, to say that talk of admiring or respecting what someone does are just elliptical ways of speaking of admiring and respecting the persons themselves (in virtue of their actions). Moreover, even assuming that these ways of speaking are literal, it remains unclear that, for instance, the admirable character of John's acts cannot—as would however seem to be the case—somehow *infuse* or *insufflate* some final value into John himself and make him admirable. In order to deal with cases where different evaluative attitudes apply to the object and the state, Zimmerman attempts his second strategy:

What exactly is it about the idea that individual objects have a different sort of value from the value that states have that requires us to say that the former value is final? Why could it not be, for instance, that, even if Diana's dress is to be treasured but its having belonged to Diana is to be valued in some other way, still the dress is to be treasured, not for its own sake, but for the sake of the state in question? That one entity derives its value from another would not appear to require that the two have exactly the same type of value (Zimmerman 2001a, p. 200; see also 2001b, p. 44).

The problematic character of this response seems even more obvious. In principle, a reduction implies that the conversion of one thing into another has no residue (no loss of information, one could say). But if Zimmerman concedes that there is some incommensurability between the input and the output of the reduction, then he seems thereby to recognize that the reduction fails due to incompleteness. Rabinowicz and Rønnow-Rasmussen make this point quite forcefully:

It seems that certain pro-attitudes that fit things or persons are not fitting or perhaps may even be impossible to hold with respect to tropes. Thus, for example, consider the attitude of respect. I can respect a person but it sounds odd, to say the least, to say that I respect a trope. I can respect Ann for her courage, but this is an attitude I hold towards Ann, and not towards her courage. [...] Conversely, some appropriate pro-attitudes towards tropes are not appropriate towards things or persons. For example, I can be exhilarated by a display of courage in Ann but it would be wrong to say that I am exhilarated by Ann, even though I may well come to value her more on account of her courage. Similarly, I can rejoice in her happiness, but I cannot rejoice in Ann herself. These examples suggest that the value of a concrete object (of a person or a thing) cannot just consist in the value of the corresponding tropes. [...] Consequently, the value of a concrete object and the value of a trope must be different from each other if the concrete object and the trope call for different responses (Rabinowicz and Rønnow-Rasmussen 2003, p. 222).

At this point, it will be illuminating to integrate some resources from Olson's version of the reductive argument. In response to Rabinowicz and Rønnow-Rasmussen, Olson (2003, p. 416) introduces a telling qualification concerning the kind of reduction pursued here. In his view, this reduction is not one of *equivalence*, but of *specification*. Thus, the reductive relation goes only one way: it says that whenever an object has value, this value is rooted in the value of a state, but it says nothing about when a state's value engenders an object's (non-final) values. This qualification may remove the requirement that the reduction have no residue. Yet the question remains: why would such rooting imply that only the state *really* has final values while the object is left with "merely *derived* values" (Olson 2003, p. 417)? Olson's answer brings us back to Zimmerman's initial contrast between derivative and nonderivative values, but through a formulation that reveals the very presupposition behind it:

Recall that we characterised final value as that which is valuable for its own sake. Now, if a person is valuable in virtue of, or because of, or for the sake of, her courage, this person is not valuable for her own sake, but rather for the sake of something else, namely her courage—the trope—which doubtless is a component of the concrete person, but a contingent one. It would be an easy thing to imagine this person lacking her courage, and thus this value (Olson 2003, p. 418).

Notice that in this passage, Olson lumps together the two types of relations that I have argued are distinct: the “for the sake of” and the “in virtue of” relations. Thus, Olson treats the relation that refers to the *target* of the evaluation as equivalent to the relation that points to its *supervenience* base. But if one does this, of course, the result will be that as soon as one values something *in virtue of* some of its features, one will thereby be valuing it *for the sake of* these features rather than *for its own sake*. This, I think, is an invalid way of denying that objects may be valued for their own sakes. For more illustration, take Olson’s ski example:

If one would go on asking the skier further questions about why he values or appreciates skiing, or what it is about skiing that he values or appreciates, one would presumably find out that there is (are) some feature(s), characteristic of skiing that the skier values or appreciates for its (their) own sake(s). The immediate object(s) of his evaluative attitude(s) would then be this (those) characteristic(s), rather than the skiing itself (Olson 2003, pp. 418-419).

Such an interpretation, I think, again conflates the *target* and the *supervenience* base of the evaluation. On Olson’s account, one really values something for its own sake only when one cannot explain *in virtue of* what features of the thing one values it. One’s being able to do so would automatically imply that one values the thing for the sake of these features rather than for its own sake. In other words, as soon as an explanation is available for a given evaluation, then the value becomes derivative and therefore nonfinal. This way of interpreting the “for the sake of” relation constitutes what Olson labels the “*strict* reading of the notion ‘value for its own sake’,” which he opposes to Rabinowicz and Rønnow-Rasmussen’s looser reading. I think that Olson’s strict reading is much too strict—indeed, *arbitrarily* strict, as will be shown shortly.

Recall that Zimmerman defends the nonderivativity criterion, which Olson here just assumes, on the grounds that applying this criterion guarantees that the value at stake really is *intrinsic to its bearer*. The criterion does this by guaranteeing that the bearer acquires its value solely in virtue of properties that are *intrinsic to its nature*. Such a defense, however, seems to conflate a *sufficient* with a *necessary* condition: nonderivativity’s guaranteeing that the value supervenes on properties intrinsic to their bearer’s nature does not imply that this happens *only when* nonderivativity occurs. In fact, the only difference between many cases of derivative value and what occurs with nonderivative value is the following: when the value is derivative, the nature of the thing in question is *transparent* to the evaluator, so that she can explain her evaluation by specifying the components of this nature that ground her evaluation; whereas when the value is nonderivative, the nature of the thing remains *opaque* to the evaluator. This opacity is, in fact, the only reason why the evaluator cannot explain her evaluation, and why her only option is then to say that she values the thing “in virtue of its own nature.” On these grounds, I take it that the derivative/nonderivative distinction does not reflect any fundamental contrast about how evaluations relate to their supervenience bases.

TOWARDS AN ESSENTIALITY PRINCIPLE?

My criticism of Zimmerman and Olson's nonderivativity criterion nonetheless concedes them something significant. It concedes them what may be called the *essentiality* principle (which must not be conflated with the *intrinsicity* principle). This principle states that a thing may only have final value in virtue of features that are *internal to its nature*. This principle underlies my rejoinder to Zimmerman and Olson that the value of a state insufflates value into an object when the state instantiates a property that is *internal to the nature* of the object. Before further explicating the essentiality principle, let's first observe how something like it seems to underlie Olson's worry that, were I to value someone for her courage, I would value her in virtue of something that is *contingent* to her (see the above quotation). Contingency seems to be the intuitive reason for Olson's worry. And there seems to be something right about this intuition: if someone values someone (or something) in virtue of properties which are contingent to that person (or thing), then it would seem right to say that the *target* of her evaluation cannot really be that person (or thing). But let's imagine that, in Olson's example, the person's display of courage can be interpreted as an enduring character trait, as a reflection of her deep personality, or of the internal nature of the person she is, so to speak. In such a case, then, it seems that Olson's worry disappears, and consequently that it is perfectly unproblematic to say that the person who is valued in virtue of her courage is valued for her own sake. And this is exactly what the essentiality principle predicts: insofar as this person is valued in virtue of features belonging to her own nature, she can be valued finally. Now take Olson's skiing example. I doubt that the features in virtue of which skiers enjoy skiing are so contingent to this activity; if there were no stable features of skiing that skiers are confident to find again and again when they go hurtling down the slopes, their general attitude towards skiing would be very hard to explain. If these stable features can be seen as part of the internal nature of skiing, however, it seems perfectly correct to say that the skiers value skiing for its own sake. Contrast this with a case where some people think that they enjoy skiing for its own sake in virtue of its being a good occasion to display their economic wealth by wearing ostentatiously expensive equipment. Presumably, such a display is not part of the internal nature of skiing, and so the essentiality principle would entail that these show-offs' evaluation is *delusively* final. And consider the case where I admire a butterfly so much in virtue of its beauty that I can't resist suffocating it so that I can preserve it forever in my collection. In such a case, my evaluative attitude is conceivably one of final evaluation: I value the butterfly *as an end* in virtue of its beauty. But given that I am sacrificing its life, presumably a property that is dearer to a butterfly's nature than is its beauty, the essentiality principle states that my attitude, like that of the show-off skiers, would have to be qualified as *delusively* final. To me, these upshots of the essentiality principle seem fairly intuitive. Although FA analysts want to give much credit to attitudinal responses, they should nonetheless, it seems to me, leave some room for a notion of *delusively final* evaluative attitudes (which are in fact concealed acts of instrumentalization).

Although, as I said, the essentiality principle is not equivalent to Moore's intrinsicity principle, it nevertheless has some kinship with it (see Bradley 2002, pp. 23-24; Rabinowicz and Rønnow-Rasmussen 1999, p. 116, p. 127). First, it seems to be a principle that Moore himself endorsed at times:

When I say, with regard to any particular kind of value, that the question whether and in what degree anything possesses it depends solely on the *intrinsic nature* of the thing in question, I mean to say [...] that it is impossible for what is strictly one and the same thing to possess that kind of value at one time, or in one set of circumstances, and not to possess it at another; and equally possible for it to possess it in one degree at one time, or in one set of circumstances, and to possess it in a different degree at another, or in a different set (Moore 2000, pp. 260-261, italics mine).

Moreover, it seems thinkable to interpret Moore's isolation test as aimed primarily at isolating *essential* properties. Given that states of affairs have their intrinsic properties necessarily, it follows that Moore's isolation test, when isolating the intrinsic properties of *states*, thereby isolates their essential ones. This equivalence does not, however, apply in the case of objects, for, as Rabinowicz and Rønnow-Rasmussen note, "[a] concrete individual, such as, say, Theseus' ship, may well have internal properties that are contingent, say, such features as its colour, or the property of containing as a part a particular plank, *a*" (Rabinowicz and Rønnow-Rasmussen 1999, p. 116). But Moore's intrinsicity principle may have in fact been meant to serve as an essentiality principle, intended to ensure that objects' value supervene not on their intrinsic but on their essential properties. Indeed, a remark by Rabinowicz and Rønnow-Rasmussen suggests that this interpretation might make Moore's views more intuitive on this matter:

[T]he intrinsic value of an object may well vary in different possible worlds, if this value supervenes on the object's *contingent* internal properties. In other words, one should not confuse intrinsic value with *essential value*, where the latter is the value that the object has by necessity, i.e., in every possible world in which it exists. This suggests, by the way, that the notion of an intrinsic value may not be as normatively interesting as many have thought. After all, what is so special about value that supervenes on the object's internal rather than relational properties, especially if the former may be just as contingent to the object as the latter? One can easily see the normative relevance of the notion of a final value (at least if such value is analysed in terms of a range of the fitting responses that the object calls for), but the concept of an intrinsic value seems to lack a special normative interest (Rabinowicz and Rønnow-Rasmussen 1999, p. 127).

At the very least, my discussion of Zimmerman and Olson's reductionist attempts suggests that an identification of final value with essential value has more *prima facie* plausibility than one with intrinsic value.

It must be noted that the essentiality principle is not necessarily more demanding than the intrinsicity principle. As far as objects are concerned, essential properties can very often be relational. For instance, many objects to which functions are assigned, like organs and tools, are defined by their causal relations with their bearers or users. Arguably, a spontaneously generated hammer on an uninhabited exo-planet is not really a hammer (see McLaughlin 2002, pp. 135-136), and likewise for a spontaneously generated orphan heart. Similarly, many symbolic objects, such as dollar bills and photographs, are defined by their signatory relations. And likewise, many things that exist as parts of wholes, like book chapters and train stations, are defined by their contributory relations. All of these cases involve relational essential properties, which, should they be taken as supervenience bases for some evaluations, would, under the essentiality principle, ground genuine final evaluations. Returning to the examples supplied by Kagan, O'Neill, and Rabinowicz and Rønnow-Rasmussen, adding a rare part to a diverse whole can make it more finally admirable *qua* diverse whole, the absence of humans in a wilderness area can make the area deserve more respect for its own sake *qua* wilderness, and a dress's having belonged to Diana can make it more finally treasurable *qua* something having belonged to a celebrity.

Admittedly, this way of applying the essentiality principle is quite liberal, as the “*qua*-clause” is malleable enough to turn any instrumental evaluation into a final one just by redefining a given token as belonging to a different kind. As liberal as it is, I think that this way of applying the principle retains a regulative role with crucial importance for delineating genuine and delusive final evaluations. As I just said, one may well treasure Diana's dress, in virtue of its signatory relation with the princess, for its own sake *qua* something that has belonged to a celebrity. But according to the essentiality principle as I apply it, one may not, in virtue of this relation, treasure the dress *qua* dress. Thus in this case, *qua* dress, the dress is instrumentalized for the sake of being an object that has belonged to a celebrity. Similarly, one may well, in virtue of its not bearing any human imprint, respect a wilderness area for its own sake *qua* instantiation of pure otherness, say. But according to the essentiality principle, one may not, in virtue of this feature, value it *qua* healthy ecosystem, which is not conceptually equivalent to wilderness. If however, the wilderness of the ecosystem is taken as signatory of its health, then the ecosystem can be valued for its own sake *qua* healthy ecosystem, in virtue of its condition of wilderness. But this applies to cases of values grounded in intrinsic properties as well. Suppose that I like my neighbourhood, “La Petite Patrie,” in virtue of the fact that most of my friends live there. Although in this case my friends' presence is an intrinsic property of the neighbourhood (my friends are constituents of it along with its streets, schools, and so on), it would nevertheless be incorrect to say that I like “La Petite Patrie” for its own sake in virtue of the fact that my friends live there. What I like for its own sake, in this case, is the neighbourhood where my friends live. And let me introduce one last case, one in which a person's evaluation undermines the nature of the thing valued. Recall my butterfly example: I may well admire a butterfly for its own sake *qua* collectible object solely in virtue of its beauty, a property which presumably belongs to its nature, yet taking such a restricted part of its nature as

the supervenience base of my evaluation somehow undermines what the butterfly is. This leads me to adopt an attitude (and a consequent behaviour) which is not responsive to the full nature of the butterfly *qua* living being, and so involves some degree of instrumentalization. The regulative role of the essentiality principle, in short, is to uncover the cases of delusively final evaluations, or to put it more sombrely, to unmask cases of instrumentalization in disguise.

CONCLUSION

This paper has examined the plausibility of G. E. Moore's intrinsicity principle from the perspective of fitting-attitudes analyses (FA analyses) of values. First, by recalling Rabinowicz and Rønnow-Rasmussen's distinction between the supervenience and constitutive bases of values, it argued that although FA analyses portray values as being partly constituted by the responses of evaluators, this relation does not make the ensuing value relational in a way that contradicts Moore's principle. Secondly, the paper introduced Korsgaard, Kagan, O'Neill, and Rabinowicz and Rønnow-Rasmussen's paradigm examples of objects which intuitively have relational final value, and argued that FA analysts should be responsive to the phenomenological evidence that these cases are indeed genuine. Then, turning to Bradley's distinction between instrumental, contributive, and signatory value, the paper argued that the relevant contrast between final and instrumental values must be situated at the attitudinal level rather than at the level of the relations occurring in the supervenience base, and furthermore that this allows for genuine cases of final value in virtue of contributive, signatory, and instrumental relations. Fourthly, the paper examined Zimmerman and Olson's reductionist argument, which contends that only states and not objects can be proper bearers of final value, and proceeded to show that these authors' nonderivativity criterion is arbitrary. Lastly, the paper explored the *prima facie* plausibility of the 'essentiality principle' as an alternative to Moore's intrinsicity principle. The latter exploration, indeed, offered just a first glance at the essentiality principle along with a cursory examination of its plausibility. Many issues still have to be tackled, but I hope that this paper has indicated some prospects for fruitful further investigations.

NOTES

- * The author would like to thank Christine Tappolet, the audience at the “Colloque Jeunes Chercheurs” of the SoPhA and at the SPQ 2014 symposium “La normativité: découverte ou invention,” and two anonymous referees for helpful comments. He would also like to thank the *Groupe de recherche sur la normativité* (GRIN), the *Centre interuniversitaire de recherche sur la science et la technologie* (CIRST) and the *Fonds de recherche – Société et Culture* (FRQ-SC) for doctoral grants.
- ¹ I take the phrase “*formal* axiology” from Olson (2003, p. 413). The term can be seen as denoting the sub-field of axiology, which focuses on the nature of final value—what final value *is*—in contrast to *substantial* axiology, which attempts to determine what final values *are*.
 - ² However, Rabinowicz and Rønnow-Rasmussen (1999, p. 116) note that, rigorously speaking, something may have “internally relational” properties, that is, “properties that it possesses in virtue of its relations to its own *parts*.” As it seems clear that by “relational properties” Moore meant “externally relational properties,” I follow Rabinowicz and Rønnow-Rasmussen’s use and employ the former term in a restricted sense that includes only internally relational properties.
 - ³ The second “error” alluded to by Moore concerns his “principle of organic unities,” to which I will return below.
 - ⁴ In the following, I use the word “thing” in a very broad sense that may include all kinds of ontological entities (like “states of affairs,” “objects,” “tropes,” etc.) when I wish to remain neutral regarding the debate over the proper bearers of value.
 - ⁵ Many of the publications involved in this debate have been usefully anthologized in Rønnow-Rasmussen and Zimmerman (2005).
 - ⁶ For more detailed introductions to FA analyses, see Rabinowicz (2013), Jacobson (2011) and Rabinowicz and Rønnow-Rasmussen (2004).
 - ⁷ Although most FA analysts interpret the fittingness relation between the attitude and its object in *deontic* terms, that is, as a relation that states the type of attitude that evaluators *ought* to adopt or *have reasons* to adopt, there are also some FA analysts who interpret this relationship as a *representational* one, that is, a relation that states the attitude that evaluators are *correct* to adopt, that represents the objects at stake *as they are* (cf. Tappolet 2011).
 - ⁸ Although some find anticipations in the earlier work of Scottish moral sense theorists like Hume and Hutcheson, the historical origin of FA analyses is usually traced back to Brentano (1902) and Ewing (1947). The theory has recently experienced a revival ever since Scanlon (1998) has turned the metaethical community back to it. See Rabinowicz and Rønnow-Rasmussen (2004, pp. 394-400) for a historical survey and Tappolet (2011, n1) for a bibliography.
 - ⁹ There is of course a challenge here, and much of the disagreement among FA analysts has been over which formulation of the fittingness relation can best meet this challenge (as an anonymous referee has remarked, few proponents of FA analyses have in fact offered such naturalistic analyses).
 - ¹⁰ Of course, objections invoking the “wrong kind of reasons” problem have been a recurring worry for FA analyses (cf. Stratton-Lake 2005; Rabinowicz and Rønnow-Rasmussen 2004; Crisp 2000; D’Arms and Jacobson 2000), but some promising solutions to this problem have recently been put forward (e.g. Danielsson and Olson 2007).
 - ¹¹ See Bradley’s (2002, pp. 30-33) defense of the legitimacy of such seemingly strange counterfactual application of the notion of contributive value.
 - ¹² Hurka speaks of “intrinsic” rather than “final” value, but the context of his discussion, which leaves open to discussion the question whether “intrinsic” value can supervene on relational properties, makes clear that the concept he has in mind is the one I call “final value.” Kagan (1998) and most environmental philosophers follow this (often confusing) terminological practice. Also, Hurka speaks of *states*, but his two interpretations, as we will see, can also be applied to other kinds of ontological entities like objects.

- ¹³ I would however be more inclined to say that the final value of the Abraham Lincoln pen is explained by its signatory relation with the Emancipation Proclamation rather than by its usefulness.
- ¹⁴ This distinction is very close to the one formulated by Rønnow-Rasmussen (2002), however, I think that my formulation in terms of relational vs. attitudinal instrumental values better emphasizes the locus of the contrast than his in terms of strong vs. weak instrumental values.
- ¹⁵ The main difference is that states of affairs are *abstract predications* of properties, e.g. “a having p,” whereas tropes are *concrete* (spatiotemporally located) *instantiations* of properties, e.g. “a’s having p,” (Rabinowicz and Rønnow-Rasmussen 2003, pp. 218-219).
- ¹⁶ As Zimmerman accepts the FA analysis of value, simply ignoring Rabinowicz and Rønnow-Rasmussen’s objection is not an option available to him (see Zimmerman 2001a, p. 197; 2001b, p. 40).

REFERENCES

- Bradley, Ben, "Extrinsic Value," *Philosophical Studies*, vol. 91, no. 2, 1998, pp. 109-126.
- Bradley, Ben, "The Value of Endangered Species," *Journal of Value Inquiry*, vol. 35, no. 1, 2001, pp. 43-58.
- Bradley, Ben, "Is Intrinsic Value Conditional?," *Philosophical Studies*, vol. 107, no. 1, 2002, pp. 23-44.
- Brentano, Franz, *The Origin of Our Knowledge of Right and Wrong*, Westminster, A. Constable, 1902.
- Chisholm, Roderick M., *Brentano and intrinsic value*, Cambridge/New York, Cambridge University Press, 1986.
- Crisp, Roger, "Value ... And What Follows by Joel Kupperman New York: Oxford University Press," *Philosophy*, vol. 75, no. 3, 2000, pp. 452-462.
- D'Arms, Justin, and Daniel Jacobson, "Sentiment and Value," *Ethics*, vol. 110, no. 4, 2000, pp. 722-748.
- Danielsson, Sven, and Jonas Olson, "Brentano and the Buck-Passers," *Mind*, vol. 116, no. 463, 2007, pp. 511-522.
- Ewing, Alfred C., *The definition of good*, New York, NY, Macmillan co., 1947.
- Hurka, Thomas, "Two Kinds of Organic Unity," *Journal of Ethics*, vol. 2, no. 4, 1998, pp. 299-320.
- Jacobson, Daniel. "Fitting Attitude Theories of Value," in Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, 2011, URL: <http://plato.stanford.edu/archives/spr2011/entries/Fitting-attitude-theories/>.
- Kagan, Shelly, "Rethinking Intrinsic Value," *Journal of Ethics*, vol. 2, no. 4, 1998, pp. 277-297; cited as reprinted in Rønnow-Rasmussen, Toni, and Michael J. Zimmerman (eds.), *Recent Work on Intrinsic Value*, Dordrecht, Springer, 2005, chap. 9.
- Korsgaard, Christine M., "Two Distinctions in Goodness," *Philosophical Review*, vol. 92, no. 2, 1983, pp. 169-195; cited as reprinted in Rønnow-Rasmussen and Zimmerman (eds.), *Recent Work on Intrinsic Value*, Dordrecht, Springer, 2005, chap. 8.
- Mackie, John L., *Ethics: Inventing Right and Wrong*, Harmondsworth, NY, Penguin Books, 1977.
- McLaughlin, Peter, "On Having a Function and Having a Good," *Analyse and Kritik*, vol. 24, no. 1, 2002, pp. 130-143.
- Moore, George E., *Principia Ethica*. Amherst, N.Y., Prometheus Books, 1988.
- Moore, George E., *Philosophical Studies*, London, Routledge and Kagan Paul, 2000.
- O'Neill, John, "The Varieties of Intrinsic Value," *The Monist*, vol. 75, no. 2, 1992, pp. 119-137.

Olson, Jonas, "Revisiting the Tropic of Value: Reply to Rabinowicz and Rønnow-Rasmussen," *Philosophy and Phenomenological Research*, vol. 67, no. 2, 2003, pp. 412-422.

Olson, Jonas, "Intrinsicism and Conditionalism About Final Value," *Ethical Theory and Moral Practice*, vol. 7, no. 1, 2004, pp. 31-52.

Rabinowicz, Wlodek, "Value, Fitting attitude Account Of," in *International Encyclopedia of Ethics*. Blackwell Publishing Ltd, 2013, URL: <http://onlinelibrary.wiley.com/doi/10.1002/9781444367072.wbiee311/abstract>.

Rabinowicz, Wlodek, and Toni Rønnow-Rasmussen, "A Distinction in Value: Intrinsic and for Its Own Sake," *Proceedings of the Aristotelian Society*, vol. 100, no. 1, 1999, pp. 33-51; cited as reprinted in Rønnow-Rasmussen and Zimmerman (eds.), *Recent Work on Intrinsic Value*, Dordrecht, Springer, 2005, chap. 10.

Rabinowicz, Wlodek, and Toni Rønnow-Rasmussen, "Tropic of Value," *Philosophy and Phenomenological Research*, vol. 66, no. 2, 2003, pp. 389-403; cited as reprinted in Rønnow-Rasmussen and Zimmerman, *Recent Work on Intrinsic Value*, Dordrecht, Springer, 2005, chap. 17.

Rabinowicz, Wlodek, and Toni Rønnow-Rasmussen, "The Strike of the Demon: On Fitting Pro-attitudes and Value," *Ethics*, vol. 114, no. 3, 2004, pp. 391-423.

Rønnow-Rasmussen, Toni, "Instrumental Values - Strong and Weak," *Ethical Theory and Moral Practice*, vol. 5, no. 1, 2002, pp. 23-43.

Rønnow-Rasmussen, Toni, and Michael J. Zimmerman (eds.), *Recent Work on Intrinsic Value*, Dordrecht, Springer, 2005.

Scanlon, Thomas M., *What We Owe to Each Other*, Cambridge, MA, Belknap Press of Harvard University Press, 1998.

Stratton-Lake, Philip, "How to Deal with Evil Demons: Comment on Rabinowicz and Rønnow-Rasmussen," *Ethics*, vol. 115, no. 4, 2005, pp. 788-798.

Tappolet, Christine, "Values and Emotions: Neo-Sentimentalism's Prospects," in Carla Bagnoli (ed.), *Morality and the Emotions*, Oxford, Oxford University Press, 2011.

Zimmerman, Michael J., "Intrinsic Value and Individual Worth," in Dan Egonsson, Jonas Josefsson, Björn Petersson, and Toni Rønnow-Rasmussen (eds.), *Exploring practical philosophy: from action to values*, Aldershot, Ashgate, 2001a, pp. 123-138; cited as reprinted in Rønnow-Rasmussen, Toni, and Michael J. Zimmerman (eds.), *Recent Work on Intrinsic Value*, Dordrecht, Springer, 2005, chap. 15.

Zimmerman, Michael J., *The Nature of Intrinsic Value*, Lanham, MD., Rowman & Littlefield, 2001b.